Supplemental material

Essential histone chaperones collaborate to regulate transcription and chromatin integrity

Olga Viktorovskaya¹, James Chuang^{1,3}, Dhawal Jain², Natalia I. Reim¹, Francheska López-Rivera¹, Magdalena Murawska^{1,4}, Dan Spatt¹, L. Stirling Churchman¹, Peter J. Park², and Fred Winston $^{\rm 1}$

Supplemental Figures S1-S6

Supplemental Tables S1, S2, S4, S5

Supplemental Materials and Methods

Figure S1 (related to Figure 1)

Figure S1 (related to Figure 1). The *spt6-YW* **mutant phenotype is suppressed by** *SPN1* **overexpression.**

- **(A)** Analysis of genetic interactions between spt6, rco1, set2, and chd1 mutations. Strains were grown to saturation in YPD, serially diluted 10-fold, spotted on YPD, and grown at the indicated temperature.
- **(B)** Analysis of effects of SPN1 overexpression on spt6-YW temperature sensitivity. Strains FY3276 and FY3277, transformed with either plasmid FB2701 [SPN1-Myc, CEN, URA3, HIS3] or an empty vector, were grown to saturation in liquid media without uracil, serially diluted 10-fold, spotted on the indicated media, and grown at the indicated temperature.

Figure S2 (related to Figure 2)

Figure S2 (related to Figure 2). Spt6, FACT, Spt5, and Spn1 functionally interact to modulate nucleosome organization *in vivo***.**

(A) Heatmaps of changes to MNase-seq dyad signal in mutants over wild-type for 3086 non-overlapping coding genes aligned by wild-type +1 nucleosome dyad and sorted by gene length, for strains grown either at 30°C or with an 80-minute shift to 37°C.

(B) Average MNase-seq dyad signal over 3,086 non-overlapping verified coding genes aligned by 30°C wild-type +1 nucleosome dyad, for wildtype (FY87), spt6-YW (FY3223), spn1-K192N (FY3272) strains grown at 30°C or with an 80-minute shift to 37°C. Values are the mean of the mean library-size normalized coverage over the genes considered, over at least two replicates.

Figure S3 (related to Figure 3)

Figure S3 (related to Figure 3). The *spt6-YW* **mutation causes altered sense and antisense transcription.**

- **(A)** Quantification of western blots for protein abundance in whole cell lysates of wild-type, spt6-YW, and spt6-1004 strains grown at 30°C or shifted to 37°C for 80 minutes. Spt6, Spn1, and Set2 signal were normalized to a Pgk1 loading control, and histone modification signal was normalized to histone H3. Error bars indicate the mean ± standard error of the replicates shown. Samples from a set2Δ strain were included as a negative control for the histone modifications (data not shown).
- **(B)** Heatmaps of sense and antisense TSS-seq signal in wild-type, spt6-YW, and spt6-1004 strains, grown either at 30°C or shifted to 37°C for 80 minutes. Data are shown for 3,087 non-overlapping verified coding genes aligned by wild-type genic TSS and sorted by length, up to 300 nt 3*′* of the cleavage and polyadenylation site indicated by the dotted line.
- **(C)** Bar plots showing the number of TSS-seq peaks differentially expressed in spt6-1004 versus wild-type, separated by genomic class.
- **(D)** Scatterplots showing changes in sense strand intragenic transcript abundance versus wild-type strains, comparing spt6-1004 to spt6-YW. TSS peaks are colored based on significant upregulation in one, both, or neither mutant.
- **(E)** Distributions of the positions of upregulated intragenic TSSs on the sense (top) and antisense (bottom) strands along the scaled length of transcripts for spt6-YW (orange) and spt6-1004 (green) at 37°C.
- **(F)** Heatmaps of change in antisense TSS-seq and NET-seq signal in spt6-YW versus wild-type at 37°C. Data are shown over 403 genes with a significantly induced antisense TSS peak in spt6-YW at 37°C, aligned by wild-type sense TSS and arranged by the distance from the sense TSS to the spt6-YW-induced antisense TSS.
- **(G)** Northern blots for the SER3 and SRG1 transcripts in yeast cultures grown at 30°C. SNR190 is shown as a loading control.
- **(H)** Quantification of Northern blots for SER3 and SRG1. SER3 and SRG1 signal were normalized to the SNR190 loading control. Error bars indicate the mean ± standard error of the replicates shown.

Figure S4 (related to Figure 4)

Figure S4 (related to Figure 4). The *pob3-E154K* **and** *spt5-QS* **suppressors do not restore the Spt6-Spn1 interaction.**

- **(A)** Western blots showing the levels of V5-Spn1, Rpb1, Spt5, Spt16 and Spt6 in V5-Spn1 immunoprecipitation (IP) samples and their corresponding inputs, detected with respective antibodies (Table S5).
- **(B)** Quantification of V5-Spn1 co-IP experiments. Error bars indicate the mean *±* standard error of the relative co-IP signal normalized to V5-Spn1 pull-down signal in the replicates shown.
- **(C)** Quantification of protein abundance from western blots of the inputs for V5-Spn1 co-IP experiments. Error bars indicate the mean *±* standard error of the relative western blot signal from the replicates shown.
- **(D)** Western blots showing the levels of Rpb3-FLAG, Spn1, Spt5, Spt16 and Spt6 in Rpb3-FLAG IP samples and their corresponding inputs, detected with respective antibodies (Table S5).
- **(E)** As in (B), but for Rpb3-FLAG.
- **(F)** As in (C), but for Rpb3-FLAG.

(A) Scatterplot of Spn1 ChIP enrichment versus Rpb1 enrichment for 5,091 verified coding genes in wild-type cells. Enrichment values are the relative log₂ enrichment of IP over input.

(B) Heatmaps of the change in Spn1 enrichment in mutants over wild-type for 3,087 non-overlapping verified coding genes aligned by TSS and arranged top to bottom by decreasing wild-type Rpb1 enrichment.

Figure S6 (related to Figure 6)

Figure S6 (related to Figure 6). Mutant FACT suppresses the Spt6-Spn1 defect by restoring the balance between FACT and Spt6 on chromatin.

- **(A)** Scatterplots showing changes in Rpb1-normalized Spt16 ChIP enrichment in mutants over wild-type versus changes in Rpb1-normalized Spt6 ChIP enrichment, for 5,091 verified coding genes.
- **(B)** Average Rpb1 ChIP enrichment over 3,087 non-overlapping verified coding genes aligned by TSS in wild-type (FY3276) and spt6-YW (FY3277) strains, grown either at 30°C or with an 80-minute shift to 37°C. For each gene, the spike-in normalized ratio of IP over input signal is standardized to the mean and standard deviation of the 30°C wild-type signal over the region. The solid line and shading are the mean and 95% confidence interval of the mean standard score over the genes considered from two to five replicates.

Table S1 (related to Figure 1). Identification and genetic verification of the suppressors for *spt6-YW* **temperature-sensitive phenotype.**

¹ Phenotype linkage was assessed by the analysis of the Ts⁺ phenotypes in the progeny segregating the indicated suppressor allele in the spt6-YW backgrounds.

² Complementation was examined by transforming the indicated spt6-YW suppressor strains with a plasmid expressing a respective wild-type allele for the suppressor gene (see Table S5) followed by the phenotype analysis of the transformants. Phenotype rescue was scored when transformants had a Ts⁻ phenotype (complete rescue), similar to the spt6-YW parent strains, or had an intermediate Ts[±] phenotype (partial rescue). "ND" stands for "not determined".

³ In case of the non-essential genes (*RCO1, SET2, and CHD1*), the null alleles were shown to suppress the Ts- phenotype of spt6-YW (Table S2, Figure S1A). In case of FACT and Spt5, the identified mutations were introduced to SPT5 and POB3 genes, reconstituting the spt5-QS and pob3-E154K alleles, and genetic suppression of *spt6-YW* by either allele was confirmed (Figure 1C). "ND" stands for "not determined".

 4 In several cases the chromosome XVI disomy was accompanied by other disomies, including chr.I, III or XI. None of the disomies other than the chr.XVI disomv were specifically associated with the Ts^* suppressor phenotype.

Table S2 (related to Figure 1). The phenotypes for different mutant combinations used to determine genetic interactions.

The phenotypes were analyzed based on the yeast growth under indicated conditions in comparison to the wild-type strain using the spot test assay and scored after two days of incubation, unless indicated otherwise. The number of "+" indicates growth of each of the 10-fold serially diluted cultures. "-" indicates a very weak growth, if any. "ND" stands for "not determined".

 1 Spt phenotype was scored as growth on synthetic media without lysine due to suppression of the *lys2-128δ* allele.

² Phleomycin was added to the YPD media at concentration of 13 μ g/ml; plates were incubated at 30°C for three days before scoring.

 3 Hydroxyurea was supplemented to YPD at 150 mM final concentration; plates were incubated at 30°C.

⁴ *spn1-K192N chd1Δ* mutant was scored for growth at 37°C after 3 days of incubation, indicating weak suppression of Ts- phenotype.

Table S4. The list of strains used in the study.

Table S5. Reagents.

Supplemental Materials and Methods.

Data Analysis

Genome builds and annotations

The genome builds used were *S. cerevisiae* R64-2-1 {Engel, 2014 #374} and *S. pombe* ASM294v2 {Wood, 2002 #375}. *S. cerevisiae* transcript coordinates were generated from TIFseq {Pelechano, 2013 #376} and TSS-seq data, as previously described {Doris, 2018 #300}.

TSS-seq library processing

Removal of adapter sequences and random hexamer sequences from the 3′ end of the read and 3′ quality trimming were performed using cutadapt {Martin, 2011 #377}. The random hexamer molecular barcode on the 5′ end of the read was then removed and processed using a Python script. Reads were aligned to the combined *S. cerevisiae* and *S. pombe* reference genomes using Tophat2 {Kim, 2013 #378} without a reference transcriptome, and uniquely mapping alignments were selected using SAMtools {Li, 2009 #379}. Alignments mapping to the same location as another alignment with the same molecular barcode were identified as PCR duplicates and removed using a Python script. Coverage of the 5′-most base, corresponding to the TSS, was extracted using bedtools genomecov {Quinlan, 2010 #380}. Due to high variability in the proportion of *S. pombe* spike-in alignments among the libraries for certain conditions, coverage was normalized to the total number of alignments uniquely mapping to the *S. cerevisiae* genome. The quality of raw, cleaned, non-aligning, and uniquely aligning nonduplicate reads was assessed using FastQC {Andrews, 2014 #381}.

TSS-seq peak calling

TSS peak calling was performed using 1D watershed segmentation followed by filtering for reproducibility by the Irreproducible Discovery Rate method (IDR=0.1) {Li, 2011 #382}, as previously described in {Doris, 2018 #300}, except using the maximum signal within a putative peak to estimate the probability of the peak being generated by noise rather than the sum of the signal within the peak. To allow for direct comparison between *spt6-YW* and *spt6-1004,* unified TSS peak sets were generated by using bedtools multiinter to combine peaks called in wildtype, *spt6-YW,* and *spt6-1004* at 30°C or 37°C.

TSS-seq differential expression analysis

For TSS-seq differential expression analyses, counts of alignments overlapping the unified TSS peak sets described above were used as the input to differential expression analysis by DESeq2 {Love, 2014 #383}, with a null hypothesis of no change between conditions and a false discovery rate of 0.1.

Classification of TSS peaks into genomic categories

TSS peak classification was performed as described {Doris, 2018 #300}, except the summit of intragenic and antisense peaks were used to determine overlap with open/closed reading frames or transcripts. In brief, a genic region was defined for each gene using its transcript and open/closed reading frame annotations. TSS peaks were classified as genic if they overlapped a genic region on the same strand. TSS peaks were classified as intragenic if they were not classified as genic and their summit overlapped an open or closed reading frame on the same strand. TSS peaks were classified as antisense if they were not classified as genic and their

summit overlapped a transcript on the opposite strand. TSS peaks not overlapping genic regions, transcripts, or reading frames were classified as intragenic.

NET-seq library processing

Removal of adapter sequences from the 3′ end of the read and 3′ quality trimming were performed using cutadapt {Martin, 2011 #377}. Reads were aligned to the *S. cerevisiae* genome using Tophat2 without a reference transcriptome (kim2013), and uniquely mapping alignments were selected using SAMtools {Li, 2009 #379}. Coverage of the 5′-most base of the read, corresponding to the 3′-most base of the nascent RNA and the active site of elongating RNA polymerase, was extracted using bedtools genomecov {Quinlan, 2010 #380}, and normalized to the total number of uniquely mapping alignments. The quality of raw, cleaned, non-aligning, and uniquely aligning reads was assessed using FastQC {Andrews, 2014 #381}.

ChIP-seq library processing

Reads were demultiplexed using fastq-multx {Aronesty, 2103 #384}*,* allowing one mismatch to the index sequence and A-tail. Cutadapt {Martin, 2011 #377} was then used to remove index sequences and low-quality base calls from the 3′ end of the read. Reads were aligned to the combined *S. cerevisiae* and *S. pombe* genome using Bowtie 2 {Langmead, 2012 #385}, and alignments with a mapping quality of at least 5 were selected using SAMtools {Li, 2009 #379}. The median fragment size estimated by MACS2 {Zhang, 2008 #386} cross-correlation over all samples of a factor was used to generate coverage of fragments and fragment midpoints by extending alignments to the median fragment size or by shifting the 5′ end of alignments by half the median fragment size, respectively. The quality of raw, cleaned, non-aligning, and uniquely aligning reads was assessed using FastQC {Andrews, 2014 #381}.

ChIP-seq normalization

For ChIP-seq coverage from IP samples, spike-in normalization was accomplished by scaling coverage proportionally to the normalization factor $N_{input, spike-in} / N_{IP, spike-in} * N_{input, experimental}$, where N_{IP, spike-in} is the number of *S. pombe* alignments in the IP sample, N_{input, spike-in} is the number of *S.* pombe alignments in the corresponding input sample, and N_{input, experimental} is the number of *S*. *cerevisiae* alignments in the input sample. Coverage from input samples was normalized to N_{input, experimental}. Relative estimates of the total abundance of each ChIP target on chromatin were also obtained by multiplying the normalization factor with the number of *S. cerevisiae* alignments in an IP sample. Coverage of the relative ratio of IP over input was obtained by first smoothing normalized IP and input fragment midpoint coverage using a Gaussian with 20 bp bandwidth, and then taking the ratio. Coverage of the relative ratio of one factor to another (e.g. Spn1 over Rpb1) was obtained as follows: For each factor, coverage of IP over input in each sample was standardized using the genome-wide mean and standard deviation over all samples, weighted such that each condition had equal contribution. Standardized coverage of the normalizing factor was then subtracted from the matched coverage of the factor to be normalized.

ChIP-seq differential occupancy analysis

For differential occupancy analyses of single factors over verified coding genes, IP and input fragment midpoints overlapping the transcript annotation of these genes were counted using bedtools {Quinlan, 2010 #380}. These counts were used to perform a differential occupancy analysis using DEseq2 {Love, 2014 #383}, at a false discovery rate of 0.1. The design formula used was a generalized linear model with variables for sample type (IP or input), condition

(strain and temperature), and the interaction of sample type with condition. Fold changes were extracted from the coefficients of the interaction of sample type with condition, and represent the change in IP signal between conditions, corrected for change in input signal. To normalize to the spike-in control, size factors obtained from *S. pombe* counts over peaks called with MACS2 {Zhang, 2008 #386} and IDR {Li, 2011 #382} were used for each sample.

MNase-seq library processing and quantification

Paired-end reads were demultiplexed using fastq-multx {Aronesty, 2103 #384} allowing one mismatch to the barcode. Filtering for the barcode on read 2 and 3′ quality trimming were performed with cutadapt {Martin, 2011 #377}. Reads were aligned to the combined *S. cerevisiae* and *S. pombe* genome using Bowtie 1 {Langmead, 2009 #402}, and correctly paired alignments were selected using SAMtools {Li, 2009 #379}. Coverage of nucleosome protection and nucleosome dyads were extracted using bedtools {Quinlan, 2010 #380} and shell scripts to get the entire fragment or the midpoint of the fragment, respectively. Smoothed nucleosome dyad coverage was generated by smoothing dyad coverage was generated by smoothing dyad coverage with a Gaussian kernel of 20 bp bandwidth. Due to differences in the proportion of *S. pombe* DNA added between sequencing runs, coverage was normalized to the total number of correctly paired *S. cerevisiae* fragments. The quality of raw, cleaned, non-aligning, and correctly paired reads was assessed using FastQC {Andrews, 2014 #381}. Nucleosome regions for each condition were called using DANPOS2 {Chen, 2013 #365}. Nucleosome 'fuzziness' was calculated for each nucleosome region in each sample by taking the standard deviation of nucleosome dyad positions in the region.