Extended Experimental Procedures

High-resolution mapping reveals a conserved, widespread, dynamic meiotically regulated mRNA methylation program

Schraga Schwartz, Sudeep D. Agarwala, Maxwell R. Mumbach, Marko Jovanovic, Philipp Mertins, Alexander Shishkin, Yuval Tabach, Tarjei Mikkelsen, Rahul Satija, Gary Ruvkun, Steven A. Carr, Eric S. Lander, Gerald R. Fink, Aviv Regev

Strain genotypes are shown in **Table S4**. To induce synchronous meiotic entry, cells were pre-selected on 1% yeast extract, 2% peptone, 3% glycerol, 2% agar for 24 hours at 30°C, grown for 24 hr in 1% yeast extract, 2% peptone, 4% dextrose at 30°C, diluted in BYTA (1% yeast extract, 2% tryptone, 1% potassium acetate, 50 mM potassium phthalate) to OD600 = 0.2 and grown for another 16 hr at 30°C, 300 rpm. Cells were then washed once with water and re-suspended in SPO (0.3% potassium acetate) at OD600 = 2.0 and incubated at 30°C at 190 rpm. Cells were isolated from SPO at the indicated times, collected by centrifugation, re-suspended in pre-warmed 1% yeast extract, 2% peptone, 2% dextrose and incubated at 30°C at 190 rpm. For ectopic expression of the MIS complex, cells were collected after 150 minutes of mitotic growth in the presence of cupric sulfate, induced with 100 µM CuSO4 in rich complete synthetic media (2% glucose). All replicates reported in the manuscript are biological, done on different meiotic inductions in different yeast cultures at different times.

RNA preparation for m^{6}A-seq

RNA was extracted from cells using a standard hot acid phenol protocol. Briefly, cell pellets were resuspended in equal volumes of acid phenol:chloroform 5:1 pH 4.3-4.7 (Sigma), buffer AE (50 mM sodium acetate, 10mM EDTA 1% SDS) and glass beads. This mixture was vortexed for 15 minutes, followed by a 15 minute incubation at 65°C. Samples were centrifuged for 10 minutes (12,000g, 4°C); the supernatant isolated, re-extracted with phenol:chloroform::5:1, and precipitated with sodium acetate and isopropanol. Enrichment of polyadenylated RNA (polyA+ RNA) from total RNA was performed using Oligo(dT) dynabeads (Invitrogen) according to the manufacturer's protocol. The mRNA was chemically fragmented into ~80-nt-long fragments using RNA fragmentation reagent (Ambion). The sample was then subjected to Turbo DNAse treatment (Ambion), followed by a phenol-chloroform extraction, and resuspension in 20 µl of IPP buffer (150 mM NaCl, 0.1% NP-40, 10 mM Tris-HCl, pH 7.5).

m^{6}A-seq

25 µl of protein-G magnetic beads were washed and resuspended in 200 µl of IPP buffer, and tumbled with 3 µl of affinity purified anti-m^6A polyclonal antibody (Synaptic Systems) at room temperature for 30 minutes. Following 2 washes in IPP buffer, RNA was added to the antibody-bead mixture, and incubated for 2 h at 4°C. The RNA was then washed twice in 200 µl of IPP buffer, twice in low-salt IPP buffer (50 mM NaCl, 0.1% NP-40, 10 mM Tris-HCl, pH 7.5), and twice in high-salt IPP buffer (500 mM NaCl, 0.1% NP-40, 10 mM Tris-HCl, pH 7.5), and eluted in 30 µl RLT (Qiagen). To purify the RNA, 20 µl MyOne Silane Dynabeads (Life Technologies) were washed in 100 µl RLT, resuspended in 30 µl RLT, and added to the eluted RNA. 60 µl 100% ethanol was added to the mixture, the mixture attached to the magnet and the supernantant discarded. Following two washes in 100 µl of 70% ethanol, the RNA was eluted from the beads in 160 µl IPP buffer. Eluted RNA was subjected to an additional round of IP, by re-incubating it with protein-A magnetic beads coupled to anti-m^6A antibody, followed by washes, elution from the protein-A beads and purification as above, followed by elution from the MyOne silane dynabeads in 10 µl H_20.

Library preparation

Strand-specific m^6A RNA-seq libraries were generated as described in (Engreitz et al., 2013). Briefly, RNA was first subjected to FastAP Thermosensitive Alkaline Phosphatase (Thermo Scientific), followed by a 3' ligation of an RNA adapter using T4 ligase (New England Biolabs). Ligated RNA was reverse transcribed using AffinityScript Multiple Temperature Reverse Transcriptase (Agilent), and the cDNA was subjected to a 3' ligation with a second adapter using T4 ligase. The single-stranded cDNA product was then amplified for 9-14 cycles in a PCR reaction. Libraries were sequenced on Illumina Miseq, HiSeq 2000 and/or HiSeq 2500 platforms generating paired end reads (25 or 30 bp from each end, depending on the platform).

Read alignment

Reads were initially mapped against all SK1 ribosomal RNA (rRNA) sequences using Bowtie (version 0.12.7), and all reads aligning to the rRNA were discarded. All remaining reads were aligned against the SK1 genome using Tophat (version 1.4.1). Parameters used were '--max-multihits 1 –prefilter-multihits' and '–transcriptome-index', for which we assigned a pre-indexed version of the SK1 transcriptome. An in-house script was then used to cast all reads aligning to the genome upon the SK1 transcriptome. Only reads fully matching a transcript structure, as defined by the transcriptome annotation, were retained. Such reads were computationally extended in transcriptome space from the beginning of the first read to the end of its mate, and coverage in transcriptome-space was calculated for each nucleotide across all transcripts. Since the SK1 transcriptome annotations for the most part do not define 3' UTRs, we used an in house script based on high-throughput strand-specific sequencing of our input samples at three hours in meiosis, to conservatively extend transcript annotations into 5' and 3' UTRs, as long as (**1**) they were covered by reads from the same strand, (**2**) coverage within these regions was within 3-fold of the median coverage within the genes, and (**3**) the total length of the extension (UTR) did not exceed 500 nt (a filter applied to eliminate merging of adjacent genes in the same orientation). This set of annotations is available as **Supplemental Table S5**. For the analysis of methylation sites on rRNA, all reads were directly mapped to yeast rRNA sequences.

Detection of putative m^{6}A sites

Putative m^{6}A sites were identified using a 3 step-approach, consisting of: (**1**) Examination of the IP sample, to identify regions within genes in the IP samples that were enriched in comparison to background gene levels; (2) Comparison of IP sample with input sample, to ensure that these regions were not enriched in corresponding input samples; and (**3**) Comparison across multiple replicates of IP and input samples, to ensure that the identified regions were reproducibly

enriched across different replicates and reproducibly depleted across the negative control *ime4*Δ/Δ samples. Below we provide a detailed description of these steps:

(1) Peak detection within genes. To search for enriched peaks in the m^6A IP samples, we scanned each gene using sliding windows of 100 nucleotides with 50 nucleotides overlap. Each window was assigned a score, corresponding to the fold of the mean coverage in the window over the median coverage across the gene. Windows with scores greater than 4 (i.e. greater than 4-fold enrichment) and with a mean coverage >10 reads were retained. Overlapping windows were merged together, and for each disjoint set of windows in transcriptome space we recorded its start, end, and peak position (corresponding to the position with the maximal coverage across the window), and its peak score (corresponding to the fold-change of enrichment in coverage over the median gene level).

(2) Ensuring that peaks were absent in input. We repeated the same steps for the input sample. We eliminated from all subsequent analysis all windows that were detected in both step 1 and in step 2.

(3) Comparison of multiple samples. To identify peaks that were robustly present across multiple replicates of wild-type samples but absent in the *ime4*Δ/Δ negative control samples, we applied the following strategy. We first merged the coordinates of all windows from all samples passing step 1 and 2, to define a set of disjoint windows passing these filters in at least one of the samples. For each such window, we recalculated the peak start, end, and peak position and peak score across each of the samples using the approach in step 1. To identify IME4 dependent peaks in *S. cerevisiae*, we required that (**i**) the peak is detected in at least two of three replicates, (**ii**) the distribution of peak scores across the wild-type samples is significantly different from its counterpart in the *ime4*Δ/Δ samples (Student's t-test, $P\ value < 0.05$), and (**iii**) the mean peak

score across the WT replicates is at least 3-fold greater than the mean peak score in the *ime4*Δ/Δ samples. To identify *IME4* dependent peaks in *S. mikatae*, we required that (**i**) the peak is present in the sample expressing *IME4*, and (**ii**) the peak score in this sample is at least 4-fold higher than in the sample not expressing *IME4*.

Control datasets

Randomized location dataset: For each identified peak passing all three filters, we generated a matching control counterpart by selecting a random position within the same gene. This set of 1,308 control peaks are the basis for the motif enrichment analysis in Figure 1C, and the plot in Figure 1D.

Non-methylated RGAC sites dataset: For the analyses presented in Figure 2B-2D examining relative position of an RGAC site within a gene, secondary structure, and the machine learning analysis, we generated a panel of 10,346 RGAC sites within the same transcripts as their methylated counterparts, and for which we did not find evidence of being methylated based on the m6A-seq experiments.

***In vitro* transcription**

150-nt long fragments tiling across 17 genes containing IME4-independent peaks were designed, such that they start at every fourth position in the gene (i.e. from position 1 to 150, 5 to 154, etc.), share identical sequences at their 5' and 3' ends, to facilitate amplification by PCR, and contain a T7 promoter at their 5' end. The oligonucleotide pool was synthesized on a CustomArray 12K Microarray using a B3 Synthesizer (CustomArray, Bothell, WA) as recommended by the manufacturer, solubilized by immersion in 28-30% ammonium hydroxide solution (Sigma, St. Louis, MO) at 60°C for 7 hours, dried, resuspended in TE buffer, purified

using a QIAquick PCR Purification Kit (Qiagen, Germantown, MD), amplified for 26 cycles, and purified using a QIAquick PCR purification kit. *In vitro* transcription reactions were performed using a MAXIscript T7 kit (Life Technologies) followed by a rigorous DNAse treatment using Turbo DNAse (Life Technologies).

De novo motif search

Motifs enriched within m^6A peaks compared with control peaks were identified by counting the occurrence of 4–6-nucleotide k-mers in a 50 nt window centered around the peak position in the IP sample, compared to the randomized control group. The total number of k-mers of each length within every group was counted and the ratio between their prevalence was used to calculate the fold change between the two groups. Fisher's exact test was used to evaluate the differences in the prevalence of each k-mer between the groups. Analysis was limited to motifs enriched more than twofold and with an associated Bonferonni-corrected P value < 0.05. To generate motifs, k-mers were clustered together, using the approach in ref. (Dominissini et al., 2012).

Methylation classifier

To train a classifier of methylatability, we made use of the non-methylated RGAC sites dataset (described above) and a stringent set of 832 methylated sites (from the catalogue of 1,308 sites) which were within 10 nt of the closest consensus site; The latter sites were each assigned to the nearest consensus site. Sites were assigned a binary methylation state value (methylated versus non-methylated). For each site, we then recorded its absolute distances from the transcription start and end sites, its relative position within the gene (on a scale of 0 to 1), the nucleotide composition of -4 to +5 with respect to the methylated position (each position was binarized into four dummy variables, denoting whether it was an 'A', 'C', 'G' or 'T'). We also calculated the

predicted secondary structure strength of each site (see below). One or more of these features were then integrated into a logistic regression machine learning scheme that learns methylation states on the basis of the different features. A 10-fold cross validation scheme was designed, and implemented using the RWeka package (Hornik et al., 2009) in R.

Secondary structure analysis

RNAfold (Hofacker et al., 1994) was used to assess local secondary structure stabilities within 50-nt windows centered around the detected m^6A and random peaks, yielding minimum free energy (MFE) of the folded sequence. To control for GC content effects, each sequence was shuffled 50 times and MFE scores were calculated for those sequences as well. Sequences were then assigned a Z score, indicative of the extent to which a sequence is more stably folded compared to the shuffled controls.

Calculation of fraction of methylated sites from all sites containing a consensus: For this analysis, presented in the discussion, we compiled a dataset of all RGAC sites in yeast transcripts that (1) were expressed above the 30^{th} percentile, and (2) were longer than 500 nt in length. The latter filter was imposed as in shorter genes our method of detecting peaks, that relies on background estimation, is of more limited power. 44,654 such sites were identified, of which 1,195 were methylated (2.6%). Limiting this analysis only to genes with an extended consensus sequence, 443/4,710 sites were methylated (9.4%).

mRNA expression analysis

To estimate expression levels, reads were aligned against the *S. cerevisiae* SK1 transcriptome using RSEM (version 1.2.1) with default parameters. For robust comparison between different

samples, we used TMM normalization of the RSEM read counts as implemented by the edgeR package (Robinson et al., 2010) in R.

GO analysis

All Gene Ontology analyses presented were based on sets of genes obtained from GOslim (http://www.geneontology.org/). To obtain a more fine-grained resolution on meiosis related gene categories, this dataset was supplemented with sets of genes from the full GO annotation including all terms matching the keywords "meiosis" or "sporulation".

m^{6}A-seq *in S. mikatae*

Due to the incomplete assembly of the *S. mikatae* genome, we aligned reads directly against a collection of *S. mikatae* ORFs with flanking intergenic sequence, downloaded from http://www.broadinstitute.org/annotation/fungi/comp_yeasts/downloads.html. These ORFs were concatenated together, interspersed with a stretch of Ns and are available as **Supplemental Table S6**. To compare transcript positions between *S. cerevisiae* and *S. mikatae*, we generated pairwise sequence alignments between *S. cerevisiae* genes and *S. mikatae* genes sharing identical gene names using NW-align (as above), and used in house scripts to map position coordinates from one organism to another based on this alignment.

Meiotic spreads and immunofluorescence

Spread meiotic nuclei were prepared using the method of FALK. Briefly, cells were spheroplasted at 37°C in solution 1 (2% potassium acetate, 1M sorbitol, 10 mM dithiothreitol, 130 µg/ml zymolyase 100T). Spheroplasting was stopped using ice-cold solution 2 (100 mM MES [pH 6.4], 1 mM EDTA, 0.5 mM MgCl2, 1 M sorbitol). Slides were washed in 0.4% Photoflo (Kodak). Fifteen microliters of spheroplast suspension were briefly prefixed on a glass

slide with 40 µl fixative (4% paraformaldehyde, 3.4% sucrose) and lysed with 80 µl 1% lipsol. After further addition of 80 µl fixative, spheroplasts were spread using a glass rod. Slides were then dried in the fume hood overnight. Slides were blocked with blocking buffer (0.2% gelatine, 0.5% BSA in PBS). Fob1 was detected using the rabbit 5778 anti-Fob1 antibody at a 1:250 dilution in blocking buffer and an anti-rabbit Alexa568 antibody (Invitrogen) at 1:200. Ime4-3xMYC and Mum2-3xHA was visualized using the mouse 4A6 anti-MYC antibody (Millipore) and mouse HA.11 antibody, respectively, at 1:100 (Covance) and an anti-mouse Alexa488 antibody (Invitrogen) at 1:250.

qPCR

Total RNA was obtained by standard phenol:chloroform:isoamyl alcohol extraction. cDNA was generated using random hexamers or strand-specific primers and the Qiagen QuantiTect Reverse Transcription Kit (cat no. 205314). Transcript abundance was quantified using reagents from Applied Biosystems and the ABI 7500 real-time PCR system. Primer sequences are: SLZ1 (Forward: gcttgaaagattgtgtatggatgaa, Reverse: cgcttgttgcatggttattcc), MUM2 (Forward: ttcatccccaccaacagtca, Reverse: ggcatcgtttctttcaccagat), IME4 (Forward: gcggcctggctggttt; Reverse: ccatttcgtaaatgcaatttcct), and ACT1 (Forward: ctccaccactgctgaaagagaa; Reverse: ccaaggcgacgtaacatagtttt).

TLC analysis

TLC analysis was carried out as in (Zhong et al., 2008); mRNA was purified with the Dynabeads mRNA purification system (Invitrogen) and analyzed on cellulose plates (20cm x 20cm) from EMD.

RNA affinity chromatography and mass spectrometry

Three RNA baits, each comprising a ~120 nt long sequence containing only a single adenine were *in vitro* transcribed from dsDNA templates. Sequences of the templates are:

IVT1:GCGTAATACGACTCACTATAGGGGCTCTCGTCCCTCTGGTCGTTGCGCGCCTG
CGTGGCTCTTGGTTCGCTTCTCTGG**A**CTCTCGTCTTGGCGCGTGCGTCGTTTGTTTCTT
GGGCTGTGGTCTCGCGGCCGC,
IVT2:GCGTAATACGACTCACTATAGGGGCTTGTTGCTGCCCTTGTTGCCTCCTTGGCC
CCTCCCGTTTGGGCTGCCTGTGCTGGG**A**CTTTTCCCCTTTCTGCGTTGGGGGTCCGCG
TTGTCTTGTGGTCGTGCTGTGTGCGGCCGC,

IVT3:GCGTAATACGACTCACTATAGGGGTCTGTTGGTTCGGTGCTCTCGGTTGTGCGC
GGCCGTTGGGCCCGGCCTTCGTCGGG**A**CTTGTCGCGCTCCTCCTTCTTTCTCGCGTCG
CTTGCGGTTTTGCTTCTTTGGCTTGCGCGGCCGC

In vitro transcription was performed in a volume of 20 µl using the MEGAshortscript™ T7 Kit (Life Technologies), with a nucleotide mix containing 150 nmol of GTP, CTP, and UTP, 75 nmol of Biotin-16-UTP (Roche), and 150 nmol of either ATP or N6-Methyl-ATP (TriLink BioTechnologies). The baits were mixed at equimolar ratios. Poly(A) tailed mRNA was extracted from meiotic cultures in prophase arrest from either WT (*ndt80*Δ/Δ, SAy841) or IME4Δ/Δ (*ime4*Δ/Δ *ndt80*Δ/Δ—SAy996) backgrounds. 25 ml of meiotic culture was harvested 3 hours after meiotic induction in the presence of protease inhibitors (Complete protease inhibitors, Roche). Cells were washed once with 1M Tris-HCl, pH 7.5 and snap-frozen. Frozen pellets were resuspended in lysis buffer (150 mM NaCl, 50 mM Tris pH 7.5, 1% NP40, 10 mM PMSF, Complete mini protease inhibitors (Roche) at 2× concentration) and glass-bead homogenized three times for 5 minutes at 4°C. Debris was pelleted by centrifugation for 10 minutes, and supernatant was incubated with either 2 micrograms of biotin-labeled *in vitro*

transcribed probes immobilized on streptavidin C1 beads (Life Technologies) or with the poly(A) mRNA immobilized on Dynabeads Oligo(dT) beads (Life Technologies) with head-over-tail rotation overnight in the presence of 50 units of SUPERasin (Life Technologies). Beads were washed 5 times in lysis buffer and boiled in reducing loading buffer. Proteins were precipitated by adding -20C cold actone to the eluate (acetone to eluate ratio 7:1) and overnight incubation at -20C. The proteins were pelleted by centrifugation at 20000xg for 15min at 4C. The supernatant was discarded and the pellet was left to dry by evaporation. The protein pellet was reconstituted in 50 ul lysis buffer (8M Urea, 75 mM NaCl, 50 mMTris/HCl pH 8.0, 1 mM EDTA) and protein concentrations were determined by BCA assay (Pierce, Rockford, IL). 5 to 10 ug total protein per affinity enrichment were obtained condition. Disulfide bonds were reduced with 5 mM dithiothreitol and cysteines were subsequently alkylated with 10 mM iodoacetamide. Samples were diluted 1:4 with 50 mM Tris/HCl (pH 8.0) and sequencing grade modified trypsin (Promega, Madison, WI; V5113) was added in an enzyme-to-substrate ratio of 1:50. After 16 h of digestion, samples were acidified with 1% formic acid (final concentration). Tryptic peptides were desalted on C18 StageTips according to (Rappsilber et al., 2007) and evaporated to dryness in a vacuum concentrator.

Desalted peptides were labeled with the iTRAQ reagent according to the manufacturer's instructions (AB Sciex, Foster City, CA) and to (Mertins et al., 2012). Briefly, for 10ug of peptide 0.33 units of iTRAQ reagent was used. Peptides were dissolved in 10 ul of 0.5M TEAB pH 8.5 solution and the iTRAQ reagent was added in 25 ul of ethanol. After 1 h incubation the reaction was stopped with 50 mM Tris/HCl (pH 8.0). Differentially labeled peptides were mixed and subsequently desalted on C18 StageTips (Rappsilber et al., 2007) and evaporated to dryness in a vacuum concentrator. Peptides were reconstituted in 20 ul 3% MeCN/0.1% formic acid. LC-MS/MS analysis was performed as described in

(Mertins et al., 2013). All mass spectra were processed using the Spectrum Mill software package v4.0 beta (Agilent Technologies, Santa Clara, CA) according to (Mertins et al., 2012).

Statistical analysis

All statistical analyses and visualizations were performed in R: Sequence logos were prepared using the SeqLogo package (Bembom, 2011), heatmaps were generated using the gplots package (Warnes, 2012), and the majority of the remaining plots were generated using the ggplot2 package (Wickham, 2009).

Supplemental Figure Legends

Supplemental Figure S1 related to Figure 1: **(A)** Scheme of m^6A-seq. Polyadenylated RNA is fragmented into ~80 nt long fragments. Methylated RNA is enriched using an anti-m^6A antibody. Libraries are produced by ligating an adapter to the 3' terminus of the RNA, followed by reverse transcription and ligation of an additional adapter to the 3' end of the cDNA. Thus, the entire fragmented RNA is captured and subjected to sequencing. **(B)** Read coverage along three genes containing MIS-independent sites at their 3' ends. Coverage is shown for m6A-Seq (red, IP) and control (blue, input) for three sets of experiments: one replicate of WT yeast and of IME4Δ/Δ, both at prophase arrest, and for the in-vitro transcribed RNA tiling across the entire genes. The precise recapitulation of the peaks in the in-vitro control demonstrates that these sites are sequence-dependent false positives. **(C)** Sequences enriched at MIS-independent peaks. For this analysis, a conservative dataset of 807 MIS-independent sites was generated comprising all sites that (1) had Peak Scores >2 in at least 2 of 3 replicates in both WT and IME4Δ/Δ strains, and (2) had insignificant differences between the Peak Score distribution of WT and IME4Δ/Δ strains (P value >0.2). K-mers enriched in these sequences compared to randomly selected control regions from the same genes were identified (left panel) and clustered (right panel) as for the MIS dependent peaks. (D) Read coverage across IME1, IME2 and IME4 genes in m6A-Seq (red, IP) and control (blue, input) experiments performed at prophase arrest in WT yeast, and in an IME4Δ/Δ background. Only one site was detected as an enriched position, in the IME2 gene, highlighted in yellow.

Supplemental Figure S2 related to Figure 2: Proportion of sites within 250 nt of 3' end of gene, as a function of peak score. The collection of methylated sites was divided into 6 equally sized bins based on peak score. Proportions are also shown for the random control group (light blue). Error bars depict the standard error of the mean (SEM).

Supplemental Figure S3 related to Figure 3: Histogram displaying the relative position within a gene of all methylated sites in *S. mikatae*. A strong bias towards the 3' of the gene is evident.

Supplemental Figure S4 related to Figure 4: Analysis of 3 temporal clusters. **(A)** Proportion of methylated sites harboring 'A' at position -4, 'G' at position -2, and U at position +4 in each of the three identified clusters. **(B)** Boxplots depicting the relative location within a gene of methylation sites in each of the three clusters. **(C)** RNA-seq derived expression values (Y axis) for each of the three MIS components, across the full meiotic time course (X axis). Values were normalized by the maximal values for each gene across the time course, to allow plotting on the same scale.

Supplemental Figure S5 related to Figure 5: (A) Representative images of immunofluorescence of whole cell IF, showing localization of epitope-tagged Ime4 (SAy914), Mum2 (SAy1235) or Slz1 (SAy1254) (first column), nucleolar marker Fob1 (second column), and DNA (DAPI—left column). The compilation (fourth column) shows DNA in blue, the MIS component in green, and Fob1 in red. **(B)** Expression (by qPCR) of *IME1* and *NDT80*

normalized by actin. Error bars: standard deviation of three replicates. **(C)** Whole cell IF images as in A, performed across an 8 hour meiotic time course.

Supplemental Figure S6 related to Figure 6: Percent cells undergoing meiotic divisions (Y axis) as assayed by DAPI staining along a meiotic time course (X axis) in four indicated strains.

Supplemental Figure S7 related to Figure 7: (A) Mean half-lives of transcripts in the sustained, intermediate, and peaked cluster, based on data from (Miller et al., 2011) measured under vegetative conditions. Error bars: SEM. **(B)** qPCR based quantification of *IME1* transcript (normalized by actin) in diploid wild-type (SAy821), *ime4*Δ/Δ (SAy771), *ime4*-cat/-cat (SAy1086) after induction into nutrient starvation medium (SPO). Error bars: standard deviation (SD), n=3. **(C)** *IME1* transcript accumulation in a wild-type haploid strain (AH223), a haploid strain inducing the MIS complex from the CUP promoter (SAy1411) or a haploid strain inducing the catalytic mutant form the of MIS complex from the CUP promoter (SAy1412). Cells were collected 30 minutes after induction with 100μM $CuSO_4$ in SPO medium to induce transcription from the *CUP1* promoter. Error bars: SD (n=3). **(D)** *IME1* transcript accumulation in wild-type (SAy821), *ime4*Δ/Δ (SAy771), *mum2*Δ/Δ (SAy1195), *slz1*Δ/Δ (SAy1206) after induction into SPO. Error bars: SD (n=3). **(E)** Heatmap of mRNA expression levels (purple: low; yellow: high) of two clusters of genes (of 8 total clusters) exhibiting differential behavior between WT and IME4Δ/Δ timecourses (the full dataset is available as **Supplemental Table S7**). The bottom cluster comprised genes induced at the latest stages of the time-course in wild-type cells, but failing to do so in the *ime4*Δ/Δ counterparts. This cluster was highly enriched for genes involved in GO terms pertaining to the two meiotic divisions and spore wall formation, consistent with previous findings that entry into the two meiotic divisions is delayed in the absence of *IME4* (Agarwala et al., 2012). The top cluster comprised a set of genes showing strong induction in wild-type cells 3 hours after induction of sporulation - corresponding to meiotic prophase -- and failing to do so in ime4Δ/Δ cells. GO analysis of genes in this cluster revealed a strong enrichment for double strand break formation and synaptonemal complex formation, two functions for which genes harboring methylation sites are also enriched. **(F)** Quantification of synaptonemal complex formation on meiotic chromosomes as assayed to immunofluorescence of Zip1 on nuclear spreads. Cells were either characterized as showing no Zip1 (green), partial polymerization (red) or full polymerization (blue) in either the wild-type (SAy841--top panel) or IME4 deletion (SAy966--bottom panel) strains. **(G)** Southern analysis yeast chromosome III as cells progress through meiosis. Meiotic double-strand break formation was assayed in wild-type (SAy821), ime4-cat/-cat (SAy1086) or IME4 deletion (SAy771) during meiosis. **(H)** Impact of methylations on steady state levels of mRNA, 3 hours after induction of meiosis. For each of the 8 strains in which a single methylation site was point-mutated and 2 WT controls, expression was assessed based on RNA-seq libraries. For each gene, a box-plot is shown displaying the distribution of expression values across all 9 strains in which the gene had not been mutated. The blue point depicts the expression levels for the respective mutated strains. In the box plot, the bottom and top whiskers extend to the minimum and maximum values, respectively.

Supplemental Table Legends

Supplemental Table S1. Set of 1308 methylated sites. The table records the genomic position of the identified peaks, the distance from the peak to the nearest RGAC consensus site, whether the genes in which the peak resides is annotated as a meiosis gene, the temporal cluster of the peak (based on the analysis in Figure 4), the log2 transformed peak score, an annotation of whether the peak is in cluster 1 or in cluster 5 (based on the analysis in Supplemental Figure S7E), and the translation efficiency of the gene in meiotic prophase, based on the ribosomal profiling dataset published in (Brar et al., 2012).

Supplemental Table S2. Set of 635 peaks identified in *S. mikatae*. The position indicated for the peaks is with respect to the concatenated set of *S. mikatae* peaks, provides as Table S6.

Supplemental Table S3. Mass-spectrometry based fold-changes, quantifying the enriched association of proteins with methylated baits (either in-vitro synthesized in the presence of N6-methyl-ATP, or poly(A) selected from WT yeast) compared to non-methylated counterparts (in-vitro transcribed with ATP, or poly(A) selected from IME4Δ/Δ strain). The number of unique peptides, and the number of quantified spectra are also indicated.

Supplemental Table S4. Strains used in this study. All strains are of the *S. cerevisiae* SK1 background unless otherwise noted.

Supplemental Table S5. An annotation of the meiotic yeast transcriptome, used for this study. cdsStart and cdsEnd refer to the beginning and end of the reading frame, whereas txStart and txEnd denote the beginning and end of the transcript.

Supplemental Table S6. The concatenated *S. mikatae* transcriptome, against which all reads were aligned, and based on which peaks were called in Table S3.

Supplemental Table S7. Gene expression values for all genes that (1) showed a minimal TMM-normalized FPKM values of 10 in >2 conditions, and (2) Showed a two-fold or greater change between corresponding WT and IME4Δ/Δ expression in >2 condition across the time course. Genes were clustered into 8 clusters using k-means. Cluster 1 and cluster 5 are presented in Supplemental Figure S7E.

Supplemental References

Agarwala, S., Blitzblau, H., Hochwagen, A., and Fink, G. (2012). RNA methylation by the MIS complex regulates a cell fate decision in yeast. PLoS genetics *8*.

Bembom, O. (2011). seqLogo: Sequence logos for DNA sequence alignments.

Brar, G.A., Yassour, M., Friedman, N., Regev, A., Ingolia, N.T., and Weissman, J.S. (2012). High-resolution view of the yeast meiotic program revealed by ribosome profiling. Science (New York, NY *335*, 552-557.

Dominissini, D., Moshitch-Moshkovitz, S., Schwartz, S., Salmon-Divon, M., Ungar, L., Osenberg, S., Cesarkas, K., Jacob-Hirsch, J., Amariglio, N., Kupiec, M.*, et al.* (2012). Topology of the human and mouse m6A RNA methylomes revealed by m6A-seq. Nature *485*, 201-206.

Engreitz, J.M., Pandya-Jones, A., McDonel, P., Shishkin, A., Sirokman, K., Surka, C., Kadri, S., Xing, J., Goren, A., Lander, E.S.*, et al.* (2013). The Xist lncRNA exploits three-dimensional genome architecture to spread across the X chromosome. Science (New York, NY *341*, 1237973.

Hofacker, I., Fontana, W., Stadler, P., Bonhoeffer, S., Tacker, M., and Schuster, P. (1994). Fast Folding and Comparison of RNA Secondary Structures. Monatshefte f Chemie *125*, 167-188.

Hornik, K., Buchta, C., and Zeileis, A. (2009). Open-Source Machine Learning: R Meets Weka. Computational Statistics *24*, 225-232.

Mertins, P., Qiao, J.W., Patel, J., Udeshi, N.D., Clauser, K.R., Mani, D.R., Burgess, M.W., Gillette, M.A., Jaffe, J.D., and Carr, S.A. (2013). Integrated proteomic analysis of post-translational modifications by serial enrichment. Nat Methods *10*, 634-637.

Mertins, P., Udeshi, N.D., Clauser, K.R., Mani, D.R., Patel, J., Ong, S.E., Jaffe, J.D., and Carr, S.A. (2012). iTRAQ labeling is superior to mTRAQ for quantitative global proteomics and phosphoproteomics. Molecular & cellular proteomics : MCP *11*, M111 014423.

Miller, C., Schwalb, B., Maier, K., Schulz, D., Dumcke, S., Zacher, B., Mayer, A., Sydow, J., Marcinowski, L., Dolken, L.*, et al.* (2011). Dynamic transcriptome analysis measures rates of mRNA synthesis and decay in yeast. Molecular systems biology *7*, 458.

Rappsilber, J., Mann, M., and Ishihama, Y. (2007). Protocol for micro-purification, enrichment, pre-fractionation and storage of peptides for proteomics using StageTips. Nature protocols *2*, 1896-1906.

Robinson, M.D., McCarthy, D.J., and Smyth, G.K. (2010). edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics (Oxford, England) *26*, 139-140.

Warnes, G.R. (2012). gplots: Various R programming tools for plotting data.

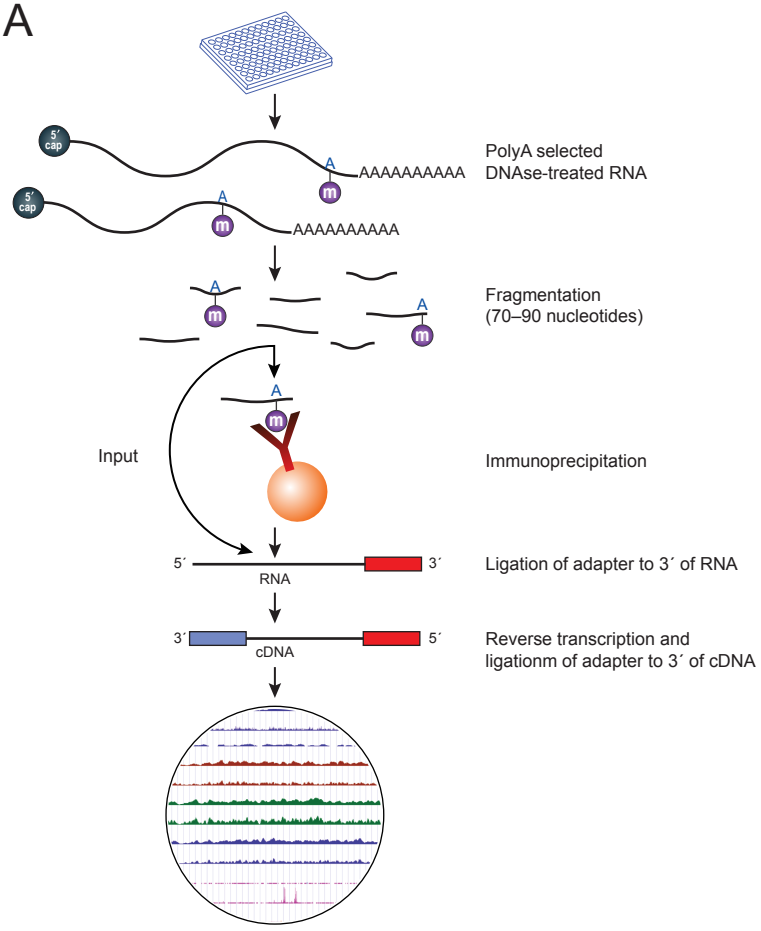
Wickham, H. (2009). ggplot2: elegant graphics for data analysis. Springer New York.

Zhong, S., Li, H., Bodi, Z., Button, J., Vespa, L., Herzog, M., and Fray, R.G. (2008). MTA is an Arabidopsis messenger RNA adenosine methylase and interacts with a homolog of a sex-specific splicing factor. The Plant cell *20*, 1278-1288.

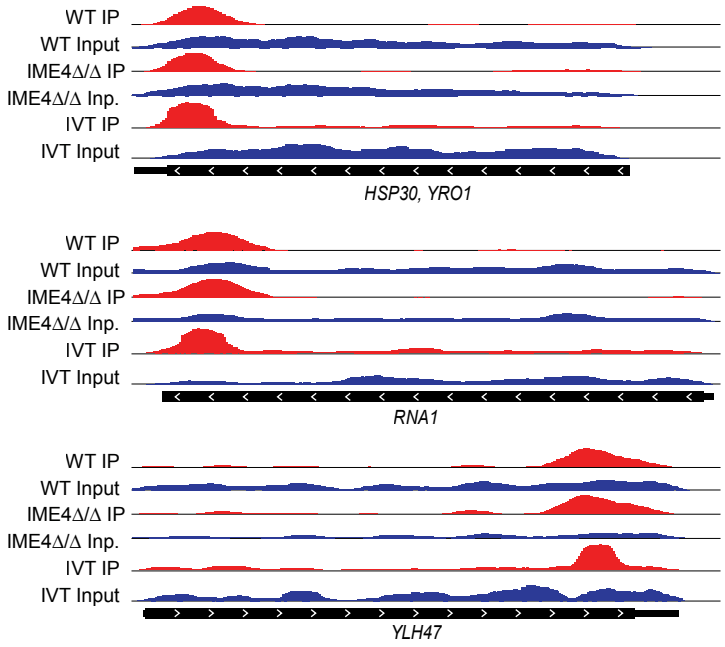
Supplemental Figure S1

SUPP. FIG. 1

A



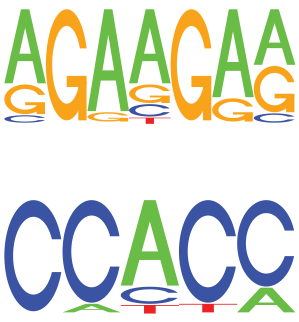
B



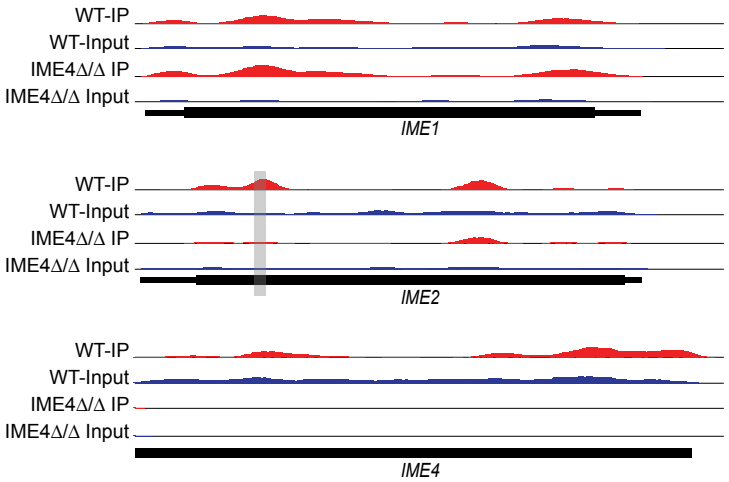
C

Sequence	Fold change	*P* value (Bonferroni corrected)
AGGAAG	5.4	1.76×10^{-11}
CCACCA	5.2	3.71×10^{-5}
GAGGAG	4.7	5.19×10^{-4}
GATGAG	4.7	5.19×10^{-4}
GGAGGA	4.5	2.44×10^{-3}
AGACGA	3.8	2.32×10^{-3}
AAGAAG	3.8	1.51×10^{-23}
GAAGAA	3.6	2.83×10^{-23}
AGAAGG	3.6	1.68×10^{-4}
AGAGGA	3.6	3.28×10^{-8}
GAGGAA	3.5	1.38×10^{-7}
AAGACG	3.5	1.93×10^{-3}
AAGAGG	3.4	3.15×10^{-7}
GGAAGA	3.4	6.41×10^{-8}
GAGGA	3.2	1.31×10^{-17}
GAAGAG	3.1	1.89×10^{-8}
AGAAGA	3.1	3.73×10^{-17}
GAAGGA	3.1	8.53×10^{-3}
GAAGA	2.9	1.06×10^{-37}
CGAAGA	2.9	1.35×10^{-3}
GGAAG	2.8	8.07×10^{-11}
AGGAC	2.8	6.02×10^{-5}
GACGA	2.8	1.09×10^{-7}
CCACC	2.7	1.26×10^{-3}
AGACG	2.6	1.59×10^{-3}
GGAGG	2.6	5.42×10^{-3}
AGAAG	2.5	1.61×10^{-20}
AGAGG	2.5	3.90×10^{-7}
CACCA	2.5	5.78×10^{-5}

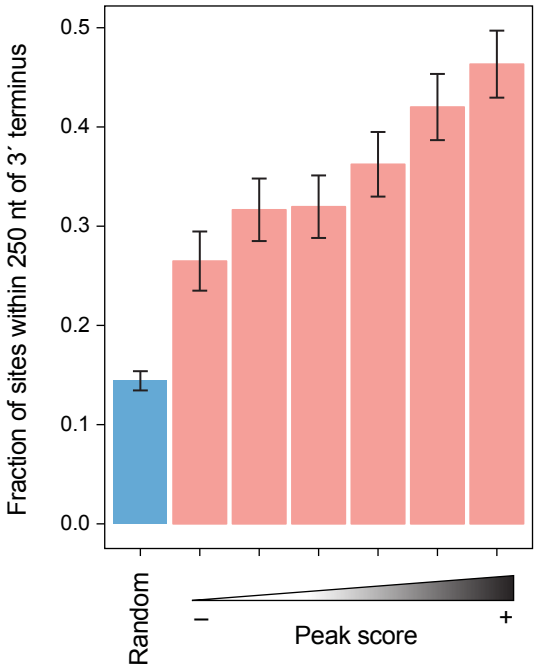
Common motifs



D

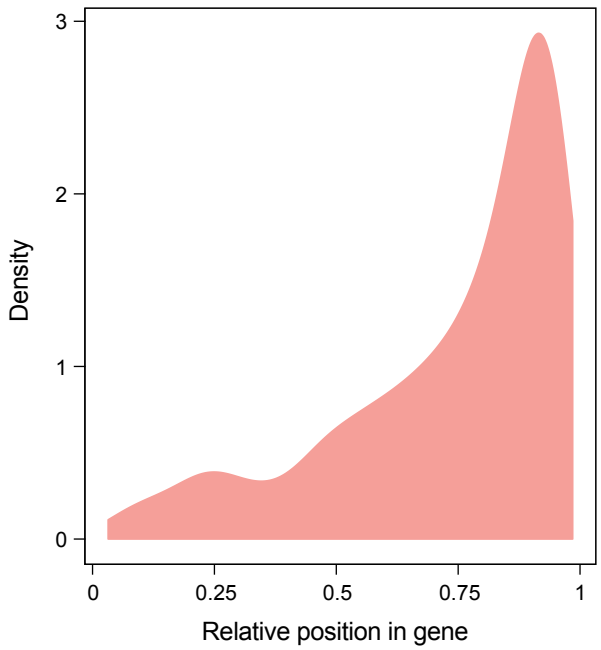


SUPP. FIG. 2

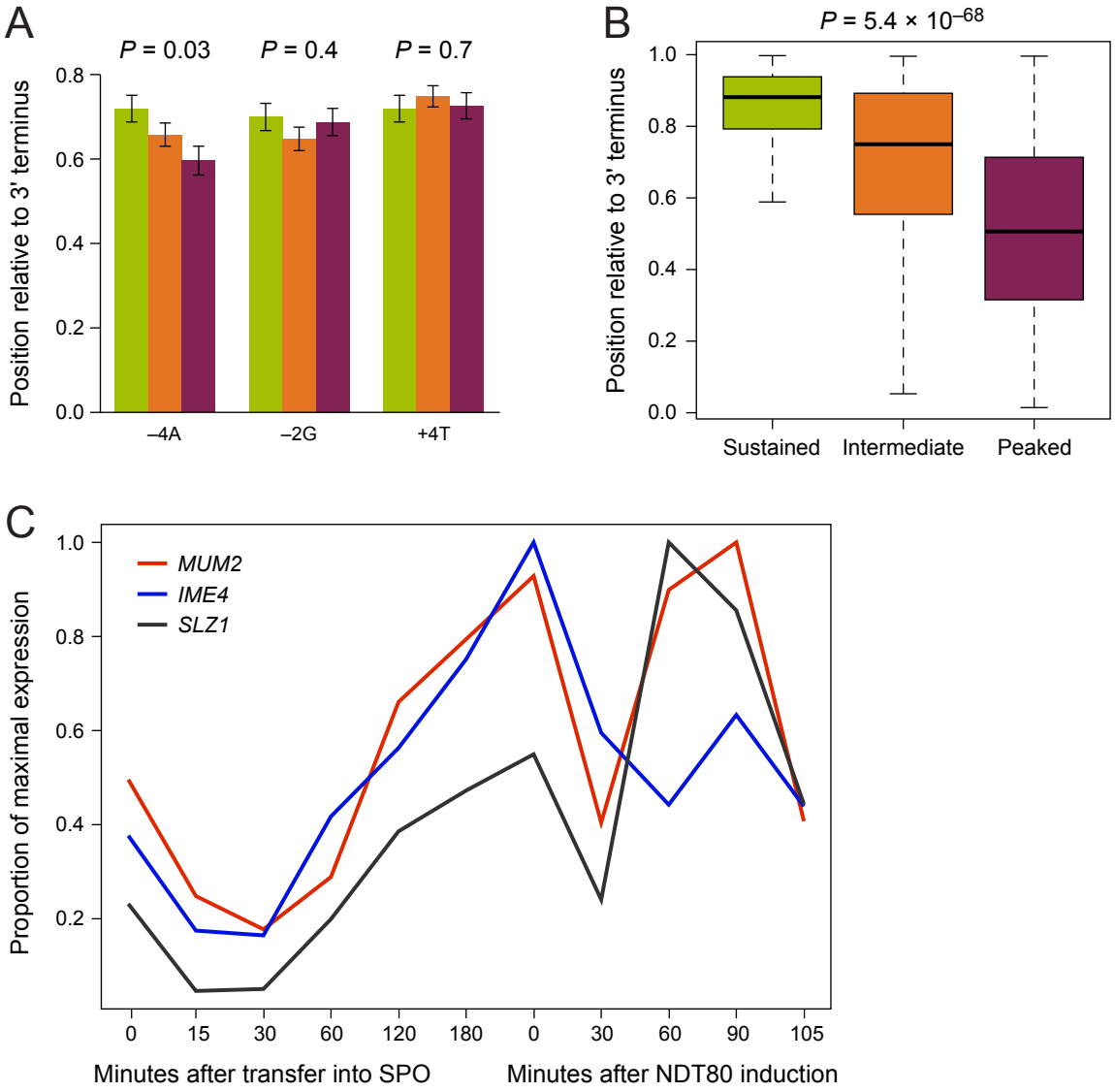

Fraction of sites within 250 nt of 3´ terminus
0.0
0.1
0.2
0.3
0.4
0.5
Random
–
+
Peak score

Supplemental Figure S2

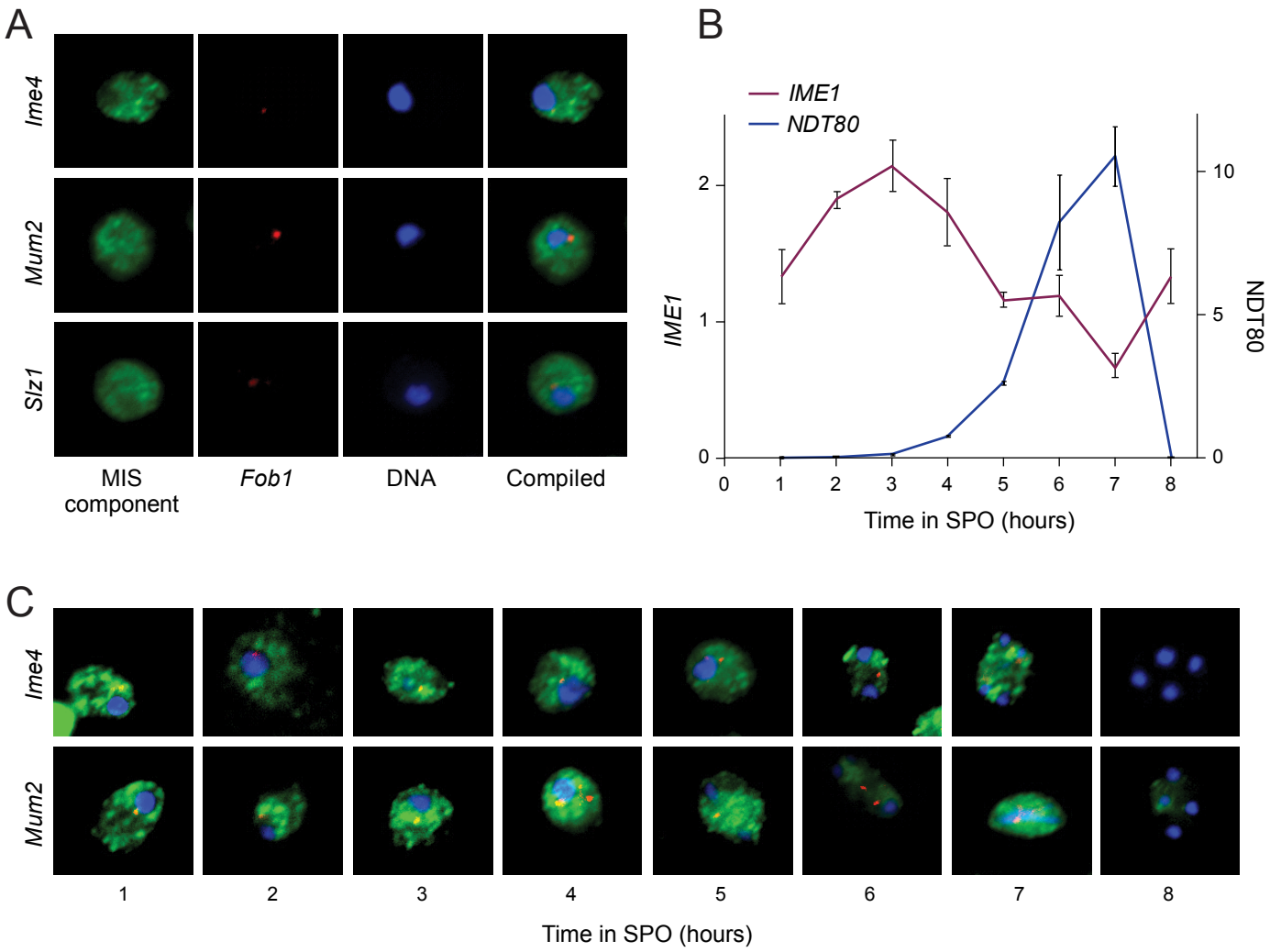
SUPP. FIG. 3



SUPP. FIG. 4

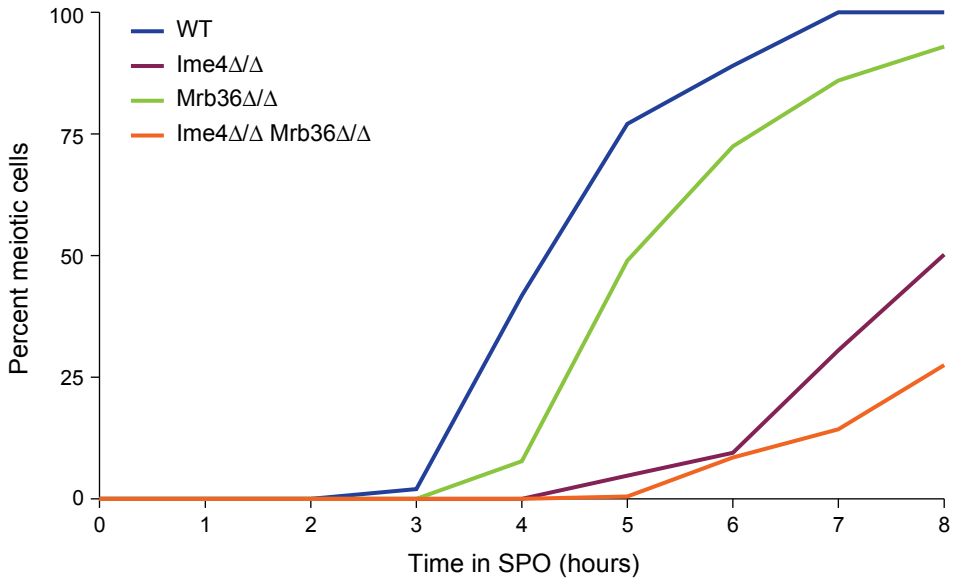

A
P = 0.03
P = 0.4
P = 0.7
Position relative to 3' terminus
0.0
0.2
0.4
0.6
0.8
–4A
–2G
+4T
B
P = 5.4 × 10–68
Position relative to 3' terminus
0.0
0.2
0.4
0.6
0.8
1.0
Sustained
Intermediate
Peaked
C
MUM2
IME4
SLZ1
Proportion of maximal expression
0.2
0.4
0.6
0.8
1.0
0
15
30
60
120
180
0
30
60
90
105
Minutes after transfer into SPO
Minutes after NDT80 induction

SUPP. FIG. 5


A
Ime4
Mum2
Slz1
MIS component
Fob1
DNA
Compiled
B
IME1
NDT80
IME1
NDT80
Time in SPO (hours)
C
Ime4
Mum2
1
2
3
4
5
6
7
8
Time in SPO (hours)

Supplemental Figure S6

SUPP. FIG. 6



Supplemental Figure S7

SUPP. FIG. 7