Meiotic cells counteract programmed retrotransposon activation via RNA-binding translational repressor assemblies

AUTHOR LIST AND AFFILIATIONS

Raphaelle Laureau¹, Annie Dyatel¹, Gizem Dursuk¹, Samantha Brown¹, Hannah Adeoye¹, Jia-Xing Yue², Matteo De Chiara², Anthony Harris³, Elçin Ünal³, Gianni Liti², Ian R. Adams⁴, Luke E. Berchowitz¹,⁵.*

¹ Department of Genetics and Development, Hammer Health Sciences Center, Columbia University Irving Medical Center, New York, NY 10032, USA.
² Université Côte d'Azur, CNRS, Inserm, IRCAN, Nice 06107, France.
³ Department of Molecular and Cell Biology, University of California, Berkeley, Berkeley, CA 94720, USA.
⁴ MRC Human Genetics Unit, MRC Institute of Genetics and Molecular Medicine, University of Edinburgh EH4 2XU, Edinburgh, United Kingdom.
⁵ Lead Contact

*Correspondance: leb2210@cumc.columbia.edu

Luke E. Berchowitz, Ph.D.
701 W 168th St. Hammer Health Sciences Building Room 1520
Columbia University Irving Medical Center
New York NY, 10032
Phone: 212-305-7003
SUMMARY

Retrotransposon proliferation poses a threat to germline integrity. While retrotransposons must be activated in developing germ cells in order to survive and propagate, how they are selectively activated in the context of meiosis is unclear. We demonstrate that transcriptional activation of Ty3/Gypsy retrotransposons and host defense are controlled by master meiotic regulators. We show that budding yeast Ty3/Gypsy co-opts binding sites of the essential meiotic transcription factor Ndt80 upstream of the integration site, thereby tightly linking its transcriptional activation to meiotic progression. We also elucidate how yeast cells thwart Ty3/Gypsy proliferation by blocking translation of the retrotransposon mRNA using amyloid-like assemblies of the RNA-binding protein Rim4. In mammals, several inactive Ty3/Gypsy elements are undergoing domestication. We show that mammals utilize equivalent master meiotic regulators (Stra8, Mybl1, Dazl) to regulate Ty3/Gypsy-derived genes in developing gametes. Our findings inform how genes that are evolving from retrotransposons can build upon existing regulatory networks during domestication.
INTRODUCTION

The relationship between organisms and transposable elements is complex and fluctuating. For the organism, transposable elements can act as a reservoir for novel genes and drive genomic plasticity, but they are also key drivers of genome instability. Insertions can rearrange regulatory networks and alter gene expression patterns (Chuong et al., 2017; Mita and Boeke, 2016; Rebollo et al., 2012). More dangerously, insertion of transposable elements can disrupt genetic information, give rise to heterologies, and can potentially provide sites for ectopic recombination (Warren et al., 2015). For the element, if it is to survive and proliferate, it must do so with minimal detriment and possible benefit to the host (Cosby et al., 2019; Haig, 2016).

Retrotransposons are a class of transposable elements that propagate through an RNA intermediate via a copy and paste mechanism. This mechanism preserves the original element while generating DNA copies which are free to invade other loci. Retrotransposon proliferation has seemingly no intrinsic limit and poses an acute burden to the host genome. This burden is reflected in the repetitive fraction of genomes, which is primarily composed of retrotransposon-derived sequences. In this study we focus on understanding the biology underpinning proliferation of a class of Long-Terminal-Repeats retrotransposons (LTR-retrotransposons), Ty3/Gypsy, during meiosis.

LTR-retrotransposons are divided into four superfamilies based on structure and phylogeny: Ty1/Copia, Ty3/Gypsy (both represented in all eukaryotic lineages), Bel-Pao (found in metazoans but not mammals), and ERVs (Endogenous RetroViruses, found only in vertebrates) (Wicker et al., 2007). Ty3/Gypsy and ERVs are closely related because many vertebrate retroviruses likely arose from Ty3/Gypsy retrotransposons that acquired the envelope gene from other viruses (Hayward, 2017). Although no active Ty3/Gypsy retrotransposons are found in mammalian genomes, several sequences derived from Ty3/Gypsy are present, many of which have been domesticated for cellular functions (Volff, 2009).

The Ty3/Gypsy life cycle is much like that of a retrovirus except it occurs entirely within the cell. It is transcribed into an mRNA which encodes 2 polyproteins: Gag3 and Gag3-Pol3 (Figure 1A). Gag3 is the precursor to the capsid (CA) and nucleocapsid (NC) proteins and Pol3 is the precursor to a protease (PR), a reverse transcriptase-RNaseH (RT-RH), and an integrase (IN). In most cases only Gag3 polyprotein is translated, while Gag3-Pol3 is only produced by a frameshift that occurs in 5-10% of translation events (Farabaugh et al., 1993) (Figure 1A). Gag3 assembles together with a Ty3 mRNA dimer and Gag3-Pol3 into a remarkable structure within the cell – the virus-like particle (VLP) – inside which the precursor proproteins mature into their final products and the Ty3 mRNA is reverse transcribed into cDNA (Garfinkel et al., 1985). This cDNA copy is bound to integrase proteins and can
enter the nucleus via the nuclear pore to integrate in the genome. Ty3 integration events require binding to polymerase III and thus frequently occur in close proximity to tRNA genes (Qi et al., 2012).

For a retrotransposon to increase its copy number and proliferate through generations it must to do so in cells that will ultimately become gametes. For this reason, retrotransposons tend to be expressed in the developing germline, including during meiosis (Crichton et al., 2014; Ribeiro-dos-Santos et al., 1997). However, as retrotransposon mobilization in these cells can cause heritable mutations and trans-generational genetic instability, organisms have evolved elegant multi-layered mechanisms which act at various stages of the retrotransposon cycle to protect their germlines (Crichton et al., 2014).

For example, transcription of retrotransposons can be repressed by DNA methylation, chromatin modifications, and transcriptional interference (for review (Crichton et al., 2014; Yang and Wang, 2016)). Post-transcriptional pathways such as miRNA, siRNA, piRNA target retrotransposon mRNA for degradation (Bao and Yan, 2012; Dechaud et al., 2019; Yang and Wang, 2016) and RNA modification enzymes such as cytidine deaminases prevent reverse transcription of the retrotransposon mRNA into cDNA (Orecchini et al., 2018). Post-translational regulatory mechanisms act in germ cells to target retrotransposon proteins for premature degradation (MacLennan et al., 2017). Beyond the obvious long-term consequences of uncontrolled retrotransposon proliferation, dysregulation of retrotransposon control can cause acute failures in gamete development. Disruption of defense mechanisms in mice or Drosophila leads to sterility (Barau et al., 2016; Bourc'his and Bestor, 2004; Wang et al., 2018) and hyperactivation of retrotransposons causes fetal oocyte attrition in mice (Malki et al., 2014).

Here, we show that Ty3/Gypsy elements can exploit the genomic environment at their integration sites to ensure transcriptional activation during the meiotic divisions. Our results illustrate a mechanism by which retrotransposons can restrict their expression to a developmental context in which activation is beneficial. We also identify an opposing mechanism by which cells defend their genome against meiotic retrotransposition. Developing budding yeast Saccharomyces cerevisiae gametes utilize amyloid-like RNA-binding protein assemblies to repress Ty3 translation thereby halting the proliferation cycle. We find that both the co-opting of meiotic regulators by Ty3/Gypsy elements and their interaction with assembly-forming RNA-binding proteins are strategies which are also employed by Gypsy-derived genes in mammals. Our findings highlight the likely hundreds of millions of years of perpetual interactions among retrotransposons, host RNA-binding proteins, and genomes.
RESULTS

Ty3 is expressed in coordination with meiotic progression

We first determined whether retrotransposons are transcriptionally active during *S. cerevisiae* meiosis (strains used in this study are described in Table S1). We performed RNAseq at hourly time points from cells undergoing synchronous meiosis using the *NDT80* block-release system (Benjamin et al., 2003; Carlile and Amon, 2008). There is only one full-length copy of Ty3 in the SK1 strain, and we observed that this locus is transcribed in the sense orientation specifically during the meiotic divisions (Figure 1A, B). We also observed an antisense RNA at the onset of meiosis that declines as meiosis proceeds to almost undetectable levels. These data led us to hypothesize that transcription of the full-length Ty3 is repressed in mitotic cells but is activated during the meiotic divisions.

Only a full-length retrotransposon mRNA can complete the steps necessary for proliferation. While our RNAseq data imply that the Ty3 locus is transcribed in meiosis, we sought to differentiate whether this was indicative of bona-fide full-length transcripts or of cryptic inactive transcripts that span the locus. Using Northern blot analysis, we observed several Ty3 RNA isoforms (Figure 1C), consistent in size with the various isoforms spanning the locus in the RNAseq (Figure 1B), including a prominent band of the size of a full-length transcript (5 kb). All of the RNA isoforms we observed were specific to Ty3 because none of them were produced in cells lacking Ty3 (ty3Δ) (Figure 1C, Figure S1A, B).

Using probe tiling, we were able to identify this top band as a full-length mRNA, containing both GAG3 and POL3 coding sequences (Figure S1C) and thus the potential capacity to propagate. Strikingly, full-length Ty3 mRNA was expressed and reached maximal levels at the first meiotic division (Figure 1C, D).

We next determined whether meiotic transcriptional activation of Ty3 occurs in other *S. cerevisiae* isolates. Because rapid, efficient, and synchronous meiosis is a feature that is not shared by most lab strains, we decided to investigate a collection of wild diploid isolates (Peter et al., 2018) for sporulation efficiency and Ty3 activation (Table S2). These strains harbor varying Ty3 copy numbers. We found that full-length Ty3 transcription switches on at the time of meiotic divisions in at least one isolate, OS673, which harbors two full-length copies of Ty3 (Figure 1E-G, Table S2). This strain was isolated from a beetle in Hungary (Peter et al., 2018) and is of a completely different genetic background than SK1 (~ 51,400 SNPs), allowing us to rule out the possibility that Ty3 activation in meiosis is a peculiarity that appeared during lab domestication. From these data we conclude that Ty3 retrotransposons have evolved to activate, and potentially proliferate, during meiosis.
**Ty3 expression is controlled by a critical meiotic transcription factor**

How is Ty3 transcription activated specifically at the onset of meiosis I? The transition from meiotic prophase to the first meiotic division is governed by the meiosis-specific transcription factor Ndt80 (Xu et al., 1995). Accordingly, we asked whether Ty3 is under the control of Ndt80. If Ty3 is a Ndt80 target, it would not be transcribed in meiotic cells deprived of Ndt80. Furthermore, Ty3 transcription should be selectively activated upon ectopic Ndt80 activity. We found that cells in which NDT80 is not induced – and are thus arrested in meiotic prophase – do not express Ty3 mRNA (Figure S2A, B). Additionally, in meiotic cells where NDT80 expression is delayed, Ty3 transcription exhibits a corresponding delay. We observed that Ty3 full-length transcripts accumulate from 15-30 min before metaphase I, and persist until exit from anaphase II (Figure 2A, B). These results imply that Ty3 is a target of Ndt80.

In further support of this idea, Ty3 is integrated downstream of two Ndt80 binding sites (79 bp and 269 bp) (Figure 2C). To investigate whether Ty3 meiotic transcription activation is governed by its environment of integration rather than elements within its 5' LTR, we asked if Ty3 is still transcribed in cells harboring a promoter-gene cassette (HIS3MX) between the 5' LTR and the first proximal Ndt80 binding site (Figure S3A). We found that Ty3 was no longer transcribed when we separated it from its upstream genetic context (Figure S3B). To test whether Ndt80 directly activates Ty3, we mutated the two Ndt80 binding sites upstream of Ty3 (CACAAA to AAAAAA) and assessed the effect of these mutations on meiotic Ty3 transcription (Figure 2C). We found that point mutation of these sequences specifically abolished full-length meiotic Ty3 transcription (Figure 2D). To assess whether Ndt80 binds upstream of Ty3 during meiotic progression, we conducted a meiotic time course analysis of Ndt80 binding by ChIP-seq. We analyzed Ndt80 binding at 0, 3, 5, and 6 hours in meiosis. We chose these time points because Ndt80 levels are highest ~ 5-6 hours in meiosis (Figure 2E). In support of the hypothesis that Ty3 is transcriptionally activated by Ndt80, we find that Ndt80 binds upstream of Ty3 as meiosis progresses (Figure 2F). We conclude that Ty3 co-opts a host regulatory circuit to coordinate its activation with meiotic division and we predict that other retroelements could have evolved a similar strategy to acquire germline-specific activation.

The Ndt80 protein sequence as well as its binding site is conserved among yeasts from Saccharomyces to Candida (Nocedal et al., 2017). We thus sought to determine whether Ty3 control by Ndt80 is conserved among other Saccharomyces. We assessed the presence of Ndt80 binding sites within and upstream of full-length Ty3 elements in seven S. cerevisiae and five S. paradoxus genomes for which all Ty elements have been mapped and annotated (Yue et al., 2017) (spanning approximately 5 million years of evolution (Shen et al., 2018)) (Figure 3A). Among twenty full-length Ty3 elements, we found that thirteen harbor one or more Ndt80 binding sites within 1 kb upstream of their integration.
sites, including five cases where the Ndt80 binding site is less than 250 bp upstream the integration site (Figure 3A). Within the 1 kb regions upstream of Ty3 elements, we observed several that harbor multiple Ndt80 binding sites. In total, we observed nineteen Ndt80 binding sites within the 20 total kb which is significantly higher than what would be expected by chance (Z-score 2.7; p < 0.01).

Because Ndt80 binding sites are abundant (about 8,000/genome), we needed to ascertain whether the distance of a Ndt80 binding site from a locus is predictive of whether that locus is truly activated by Ndt80. Ndt80-responsive genes are well defined in SK1 meiosis (Chu et al., 1998). We first compared the enrichment of Ndt80 binding sites upstream of Ndt80 target genes and non-target genes, and found a peak of enrichment within the -50 to -250 bp region upstream of the transcription start sites (TSS) in target genes (Figure 3B). We next asked whether the association between Ty3 elements and Ndt80 binding site resembles that of Ndt80 targets or non-targets. Because the small number of full-length Ty3 found in the Saccharomyces mentioned above gave us limited statistical power, we combined full-length elements with Ty3 solo-LTRs (n=498), which are the vestiges of Ty3 integrations lost by homologous recombination between LTRs. We found that taken together, Ty3 elements showed a significant enrichment of Ndt80 binding sites 200 bp upstream of their integration site (p < 0.001, Figure 3B). This result strongly supports the hypothesis that the co-opting of Ndt80 by Ty3 elements is a recurring feature.

Ty3 translation is repressed specifically in meiosis

Ty3 has co-opted a meiotic control element to activate its transcription specifically during meiosis. We next sought to determine how yeast responds to protect its developing gametes from the potential threat of Ty3 proliferation. Because Ty3 copy number seems to be relatively low and stable in yeast strains (Table S2), we reasoned that yeast cells might have evolved a defense mechanism that acts post-transcriptionally. We first wanted to ascertain whether Gag3 protein is produced in meiosis. We found that Gag3 was produced in our positive control – exposure of MATa SK1 haploid cells to α-factor mating pheromone (Figure 4A) – a condition in which full length Ty3 RNA was shown to be both transcribed and translated in another lab strain via the transcription factor Ste12 (Bilanchone et al., 1993). Hence, we confirmed that Ty3 is an active element in SK1. However, despite the presence of full-length Ty3 mRNA in meiosis, no detectable Gag3 protein was produced (Figure 4A). The absence of Gag3 protein in meiosis could be due to a lack of translation or rapid protein degradation. To test the latter possibility, we inhibited proteasomal function in meiotic cells. We found that either genetic (rpn6-1 (Isono et al., 2005)) or chemical inhibition (MG-132) of the proteasome did not result in Gag3 protein accumulation (Figure S4A). We concluded that the absence of Gag3 in meiosis is not due to
degradation by the proteasome. Our data instead support a model by which in response to Ty3 transcription in meiosis, yeast cells respond by preventing translation of the mRNA thereby halting the retrotransposition cycle.

Because Ty3 mRNA can be translated upon exposure of MATα cells to α-factor pheromone, we suspected that translational repression of Ty3 is specific to meiosis. To test this hypothesis, we ectopically expressed NDT80 in haploid mitotic cells to drive Ty3 transcription. In parallel, we depleted the mitotic repressor Sum1 which binds upstream of Ndt80 target genes to inhibit their inappropriate expression (Klutstein et al., 2010; Pierce et al., 2003; Winter, 2012). Selective depletion of Sum1 was achieved using the auxin-inducible degron system (Morawska and Ulrich, 2013). We found that when we expressed NDT80 in mitotic cells either in the presence or absence of Sum1 (Figure S4B), Ty3 was both transcribed and translated (Figure 4B). Ty3 mRNA and protein levels were much higher when Sum1 was depleted. We also observed that upon NDT80 activation in meiosis or mitosis, Ty3 mRNA levels were similar, while translation occurred only in the mitotic context. In addition we observed that Ty3 mRNA dimers, which are indicative of VLP assembly (Cristofari et al., 1999; Feng et al., 2000), form upon NDT80 induction in mitosis but not meiosis (Figure S4C). We conclude that the mechanism by which yeast represses translation of Ty3 mRNA is meiosis-specific. These data also provide additional support to the notion that this retrotransposon has co-opted the Sum1/Ndt80 regulon to restrict its expression to meiosis.

**Meiotic Ty3 repression is mediated by the translational repressor Rim4**

We next sought to elucidate the mechanism of Ty3 translational repression. Translational control of mRNA plays a central role in regulating gene expression during gametogenesis in virtually all sexual organisms (reviewed in (Kronja and Orr-Weaver, 2011)). In most organisms, transcription is shut off or limited during meiosis and regulation of gene expression relies on RNA-binding proteins that govern translation of previously transcribed mRNA. In *S. cerevisiae*, the RNA-binding protein Rim4 binds to and inhibits translation of mRNAs that encode protein products that are toxic when prematurely expressed but are necessary for late meiotic events (Berchowitz et al., 2013). We thus wanted to test whether repression of Ty3 translation was mediated by Rim4. Because Rim4 is required for meiotic entry we designed a selectively degradable Rim4 using the auxin-inducible degron system (Morawska and Ulrich, 2013) such that we could deplete it just before activation of NDT80. We found that Ty3 was translated upon premature Rim4 depletion (Figure 5). Rim4 must be assembled into an amyloid-like state in order to repress translation of its targets which requires the intrinsically disordered region (IDR) in its C-terminus (Berchowitz et al., 2015). We found that rim4ΔIDR mutants also translate Ty3 (Figure...
While our genetic analyses do not allow us to distinguish whether Rim4 directly or indirectly regulates Ty3 translation, we conclude that Rim4 amyloid-like repressors function in part to protect the germline from the proliferation of retrotransposons.

**Mammalian Gypsy-derived genes are regulated by meiotic transcription factors**

We next asked whether mammalian Ty3/Gypsy-like elements also exploit meiosis-specific transcription factors to activate during gametogenesis. In mammals most Ty3/Gypsy LTR-retrotransposon/retrovirus-derived elements no longer retain their retrotransposition capacity, and are classified based on whether Gypsy-internal sequences are present (28 loci in mouse and 693 loci in human contain Gypsy-internal sequences) (Campillos et al., 2006; Kojima, 2018). To test if any mammalian Gypsy-derived genes are switched on during meiosis, we analyzed RNA sequencing data from adult mouse testes (Soumillon et al., 2013). We found that several Gypsy-derived elements including those that contain Gag-like internal sequence (MamGyp-int) are transcribed in testes during the meiotic stage of spermatogenesis (spermatocytes) and in post-meiotic round spermatids (Figure S6A). When we analyzed uniquely mapping RNA sequencing reads of specific MamGyp-int loci, we found that two genes – Moap1 and Pnma1 – are highly upregulated in mouse meiotic spermatocytes (Figure 6A). Both Moap1 and Pnma1 are characterized as domesticated in mammals, and their open reading frames encode Gypsy-derived Gag-like proteins. Based on these analyses we propose that, analogous to Ty3 regulation in yeast, the Gypsy-like sequences Moap1 and Pnma1 are being transcriptionally activated in male mouse meiosis.

In eutherian mammals, there has been massive expansion of the Pnma gene family (15 genes in humans, 12 in mice (Campillos et al., 2006)). We wanted to test whether meiotic transcription of Pnma can be observed in marsupials which did not experience this expansion of Pnma genes and harbor only a single Pnma gene: M-PNMA (Kočošar and Kordiš, 2013). We analyzed tissue-specific RNA sequencing data from opossum (Lesch et al., 2016; Marin et al., 2017) and found that M-PNMA mRNA is highly and specifically upregulated in testes particularly in pachytene spermatocytes (Figure S6B). These data suggest that meiotic transcription of Pnma was either present in the common therian ancestor or has convergently evolved independently in both eutherians and marsupials.

We next wished to investigate if the transcriptional activation of Gypsy-derived elements in meiosis reflects the capture of element-external meiosis-specific transcription factor binding sites. Motivated by our findings in yeast, we focused our investigation on two master transcriptional regulators of meiosis: Stra8 and Mybl1 (also known as A-myb). Stra8 is required for meiotic initiation in mammals and promotes transcription of early meiotic genes (Anderson et al., 2008) whereas Mybl1 becomes active early in meiosis and is required for meiosis to progress through pachytene (Bolcun-Filas et al.,...
We assessed whether the meiotic transcription factors Stra8 and Mybl1 bind upstream of Moap1 and/or Pnma1 using ChIP-seq data from adult mouse testes (Kojima et al., 2019; Zhou et al., 2017). We found that both Mybl1 and Stra8 ChIP-seq peaks are present upstream of multiple Moap1 TSSs (Figure 6B, Figure S6C, D). Similarly, a Mybl1 ChIP-seq peak overlaps the Pnma1 TSS external to the MamGyp-int sequence (Figure 6C, Figure S6C, D).

If Moap1 and Pnma1 are under the direct control of Stra8 and Mybl1, developing gametes that do not express these transcription factors should not produce these transcripts. To test this prediction, we analyzed whether Moap1 and Pnma1 are expressed in preleptotene Stra8+/− and preleptotene Stra8−/− spermatocytes (Kojima et al., 2019) or in P14 Mybl1+/− and Mybl1−/− testes (Zhou et al., 2017). In concordance with the ChIP-seq analysis, Moap1 transcript abundance is significantly reduced in the absence of Stra8 (Figure 6D). In addition, both Pnma1 and Moap1 transcript abundance is reduced in the absence of Mybl1 (Figure 6E). Taken together, our data support the hypothesis that Gypsy elements co-opt critical meiotic transcription factors to ensure their expression in gametogenesis and that this regulation exists from yeast to mammals.

**Mammalian meiotic translational repressor proteins bind to Gag-containing mRNA**

In mammals, the functional orthologs of Rim4 are the DAZ family of RNA-binding proteins (Berchowitz et al., 2015). DAZ family proteins act as regulators of translation and exhibit a similar organization to Rim4 in that they harbor N-terminal RRM domains and a C-terminal prion-like domain (King et al., 2012). Furthermore, Dazl (DAZ-like) forms SDS-resistant amyloid-like assemblies specifically in mouse testes (Berchowitz et al., 2015). Thus, we asked whether meiotic Gypsy-derived RNAs might be regulated by DAZ family proteins. We analyzed Dazl CLIP data generated from adult mouse testes (Zagore et al., 2018) and observed clear enrichment peaks within the 3' UTRs of Moap1 and Pnma1 transcripts (Figure 6B, C). In addition, we analyzed DAZL-RNA complexes isolated by RIP in oocyte-containing human fetal ovaries (Rosario et al., 2017). We found that among human Gypsy-derived elements, the cognate human MOAP1 and PNMA1 transcripts are significantly enriched in DAZL immunoprecipitations (Figure 6F, Figure S6E). Of all the 802 different types of retrotransposon expressed in this tissue, only five were enriched in the DAZL RNA immunoprecipitations (Figure S6F, G). We propose that the physical interaction between DAZ family proteins and transcripts containing Gypsy-derived sequences exists in both mice and humans and in both male and female gametogenesis. Taken together, these data support an ancient association between amyloid-like RNA-binding assemblies and retrotransposon and retrotransposon-derived mRNA. We propose that this
association possibly represents a functional repurposing of the retrotransposon defense mechanism we observed in yeast.
DISCUSSION

In this study, we elucidate mechanisms underlying an evolutionary battlefront where retrotransposons attempt to proliferate in the germline and host organisms respond to defend the genomes of their progeny. Ty3/Gypsy retrotransposons in yeast and domesticated Gypsy-derived genes in mammals co-opt host master meiotic regulators to transcriptionally activate during gametogenesis. Based on these findings we propose that the integration site of a retrotransposon can encode the critical regulatory information that the element requires to activate specifically during gamete development. We further show that yeast cells utilize a meiosis-specific translational repressor to thwart further progress in the retrotransposon life cycle. Likewise, transcripts from domesticated Ty3/Gypsy elements in mammals are bound by a cognate translational regulator. In the yeast strain SK1, we find that meiotic cells efficiently repress translation of Ty3 mRNA which very likely prevents the element from proliferating in this context. However, because the control of retrotransposons by meiotic transcription factors occurs frequently in yeast, we expect this activation strategy to be successful in other strains and other organisms. Future studies will be necessary to determine the extent to which meiotic activation and host translational repression of Ty3/Gypsy elements can be generalized across organisms.

Arkhipova and Meselson put forth the idea that one of the major forces driving the pervasive maintenance of meiosis and sex is to allow organisms to focus defense strategies against transposable elements to a subset of cells (i.e. the germline) that will provide the genetic material for generations to come (Arkhipova and Meselson, 2005). In multicellular organisms, it is of obvious benefit to the retrotransposon to direct activation to the germline with the aim to elude host defense mechanisms. Why does the Ty3 retrotransposon activate only during the sexual cycle (meiosis and mating) in a unicellular organism such as yeast? In unicellular organisms, the entire organism is the germline and thus coordination of retrotransposon activation with meiosis may seem less obvious. Retrotransposon activation poses a resource cost on the cell (reviewed in (Bourque et al., 2018)). In rapidly proliferating cells, this cost could indeed translate into a fitness disadvantage. By restricting activation to meiosis and mating in yeast, Ty3 remains latent, thus minimizing its fitness cost and focusing its efforts to proliferate in cells that are not cycling. Furthermore, linking retrotransposon proliferation to meiotic recombination i.e. mixis could be valuable for the element. Organisms can generate favorable allelic combinations via meiotic recombination. Retrotransposons could propagate into advantageous allelic combinations by activating during and post-recombination which could offset the fitness cost of an additional element. Lastly, the programmed double-strand breaks that initiate meiotic recombination could provide access points for retrotransposon cDNA.
The transcriptional activation of LTR-retrotransposons is often governed by regulatory sequences within the LTRs themselves. For example, the 5’ LTR of Ty1/Copia contains binding sites for several transcriptional activators including Gcn4, Gcr1, Ste12, Tec1, and Tea1 and the repressor Mot3 (Servant et al., 2008). Likewise, the ability of Ty3/Gypsy to activate during mating (or upon α-factor exposure) relies on Ste12 binding within its 5’ LTR (Bilanchone et al., 1993). In the situation described here however, Ty3 utilizes regulatory sequences upstream of its integration site. What selective pressures then would drive Gypsy-like elements to utilize external regulatory sites to activate during meiosis? We propose that there may be some advantage to decouple meiotic activation from the intrinsic properties of the element. Meiosis-specific genes are repressed during vegetative/somatic growth. As we mentioned above, in yeast the Ndt80 binding site is also recognized by the repressor Sum1 (Klutstein et al., 2010; Pierce et al., 2003). If this sequence was harbored in the 5’ LTR, the element could be condemned to a single strategy by virtue of the presence of Sum1 in non-meiotic contexts. By utilizing external regulatory elements, the element can avoid intrinsic repressor binding sites while gaining the flexibility to utilize other transcription activators that may be advantageous in other selective contexts.

As illustrated by decades of previous work, the study of Ty elements is a powerful model for the discovery and elucidation of host mechanisms that restrict retrotransposition (Maxwell and Curcio, 2007). Genetic screens designed to identify factors that influence Ty1 (Griffith et al., 2003; Scholes et al., 2001) and Ty3 (Aye et al., 2004; Irwin et al., 2005) retrotransposition have shown that RNA processing factors can positively and negatively influence Ty retrotransposition. Three RNA processing factors suppress both Ty1 and Ty3 retrotransposition: The DExH-box RNA helicase Dbp3, The ribosomal protein Rpl6A, and the lariat debranching enzyme Dbr1. Studies motivated in part from these screens have identified that mRNA processing bodies (P-bodies) (Beliakova-Bethell et al., 2006) and nucleoporins (Rowley et al., 2018) play important roles in the life cycle of Ty3 elements. Here, we show that host repression of Ty3 is either directly or indirectly carried out by amyloid-like assemblies of the RNA-binding protein Rim4 – rim4::IDR mutants that fail to assemble also fail to repress Ty3 translation. It will be exciting in future studies to assess whether and how Rim4 assemblies interact with these factors to defend the host genome against retrotransposon proliferation.

Rim4 binds to and represses translation of numerous mRNAs that encode proteins required for late meiotic events. While many of these targets are translated precisely at meiosis II onset, several others are translated later (Berchowitz et al., 2013; Brar et al., 2012) and apparently some, such as Ty3, are seemingly never translated. We hypothesize that the fate of Rim4 targets is dictated by interaction with other factors. This hypothesis is consistent with the developmentally-regulated clearance of Rim4 multimeric assemblies at meiosis II onset (Carpenter et al., 2018). Targets that are
translated significantly after meiosis II onset likely interact with the RNA-binding proteins Pes4 and Mip6 (Jin et al., 2017) and we speculate that the terminal repression that we observe for Ty3 is carried out by a factor unknown to us.

To escape Rim4-mediated translational repression, Ty3 could take the advantage of a particular feature of the meiotic stage at which it is expressed. Prior to meiotic prophase exit, cells can return to a mitotic cycle if nutrients are sensed by the cell in a process known as return to growth which is a common event in the life cycle of wild yeast (Brion et al., 2017; Tsuchiya et al., 2014). Ty3 mRNA accumulates 15-30 minutes before the meiotic commitment point. Upon return to growth, Rim4-mediated translational control is relieved (Jin et al., 2017) and therefore Ty3 would recover the ability to propagate in a post-mixis environment. Similarly, if Ty3 mRNA persists until the end of gametogenesis and into spores, germination would likely provide a permissive environment for Ty3 propagation. This strategy would be analogous to what is observed in LINE-1 elements which are expressed during mouse gametogenesis and then only re-integrate after fertilization during early embryonic development (Kano et al., 2009).

While the proliferation of Ty3/Gypsy elements is often deleterious it is possible that their evolutionary cost is offset by providing an activity or function that is beneficial to the host (Cosby et al., 2019). Our preliminary results suggest that Ty3 deletion mutants do not exhibit a detectable defect (or advantage) in any aspect of gametogenesis including meiosis, spore formation, and/or spore viability. However, we speculate mutations in other Ty elements could be different in this regard. One exciting speculation is that by integrating downstream of meiosis-specific regulatory elements, a new coding sequence is now readily available to be domesticated for a meiosis-specific function. Similarly, the interaction of retrotransposons with regulatory RNA-binding proteins could connect the evolving gene to additional regulatory pathways. Retrotransposons are ripe for domestication because they encode several proteins with diverse and useful domains (Volff, 2006). Possible examples of meiotic genes evolving from Gypsy sequences include the mammalian Gypsy-like elements Moap1 and Pnma1 which do not represent active retrotransposons and are likely domesticated. We observe meiotic transcriptional activation of Pnma genes in mouse, human, and opossum and while meiosis-specific activation of Gypsy-like elements is unlikely to be generalizable among all copies of these elements within a species, it appears to be generalizable to some degree among species. Remarkably, the single Pnma gene (M-PNMA) in opossum is highly upregulated in pachytene spermatocytes. Whether this represents the ancestral expression pattern of Pnma or convergent evolution, we speculate that meiotic activation of Pnma is providing some benefit to the organism and is being selected for. Because Pnma is inactive for retrotransposition, we propose that during the time when Gypsy elements were still retrotranspositionally active and co-opting meiotic transcription factors for their own benefit, they
acquired beneficial functions in the host meiotic cells. It is exciting to us that the regulatory circuitry established during the evolutionary battle between host and retrotransposon could also give direction as to how and when the emerging gene will function. It will be important to investigate the emerging roles of these elements and how they are repurposing retrotransposon regulatory mechanisms in the process of acquiring functions and evolving into new genes.

We propose that the co-opting of meiosis-specific transcription factors by retrotransposons represents an ancient strategy to coordinate retrotransposition with germline development. Correspondingly, translational repression of retrotransposons by host RNA-binding proteins could also be an ancient and preserved feature of meiosis. We can add this mechanism to the elaborate repertoire of retrotransposon defense systems such as piRNA, DNA methylation, and heterochromatic DNA condensation. It will be interesting to assess whether other assembly-forming RNA-binding proteins intervene in the life cycles of other retrotransposons in other organisms.
ACKNOWLEDGMENTS

We would like to thank Virginia Bilanchone and Suzanne Sandmeyer for providing the anti-Ty3 VLP crude serum, Kayla Carpenter for RNA sequencing prep, and S. Grace Herod for the his3::pRIM4-OsTIR1::His3MX6 construct. We would like to thank Albert Serra Cardona and Zhiguo Zhang for help with Ndt80 ChIP. We would also like to thank Stephen Goff and members of the Berchowitz lab for valuable discussions and Jonathan Dworkin, Diana Ottoz, and Tanguy Lucas for critical reading of the manuscript. This research is supported by The Schaefer Research Scholars Program, The Hirschl Family Trust, NIH grant R35 GM124633-01 to L.E.B., MRC University Unit grant MC_UU_00007/6 to I.R.A., Agence Nationale de la Recherche grant ANR-16-CE12-0019 to G.L. and National Institutes of Health DP2 AG055946-1250 01 to EÜ. Finally, we would like to thank Harmit Malik who suggested an idea to L.E.B. that provided the motivation to begin this project.

AUTHOR CONTRIBUTIONS


DECLARATION OF INTERESTS

The authors declare no competing interests.
Figure 1: Ty3 expression is coordinated with meiotic progression

(A) Diagram of the SK1 Ty3 element DNA locus, full-length mRNA, and two polyproteins Gag3 (precursor of CA-capsid SP-spacer and NC-nucleocapsid proteins) and Gag3-Pol3 (precursor of the aforementioned proteins and of PR-protease, RT-RH-reverse transcriptase – RnaseH and IN-integrase proteins). Colored lines indicate the positions of the Northern probes used in this study: Ty3 probe (maroon), CA probe (orange), PR probe (yellow), RT probe (green), IN probe (blue).

(B) log2 tag density of polyA-RNAseq reads of Ty3 locus from synchronized meiotic cells (Crick strand top). Cells harboring NDT80-IN (NDT80 inducible strain B119) were sporulated at 30º C. Shown are the time points 0, 3, 6, 7, 8, and 9 h post transfer to sporulation medium. At 6 h when cells had arrested in G2 due to the lack of Ndt80 meiotic transcription factor, they were released from the block by addition of 1 µM β-estradiol (biological replicates = 2).

(C, D) Full-length Ty3 mRNA is expressed in meiosis. Sporulation was induced in wild type diploid strain (B47, purple) and in ty3Δ diploid strain (B837, gold). (C) Ty3 mRNA and rRNA (loading control) levels are shown (S = pre-meiotic DNA replication, G2 = meiotic prophase, MI = first meiotic division, MII = second meiotic division). Full-length Ty3 mRNA (5 kb) is indicated by a black arrow. Empty arrowheads indicate 18S and 25S rRNA (1.8 kb, 3.4 kb), both reported on the probed blot and serve as size markers. (D) Kinetics of meiotic progression was determined via DAPI nuclear staining. Shown are representative results of ≥ 5 biological replicates.

(E-G) Ty3 is activated during meiosis in wild S. cerevisiae. Wild type SK1 (B47, purple) and wild European strain OS673 (green, described in (Peter et al., 2018)) were induced to sporulate as in C. (E) Ty3 mRNA and rRNA levels are shown. (F) The number of Ty elements in SK1 and OS673 are shown. (G) Kinetics of meiotic progression determined by DAPI nuclear staining (biological replicates = 2).

Figure 2: Ty3 exploits binding sites of a meiotic transcription factor from the integration environment

(A, B) Ty3 expression is controlled by Ndt80. (A) Sporulation was induced in parallel in wild type (B47, purple), and in NDT80-IN (B48, blue, released from the G2 block at 6 h). Ty3 mRNA and rRNA levels are shown. Asterisk indicates a degraded RNA sample, which is excluded from the quantification shown in B. (B) Solid lines indicate quantification of RNA levels from A, dashed lines indicate kinetics of
meiotic divisions analyzed by tubulin immunofluorescence (quantification of cells containing meiotic spindles). Shown are representative results of ≥ 5 biological replicates.

(C, D) Ty3 transcription is directly activated from Ndt80 binding sites upstream of the integration site. Diagram of endogenous and mutated Ty3 upstream context. (C) Blue triangles indicate Ndt80 binding sites (CACAAAA) and yellow triangles indicated mutated binding sites (AAAAAAA) (-79 and -269 bp upstream of Ty3). (D) Sporulation was induced in NDT80-IN strains containing the wild type Ty3 Ndt80 binding sites (B1864, blue), or with mutated Ndt80 binding sites (B1885, yellow) as in A (in both strains, a kanMX marker was inserted 353 bp upstream the Ty3 5' LTR). Ty3 mRNA and rRNA levels are shown. Shown are representative results of 2 biological replicates.

(E, F) Ndt80 binds directly upstream of Ty3 integration sites as meiosis progresses. Sporulation was induced in wild type diploid strain (B47 no tag) and in NDT80-3V5 diploid strain (B1674). Shown are the time points 0, 3, 5, 6 h post-transfer to sporulation medium. (C) Analysis of Ndt80 protein expression. Ndt80 (tagged with 3V5) and Pgk1 (loading control) protein levels are shown. Migration of molecular size standards is indicated on right as kDa values. (D) Browser tracks showing Ndt80-3V5 ChIP-seq peaks (blue bars). Ty3 (left) and NDT80 (right, positive control) loci are shown.

**Figure 3: Integration of Ty3 downstream of Ndt80 is a recurring phenomenon**

(A) Left: Phylogenetic relationships among seven S. cerevisiae (S. c.) strains and five S. paradoxus (S. p.) strains (modified from (Yue et al., 2017)), along with their respective number of full-length Ty3 elements (total 20). Right: 2-kb window map of 1000 bp upstream of full-length Ty3 elements with Ndt80 binding sites (CACAAAA) indicated as blue triangles. Ndt80 binding sites are significantly enriched in this region (determined by Z-test statistic of 100 randomized simulations; Z-score 2.7; p < 0.01) (B) Left: 260 SK1 genes were identified as Ndt80 target genes (Chu et al., 1998), and 5291 as non-target genes. Within the 13 S.c. and S.p. strains analyzed above there are 518 Ty3 elements (20 full-length and 498 solo-LTRs). Right: Comparison of the distribution of Ndt80 binding sites in the region 500 bp upstream of the TSS among target genes (teal), non-target genes (grey), and Ty3 elements (blue). Statistical significance was determined by bootstrap analysis (p < 0.001 at the -200 position).

**Figure 4: Ty3 is post-transcriptionally repressed in meiosis**

(A) Gag3 protein encoded by Ty3 is not detectable during meiosis. Sporulation was induced in NDT80-IN (B48) as in 2A. MATa wild type (B1) and ty3Δ (B827) haploid strains were exposed to α-factor (0, 0.5µM, 10µM) to induce Ty3 activation. Left: Ty3 mRNA (black arrow) and rRNA (white arrows) levels,
and Gag3 (and its processed forms CA-SP, CA) and Pgk1 (loading control) protein levels are shown. Migration of molecular size standards is indicated on right as kDa values. Right: Quantification of Ty3 full-length RNA (maroon) and Gag3 protein (green) normalized over their respective loading controls. Shown are representative results of \( \geq 5 \) biological replicates.

(B) Post-transcriptional repression of Gag3 is specific to meiosis. To express Ty3 in both meiosis and mitosis, we expressed NDT80 in sporulating diploid cells (wild type B48, ty3\( \Delta \) B839) and in haploid mitotically dividing cells with and without the Sum1 transcriptional repressor (wild type B477, Sum1-deg B829, and ty3\( \Delta \) B827). Ty3 mRNA (black arrow) and rRNA (white arrows) levels and Gag3 and Pgk1 (loading control) protein levels and quantification are shown, as in A. Asterisks show unspecific bands (e.g. band in \( \alpha \)-VLP present in ty3\( \Delta \) - see also Figure S3B). Shown are representative results of 5 biological replicates.

**Figure 5: Ty3 is translationally repressed by the RNA-binding protein Rim4**

Gag3 protein accumulates upon Rim4 depletion in meiosis. Meiosis was induced in three NDT80-IN diploid strains: wild type (B48), Rim4-deg (B1022, Rim4-degron fusion strain, induced by auxin addition to 1 mM at \( t = 5 \) h), and ty3\( \Delta \) (B389). Meiotic cells were released from the G2 block by addition of 1 \( \mu \)M \( \beta \)-estradiol at 6 h. Mitotic samples were collected from the experiment shown in 3B and Figure S3B. Ty3 mRNA (black arrow) and rRNA (white arrows) levels and Gag3, Rim4 (Rim4-deg is tagged with FLAG, wild type Rim4 is tagged with 3V5 in the first strain, and not tagged in the remaining strains), and Pgk1 (loading control) protein levels. Migration of molecular size standards is indicated on right as kDa values. Quantification of Ty3 mRNA and Gag3 protein relatively to their respective loading controls are shown. Asterisks show unspecific bands (e.g. band in \( \alpha \)-VLP present in ty3\( \Delta \) - see also Figure S3B). Shown are representative results of 4 biological replicates.

**Figure 6: Mammalian Gypsy-derived genes are activated by meiotic transcription factors and regulated by meiotic RNA-binding proteins**

(A) Analysis of uniquely mapping MamGyp-int RNAseq reads for MamGyp-int loci expressed during spermatogenesis in adult mouse testes. biological replicates = 3 for each stage.

(B, C) Browser tracks showing mouse testis Stra8-FLAG ChIP-seq peaks (orange bars, MACS2 \( q < 0.05 \)), mouse testis Stra8-FLAG ChIP reads (orange line plot, fold change in ChIP relative to input, data from wild-type animals subtracted as background), mouse testis Mybl1 ChIP-seq peaks (blue bars, MACS2 \( q < 0.05 \)), mouse testis Mybl1 ChIP reads (blue line plot, fold change relative to input), mouse
testis Dazl CLIP clusters (green bars, biologically reproducible CLIP clusters), mouse testis Dazl CLIP (green line plot, number of CLIP tags), mouse spermatocyte RNA-seq (black histogram, counts per million mapped reads), Gencode VM18 gene models encompassing (B) Moap1, or (C) Pnma1 (navy rectangles), and the genomic location of Gypsy-derived sequences (maroon rectangles).

(D) Expression of Moap1 and Pnma1 in Stra8⁻/⁻ preleptotene mouse spermatocyte RNAseq data. CPM, counts per million mapped reads (biological replicates = 2 for Stra8⁺/⁻ spermatocytes and = 3 for Stra8⁻/⁻ spermatocytes).

(E) Expression of Moap1 and Pnma1 in P14 Mybl1⁻/⁻ mouse testis RNAseq data (biological replicates = 3).

(F) Analysis of MamGyp-int containing transcripts from MOAP1 and PNMA1 in anti-DAZL RIP from human foetal ovaries (biological replicates = 3).

Asterisks indicate p < 0.05 as determined by Student’s t-test.

MAIN TABLES AND LEGENDS

N/A.
RESOURCES AVAILABILITY

Lead Contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact, Luke E. Berchowitz (leb2210@cumc.columbia.edu).

Materials Availability

All strains and reagents used in this study are available upon request.

Data and Code Availability

The accession number for the RNA sequencing and ChIP sequencing dataset generated during this study is EBI ENA / NCBI BioProject: PRJNA669383 (https://www.ncbi.nlm.nih.gov/bioproject/?term=PRJNA669383; https://www.ebi.ac.uk/ena/browser/view/PRJNA669383). Custom code and algorithms are available upon request.

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Yeast strains and constructs

Lab strains are described in Supplementary Table S1, and wild strains in Supplementary Table S2. Strains were constructed by crossing and/or using the PCR-based method described in (Longtine et al., 1998) using the SK1 genome reference and annotation from (Yue et al., 2017). PCR and fusion PCR involved templates pFA6a-kanMX6 and pFA6a-His3MX6 (Longtine et al., 1998) and pUB976 (Sawyer et al., 2019). Seamless deletion of Ty3 was introduced by pop-in pop-out procedure as described in (Storici et al., 2001), using pRG206MX (Gnügge, 2015) as template to delete Ty3 at the pop-in step. The origins of all constructs are listed in the key resources table (this study, or (Benjamin et al., 2003; Berchowitz et al., 2013, 2015; Carlile and Amon, 2008; Carpenter et al., 2018; Torres et al., 2007)).
METHOD DETAILS

Yeast growth and sporulation conditions

All yeast strains were grown at 30° C. For sporulation of lab strains, strains were pulled from -80° C stock and grown on YPG (3% glycerol) plates overnight, were then transferred to YPD 4% (YPD + 4% glucose) plates for 24 h, were then grown to saturation in YPD for 24 h, diluted in BYTA (1% yeast extract, 2% tryptone, 1% potassium acetate, 50 mM potassium phthalate) to OD$_{600}$ = 0.3 and grown overnight. Cells were washed with water, resuspended in sporulation medium (0.3% potassium acetate [pH 7.0], 0.02% raffinose) to OD$_{600}$ = 1.8. Following inoculation into sporulation medium, cells were incubated with vigorous shaking (t = 0 h). For Rim4-deg strain, at t = 5 h, auxin (indole-3-acetic acid, Sigma) was added to a final concentration of 1 mM in order to induce degradation of Rim4. For pdr5 mutant strain, proteasome inhibitor MG-132 was added at 6 h to 50 µM. For NDT80-IN (Benjamin et al., 2003) strains, at t = 6 h, β-estradiol was added to a final concentration of 1 µM (5 mM stock in ethanol, Sigma E2758-1G) to induce expression of NDT80 from the GAL1-10 promoter by activating the Gal4.ER fusion protein. For sporulation of wild strains, the procedure was essentially the same with the exception that if tetrads were already observed after overnight on YPG plates, they were pulled again from -80° C directly on YPD 4%, and the protocol was continued from there. Time points were taken for Western and Northern blot analysis and immunofluorescence/DAPI as indicated.

For mitotic Ndt80/Sum1 experiments, strains were grown overnight in YPG (glucose in YPD would inhibit the GAL1-10 promoter driving NDT80) and diluted to OD$_{600}$= 0.3. Following transfer to fresh YPG medium, cells were incubated with vigorous shaking for 3 h. Then auxin (indole-3-acetic acid, Sigma) was added (t = 0h) to a final concentration of 1 mM in order to induce degradation of Sum1. After 1 hour, β-estradiol was added to a final concentration of 1 µM (5 mM stock in ethanol, Sigma E2758-1G) to induce expression of NDT80 from the GAL1-10 promoter by activating the Gal4.ER fusion protein. Time points were taken for Western and Northern blot analysis as indicated (collected volumes were adjusted to the OD$_{600}$ at each time point).

For pheromone exposure experiments, strains were grown overnight in YPD and diluted to OD$_{600}$= 0.2. Following transfer to fresh YPD medium, cells were incubated with vigorous shaking for 3 h. Cells were pelleted and resuspended at an OD$_{600}$ of 0.36 (t = 0 h), then α-factor (Genscript) was added to a final concentration of 0.8µg/mL (0.5 µM, “+”) or 16µg/mL (10 µM, “++”) in order to induce Ty3 activation. Time points were taken for Western and Northern blot analysis as indicated (collected volumes were adjusted to the OD$_{600}$ at each time point).
Meiotic progression analysis

Cells were fixed in 3.7% formaldehyde. For analysis of nuclear divisions and spindle visualization, indirect immunofluorescence was performed as in (Carpenter et al., 2018) with minor modifications. To visualize spindles, we used rat anti-tubulin antibody (Bio-Rad) at a dilution of 1:200, and anti-rat-FITC antibody (Invitrogen) at a dilution of 1:200. Immunofluorescence samples were mounted in ProlongGold (Invitrogen) that included DAPI. Acquisition of images was conducted using a DeltaVision microscope (GE Healthcare). Spindle morphologies (n = 100 cells per time point) were classified as in (Lee and Amon, 2001). Briefly, Metaphase I cells were defined as cells with a single DAPI mass spanned by a short, thick, bipolar, meiotic spindle. Anaphase I cells were defined as cells with two distinct (though not always separated) DAPI masses, and a single long spindle that spans both DAPI masses. Metaphase II cells were defined as cells with two separate DAPI masses with each spanned by a bipolar, short, thick, meiotic spindle. Anaphase II cells were defined as cells with four distinct (though not always separated) DAPI masses with two long spindles. Then, these four categories of cells were grouped under the label “cells in meiotic divisions”.

For analysis of nuclear divisions only, cells were washed once with 100 mM phosphate, 1.2 M sorbitol buffer [pH 7.5], permeabilized with 1% Triton X-100 and stained with 0.05 μg/ml DAPI.

Northern blot analysis

Samples were prepared by snap freezing the pellet of 7.2 OD$_{600}$ units of cells, and total RNA was isolated using a (400 μL: 400 μL) mixture of TES buffer (10 mM Tris [pH 7.5], 10 mM EDTA, 0.5% SDS) and acid phenol: chloroform 5: 1 (Ambion) with zirconia/silica beads (BioSpec) while shaking at 1400 rpm (Thermomixer, Eppendorf) at 65°C for 30 min. After centrifugation at 16,000 g for 5 min, the supernatant was ethanol precipitated at -20°C overnight (1 ml 100% ethanol and 40 μl 3M sodium acetate [pH 5.5]), centrifugated at 16,000 g for 20 min, washed in 80% ethanol, dried, and resuspended in DEPC–treated water. For wild strains, RNA extraction required two chloroform purification. 5 μg of RNA was denatured at 55°C for 15 min in 50% formamide, 18.3% formaldehyde, and 5% MOPS and was resolved on a denaturing agarose gel (1.9% agarose, 3.7% formaldehyde, 1x MOPS buffer) for 2.5 hours at 80 V. RNA was transferred to a Hybond-N+ membrane (GE Healthcare) by capillary transfer in 10x SSC (1.5 M NaCl, 0.15 M trisodium citrate dihydrate, [pH 7]). The membrane was stained with methylene blue, incubated in hybridization Buffer (0.25 M Na-Phosphate [pH 7.2], 0.25 M NaCl, 1 mM EDTA, 7% SDS, and 5% dextran sulfate) at 65°C and probed with $^{32}$P–labeled PCR products prepared via Amersham Megaprime DNA labeling kit (GE Healthcare) and Illustra ProbeQuant columns (GE
Healthcare), transferred to a phosphor screen, and imaged on a Typhoon Trio variable mode imager (Amersham).

To observe Ty3 dimers we conducted semi-denaturing Northern blot analysis (adapted from (Nymark-McMahon et al., 2002)). Briefly, the procedure is as above except we omitted formamide in the denaturing mix and incubated 15 minutes at 30°C.

**Yeast RNA sequencing and sequencing analysis**

Briefly, poly-A pull-down was used to enrich mRNAs from total RNA samples and proceed on library preparation by using Illumina TruSeq RNA prep kit. Libraries were sequenced using single-end sequencing (100 bp) in multiplex using Illumina HiSeq4000 at Columbia Genome Center. RTA (Illumina) was used for base calling and bcl2fastq2 (version 2.17) for converting BCL to fastq format, coupled with adaptor trimming. Sequencing reads were mapped onto SK1 genome reference (Yue et al., 2017) using STAR (Dobin et al., 2013) with the following parameters: Multiple alignments allowed: 500 (takes into account the number solo-LTR loci in the genome); Multiple alignment retained: 1; Order of multiple alignments: Random (--outFilterMultimapNmax 500 --outSAMmultNmax 1 --outMultimapperOrder Random --outFilterScoreMinOverLread 0.3 --outFilterMatchNminOverLread 0.1 --seedSearchStartLmax 20 --seedSearchStartLmaxOverLread 0.2) while other parameters were set to defaults. Reads count were scaled by CPM (count per millions) and visualized using Ving software (Descrimes et al., 2015).

**Immunoblot analysis**

The pellet of 7.2 OD_{600} units of cells was resuspended in 5% TCA, incubated overnight at 4° C, washed with acetone, and dried. Cells were broken using 50 μl acid-washed glass beads (Sigma), 100 μl lysis Buffer (10 mM Tris-HCl [pH 11], 1 mM EDTA [pH 8], 2.75 mM DTT, Halt protease inhibitors (Thermo Fisher)), and a 45 sec process in a FastPrep-24 (MP Biomedicals) at max speed. 50 μL Loading Buffer (9% SDS, 0.75 mM Bromophenol blue, 187.5 mM Tris-HCl [pH 6.8], 30% glycerol, and 810 mM β-mercaptoethanol) was added, samples were heated at 100°C for 5 mins, and centrifuged 5 mins at 20,000 g. Polyacrylamide gels (precast 10% Criterion TGX Precast Midi Protein Gel 26 wells, Bio-Rad) were run on a midi gel system (Criterion Vertical Electrophoresis Cell, Bio-Rad) with SDS Running Buffer (190 mM glycine, 25 mM Trizma base, 3.5 mM, 1% SDS), with 4 μl sample per well. They were transferred using a semi-dry transfer apparatus (Trans-Blot Turbo Transfer System, Bio-Rad) to a
nitrocellulose membrane (Trans-blot turbo transfer kits – Nitrocellulose, Bio-Rad), and stained in Ponceau.

α-VLP antibody was purified from α-VLP (Ty3) crude serum (gift from Sandmeyer lab), using Gag3 protein produced from pLZL2422 (Larsen et al., 2008) coupled to an AminoLink column (Thermo).

Antibodies were prepared in 1X TBST with 1% milk 1% BSA. Primary antibodies were used at the following concentrations: α-VLP (purified from α-VLP (Ty3) crude serum, gift from Sandmeyer lab) 0.1ug/mL, α-HA.11 1: 1,000 (BioLegend), α-V5 1: 2,000 (Invitrogen), α-FLAG (Sigma) 1: 800, α-Pgk1 1: 20,000 (Novex). α-mouse or α-rabbit HRP-conjugated secondary antibodies (GE Healthcare) were used.

Ndt80 ChIP-seq

NDT80-3V5 and no tag control cells were sporulated at 30° C. Cells were fixed in 1% formaldehyde for 15 minutes at room temperature and quenched with 0.1 M glycine (final concentration). Samples were prepared by snap freezing the pellet of 20 OD$_{600}$ units of cells. Cells were resuspended in 0.4 ml lysis buffer (50 mM HEPES/KOH pH [7.5], 140 mM NaCl, 1 mM EDTA, 1% Triton-X, 0.1 % Na-deoxycholate, 1 mM PMSF) lysed with zirconia/silica beads (BioSpec) in a FastPrep homogenizer. The homogenate was centrifuged at 2,500 rpm for 1 minute and the supernatant was discarded. The pellet was resuspended in 0.25 ml lysis buffer and sonicated (15 sec on, 30 sec off, 10 cycles). Lysate was centrifuged at 10,000 g for 5 minutes at 4° C and supernatant was transferred to a new tube. 50 μL was taken for DNA input sample and the remaining sample was processed for ChIP.

Ndt80-DNA complexes were immunopurified using anti V5 agarose beads (Sigma) for 1 h at 4° C. Beads were washed and crosslinks were reversed using Chelex resin (Bio-Rad). Elution was carried out by boiling samples for 10 minutes followed by Proteinase K digestion (1 mg/ml 30 min at 55° C). Beads were boiled again for 10 minutes and the eluate was collected and purified/concentrated using a Qiagene Minelute kit. Sequencing libraries were prepared from the eluate using an Accel-NGS 1S plus (Swift Biosciences) per the manufacturer’s protocol. Libraries were sequenced using paired-end sequencing (2x100 bp) in multiplex using NextSeq 550.

QUANTIFICATION AND STATISTICAL ANALYSIS

Immunoblot and Northern blot were quantified using FIJI software (Schindelin et al., 2012). Signal intensity was normalized to a loading control (Pgk1 for immunoblots, 25S rRNA for Northern blots).
Values were multiplied by an arbitrary scaling factor, which was held constant within each figure panel for each technique. Statistical significance was determined by Student’s T-test or Z-test statistic as indicated. Number of biological replicates is indicated in figure legends where applicable.

**Statistical analysis for Ndt80 binding site enrichment upstream Ty3 insertions**

Data analysis was performed using the R project for statistical computing version 3.6.1 [https://www.r-project.org](https://www.r-project.org) and RStudio version 1.1.383 [https://www.rstudio.com](https://www.rstudio.com), utilizing the following packages: dplyr, tidyr, stringr, ggplot2, plotly, tools, boot, Biostrings, GenomicRanges, janitor, seqinr, BiocManager, TFBSTools. Briefly, yeast genomes were downloaded from (Yue et al., 2017) and Ndt80 binding sites (defined as CACAAA by (Nocedal et al., 2017)) were identified. Then the positions of Ndt80 binding site upstream Ty3 elements (full-length or solo-LTRs) were recorded. For Ndt80 target genes and non-target genes in SK1, the same process was applied, using the SK1 reference genome and annotation from (Nocedal et al., 2017). For bootstrap, the R function sample_n was used to take samples with replacements. Then the data was visualized using 50 bp long non-overlapping bins. We used the distribution of outlier boundaries for each bootstrap experiment to identify region of significant difference in Ndt80 binding site enrichment.

**Analysis of yeast ChIPseq data**

Reads were trimmed of their adaptors using cutadapt (Martin, 2011) and were aligned to the S. cerevisiae SK1 genome (Yue et al., 2017) using Bowtie2 version 2.3.4 (Langmead and Salzberg, 2012). Alignments were filtered using samtools v1.6 (Li et al., 2009) (samtools view -f 2 -F 1804 -q 30), and PCR duplicates were filtered out using Picard tool MarkDuplicates v2.23.3 (Broad Institute). Deeptools bamcoverage (Ramírez et al., 2016) was used to generate a coverage track of number of reads per bp after reads were extended to their mates, to reflect the actual fragment length. Coverage was then normalized using a script detailed in (De Muyt et al., 2018): upon the assumptions that for a transcription factor, peaks are expected to be narrow, they are expected to fall in the top outliers of the coverage distribution, while the background regions are expected to fall in the bulk of the distribution, between Q1-1.5IQR and Q3+1.5IQR. Hence, for untagged immunoprecipitate (UIP), the coverage was divided by the average read count in non-null and non-outlier positions, while for tagged immunoprecipitate (IP), the coverage was divided by the average read count in non-null and non-outlier positions and multiplied by the average read count of the same set of positions in the normalized UIP. Normalized profiles were used for peak detection using bpeaks v1.2 (Merhej et al., 2014) using as
thresholds: T1: IP>(Q3_{UIP} + 1.5IQR_{UIP})*mean(IP), T2: UIP>(Q1_{UIP} − 1.5IQR_{UIP})*mean(UIP), T3: log2FC>0.5, and T4: ((log2(IP)+1)+(log2(control)+1))/2>0. For visualization on IGV (Thorvaldsdóttir et al., 2013), normalized UIP profiles were subtracted to their corresponding normalized IP profiles using Deeptools bigwigCompare (Ramírez et al., 2016).

**Analysis of mammalian RNAseq data (mouse, opossum)**

RNAseq from enriched populations of adult mouse testicular cell types (Soumillon et al., 2013) was downloaded from EBI ENA (https://www.ebi.ac.uk/ena, accession PRJNA187158). Reads were trimmed to remove adapter sequences (--stringency 5 --length 20) using TrimGalore 0.4.1 (https://github.com/FelixKrueger/TrimGalore) then aligned to the mm10 reference assembly of the mouse genome with Tophat v2.1.0 (Kim et al., 2013) (command line options -g 1 --no-coverage-search --b2-sensitive) using the GencodeBasic VM18 transcriptome as a transcriptome index (Frankish et al., 2019). Accepted hits from Tophat were filtered with samtools v1.6 (Li et al., 2009) to remove unmapped reads and non-primary alignments. For locus-specific analysis, uniquely mapping reads were selected by removing reads with a quality score below 5. Reads overlapping Repeatmasker-annotated loci in the mouse genome were counted for each repeat locus, and aggregated at the level of the repeat name (repName). EdgeR (Robinson et al., 2010) was then used to normalize the counts per million mapped fragments in each sample using trimmed means. CPMs were converted to fragments per kilobase of transcript per million mapped reads (FPKM) by dividing CPM by the length of the repeat locus for locus-specific analysis, or by the mean length of repeat loci annotated with that repeat name for repName-level analysis.

RNAseq for P14 *Mybl1* testes (Zhou et al., 2017) and *Stra8* preleptotene spermatocytes (Kojima et al., 2019) were downloaded from EBI ENA (accessions PRJNA321732, PRJNA476515), trimmed and mapped to mm10 as described above. The paired end *Mybl1* data was mapped using mate inner distance and mate standard deviations of 30 and 68 respectively. Accepted hits from Tophat were filtered as described above and uniquely mapping reads falling within GencodeBasic VM18 genes counted using htseq-count v0.11.2 (Anders et al., 2015) in intersection-nonempty mode. The ENSMUST00000174651.1 transcript in the GencodeBasic VM18 annotation, which appears to be a *Moap1* transcript originating from an alternative upstream promoter, was manually re-assigned from the RP24-234J3.3 gene to *Moap1* to circumvent ambiguity in htseq-count assigning reads to *Moap1*. EdgeR was used to normalize the counts per million mapped fragments in each sample using trimmed means, data for genes of interest (*Moap1, Pnma1*) selected from the dataset, and Student’s t-test used.
to assess statistical significance between Mybl1\textsuperscript{+/-} and Mybl1\textsuperscript{-/-} P14 testis samples, or between Stra8\textsuperscript{+/-} highSTRA8 preleptotene spermatocyte and Stra8\textsuperscript{-/-} preleptotene spermatocyte samples.

FPKMs from RNA-seq for opossum from purified spermatocyte/spermatid and multitissue (Lesch et al., 2016; Marin et al., 2017) were downloaded from NCBI GEO with the following accessions: GSE68507, GSE97367.

**Analysis of mammalian ChIPseq data (mouse)**

Mybl1 and Stra8 ChIPseq data from mouse testes (Kojima et al., 2019; Zhou et al., 2017) were downloaded from EBI ENA (accessions PRJNA321732, PRJNA476515), trimmed as described for RNAseq data, then aligned to the mm10 assembly of the mouse genome using Bowtie v2.2.6 (Langmead and Salzberg, 2012). Bowtie output was filtered to remove PCR duplicates marked by PicardTools v2.17.11 (http://broadinstitute.github.io/picard/), and to remove reads mapping to mitochondrial or blacklisted regions of the genome (https://www.encodeproject.org/files/ENCFF023CZC/), unmapped reads, non-primary alignments and reads with a quality score below 5. Bedtools (Quinlan and Hall, 2010) was used to convert the Bowtie alignments to ENCODE tagAlign format, and MACS v2.1.1 (Zhang et al., 2008) used to call peaks on pooled ChIP samples relative to pooled input controls using default settings (q < 0.05) and to generate Bedgraphs showing fold enrichment in pooled ChIP relative to input samples. For Stra8 ChIPseq, the fold-enrichment for anti-FLAG ChIP in control wild-type samples was then subtracted from the fold-enrichment for the anti-FLAG ChIP in experimental Stra8\textsuperscript{FLAG/FLAG} samples. For Mybl1 ChIPseq, anti-Mybl1 ChIPseq from Mybl1\textsuperscript{+/-} testes was compared to Mybl1\textsuperscript{-/-} input controls. ChIPseq bedgraphs and peak co-ordinates were visualised using Gviz (Hahne and Ivanek, 2016). Distances from Repeatmasker-annotated MamGyp-int loci and GencodeBasic VM18 genes overlapping those loci to the nearest Stra8 and Mybl1 ChIPseq peaks were determined using BEDtools v2.26.0 (Quinlan and Hall, 2010).

**Analysis of Dazl CLIP and DAZL RIP mammalian data (mouse, human)**

Dazl CLIP tag locations from mouse testes (Zagore et al., 2018) were downloaded from GEO (accession GSE108181) and the genomic co-ordinates lifted over from mm9 to mm10 using the UCSC liftOver tool (Haeussler et al., 2019). DAZL RIP data from human ovaries (Rosario et al., 2017) (available in EBI ENA, PRJNA321858) was trimmed and aligned to the hg38 assembly of the human genome as described for the RNAseq data using a mate inner distance of 260 and mate standard
deviation of 110. Tophat output was filtered, then reads in genes or repeats counted using htseq-count or BEDTools and normalized in EdgeR using the total number of mapped fragments as the library size as described for the RNAseq data. To assess MOAP1 and PNMA1 behaviour in the htseq-count data, EdgeR, low abundance genes were removed using \texttt{filterByExpr(count=15, min.count=10)}, \texttt{glmLRT} used to determine differentially abundant genes in a paired experimental design, MOAP1 and PNMA1 data extracted from the resulting DGELRT object, and false discovery rate used to correct p-values for multiple-testing. To determine enriched retrotransposon sequences in the anti-DAZL samples, \texttt{repName-aggregated repeat count data for retrotransposons (LINE, SINE and LTR repeat classes)} were selected then low abundance retrotransposons removed using \texttt{filterByExpr(count=15, min.count=10)}. \texttt{glmLRT} was used to determine differentially abundant retrotransposons in a paired experimental design, and false discovery rate used to correct p-values for multiple-testing. Data for retrotransposons belonging to the Gypsy repeat family were plotted.
REFERENCES


Genomics, (Springer), pp. 335–351.


<table>
<thead>
<tr>
<th>REAGENT or RESOURCE</th>
<th>SOURCE</th>
<th>IDENTIFIER</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Antibodies</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>α-HA.11 Clone:16B12 (mouse)</td>
<td>BioLegend</td>
<td>Cat# 901514</td>
</tr>
<tr>
<td>α-Pgk1 (Phosphoglycerate Kinase monoclonal) (mouse)</td>
<td>Novex (Life Technologies)</td>
<td>Part# 459250</td>
</tr>
<tr>
<td>α-v5 monoclonal (mouse)</td>
<td>Invitrogen</td>
<td>Cat# 46-0705</td>
</tr>
<tr>
<td>α-FLAG (rabbit)</td>
<td>Sigma-Aldrich</td>
<td>Cat# F7425</td>
</tr>
<tr>
<td>α-VLP (Ty3) crude serum (rabbit)</td>
<td>Sandmeyer lab (UC Irvine)</td>
<td>α-VLP (Ty3) crude serum</td>
</tr>
<tr>
<td>α-VLP (Ty3) affinity purified antibody, using <em>E. coli</em> produced Gag3 protein</td>
<td>This study</td>
<td>α-VLP (Ty3) affinity purified antibody</td>
</tr>
<tr>
<td>α-mouse HRP-conjugated secondary</td>
<td>GE Healthcare</td>
<td>Cat# NA931-1ML</td>
</tr>
<tr>
<td>α-rabbit HRP-conjugated secondary</td>
<td>GE Healthcare</td>
<td>Cat# NA934V</td>
</tr>
<tr>
<td>α-V5-coupled agarose</td>
<td>Sigma-Aldrich</td>
<td>Cat# A7345-1ml</td>
</tr>
<tr>
<td>Rat α-tubulin alpha</td>
<td>Bio-Rad</td>
<td>Cat# MCA77G</td>
</tr>
<tr>
<td>α-rat-FITC</td>
<td>Invitrogen</td>
<td>Cat# 31629</td>
</tr>
<tr>
<td><strong>Chemicals, Peptides, and Recombinant Proteins</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Halt Protease Inhibitor</td>
<td>ThermoFisher</td>
<td>Cat# 1861279</td>
</tr>
<tr>
<td>ProlongGold anti-fade reagent w/ DAPI</td>
<td>Invitrogen</td>
<td>Cat# P36935</td>
</tr>
<tr>
<td>β-Estradiol</td>
<td>Sigma-Aldrich</td>
<td>Cat# E8875</td>
</tr>
<tr>
<td>MG-132</td>
<td>Cayman Chemical</td>
<td>Cat# 10012628</td>
</tr>
<tr>
<td>Auxin (indole-3-acetic acid)</td>
<td>Sigma-Aldrich</td>
<td>Cat# 12886-25G</td>
</tr>
<tr>
<td>α-factor mating pheromone</td>
<td>Genscript</td>
<td>Cat# RP101002</td>
</tr>
<tr>
<td>Acid-washed glass beads, 425-600µm</td>
<td>Sigma-Aldrich</td>
<td>Cat# G8772-500G</td>
</tr>
<tr>
<td>0.5mm dia Zirconia/silica beads</td>
<td>BioSpec Products</td>
<td>Cat# 11079105Z</td>
</tr>
<tr>
<td>Chelex resin</td>
<td>Bio-Rad</td>
<td>Cat# 1422832</td>
</tr>
<tr>
<td>Hybond-N+ membrane</td>
<td>GE Healthcare</td>
<td>Prod# RPN203B</td>
</tr>
<tr>
<td>10% Criterion TGX Precast Midi Protein Gel, 26 well, 15 µl</td>
<td>Bio-Rad</td>
<td>Cat# 5671035</td>
</tr>
<tr>
<td>Trans-blot turbo transfer kits (Nitrocellulose)</td>
<td>Bio-Rad</td>
<td>Cat# 1704271</td>
</tr>
<tr>
<td>Illustria Probequant Columns</td>
<td>GE Healthcare</td>
<td>Prod# 28903408</td>
</tr>
<tr>
<td><strong>Critical Commercial Assays</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AminoLink Plus Immobilization kit</td>
<td>Thermo Scientific</td>
<td>Cat# 44894</td>
</tr>
<tr>
<td>Amersham MegaPrime DNA Labeling Kit</td>
<td>GE Healthcare</td>
<td>Prod# RPN1604</td>
</tr>
<tr>
<td>Minelute kit</td>
<td>Qiagen</td>
<td>Cat# 28004</td>
</tr>
<tr>
<td>Accel-NGS 1S plus DNA library kit</td>
<td>Swift Biosciences</td>
<td>Cat# 10096</td>
</tr>
<tr>
<td>TruSeq Stranded Total RNA Kit</td>
<td>Illumina</td>
<td>Cat# RS-122-2302</td>
</tr>
<tr>
<td><strong>Deposited data</strong></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
| SK1 reference genome and other *S. cerevisiae* and *S. paradoxus* reference genomes | (Yue et al., 2017) | [https://yix1217.github.io/Yeast_Pa cBio_2016/data/](https://yix1217.github.io/Yeast_PacBio_2016/data/)
<p>| Mouse reference genome: mm10 assembly | Mouse Genome Sequencing Consortium | <a href="https://genome.ucsc.edu/">https://genome.ucsc.edu/</a> |</p>
<table>
<thead>
<tr>
<th>REAGENT or RESOURCE</th>
<th>SOURCE</th>
<th>IDENTIFIER</th>
</tr>
</thead>
<tbody>
<tr>
<td>RNAseq adult mouse testicular cell types, EBI ENA / NCBI BioProject accession PRJNA187158</td>
<td>(Soumillon et al., 2013)</td>
<td><a href="https://www.ebi.ac.uk/ena/browser/view/PRJNA187158">https://www.ebi.ac.uk/ena/browser/view/PRJNA187158</a></td>
</tr>
<tr>
<td>RNAseq for P14 Mybl1 testes EBI ENA / NCBI BIOPROJECT accession PRJNA321732</td>
<td>(Zhou et al., 2017)</td>
<td><a href="https://www.ebi.ac.uk/ena/browser/view/PRJNA321732">https://www.ebi.ac.uk/ena/browser/view/PRJNA321732</a></td>
</tr>
<tr>
<td>RNAseq Stra8 preleptotene spermatocytes EBI ENA / NCBI BIOPROJECT accession PRJNA476515</td>
<td>(Kojima et al., 2019)</td>
<td><a href="https://www.ebi.ac.uk/ena/browser/view/PRJNA476515">https://www.ebi.ac.uk/ena/browser/view/PRJNA476515</a></td>
</tr>
<tr>
<td>Repeatmasker-annotated loci in mm10</td>
<td>Institute for Systems Biology</td>
<td><a href="http://repeatmasker.org/species/mm.html">http://repeatmasker.org/species/mm.html</a></td>
</tr>
<tr>
<td>GencodeBasic VM18 annotation NCBI GEO accession GSE108181</td>
<td>GENCODE project</td>
<td><a href="https://www.gencodegenes.org/mouse/release_M18.html">https://www.gencodegenes.org/mouse/release_M18.html</a></td>
</tr>
<tr>
<td>Mybl1 ChiPseq data from mouse testes EBI ENA / NCBI BIOPROJECT accession PRJNA321732</td>
<td>(Kojima et al., 2019)</td>
<td><a href="https://www.ebi.ac.uk/ena/browser/view/PRJNA321732">https://www.ebi.ac.uk/ena/browser/view/PRJNA321732</a></td>
</tr>
<tr>
<td>Stra8 ChiPseq data from mouse testes EBI ENA / NCBI BIOPROJECT accession PRJNA476515</td>
<td>(Kojima et al., 2019)</td>
<td><a href="https://www.ebi.ac.uk/ena/browser/view/PRJNA476515">https://www.ebi.ac.uk/ena/browser/view/PRJNA476515</a></td>
</tr>
<tr>
<td>Human reference genome: hg38 assembly</td>
<td>Genome Reference Consortium</td>
<td><a href="https://genome.ucsc.edu/">https://genome.ucsc.edu/</a></td>
</tr>
<tr>
<td>Repeatmasker-annotated loci in hg38</td>
<td>Institute for Systems Biology</td>
<td><a href="http://repeatmasker.org/species/hg.html">http://repeatmasker.org/species/hg.html</a></td>
</tr>
<tr>
<td>DAZL RIP from human ovaries EBI ENA / NCBI BIOPROJECT accession PRJNA321858</td>
<td>(Rosario et al., 2017)</td>
<td><a href="https://www.ebi.ac.uk/ena/browser/view/PRJNA321858">https://www.ebi.ac.uk/ena/browser/view/PRJNA321858</a></td>
</tr>
</tbody>
</table>

**Experimental Models: Organisms/Strains**

All lab strains are S. cerevisiae, of SK1 background (see table S1) | (Kane and Roth, 1974) | ATCC: 204722 |

**Oligonucleotides**

<table>
<thead>
<tr>
<th>Primer</th>
<th>Sequence</th>
<th>Source</th>
<th>N/A</th>
</tr>
</thead>
<tbody>
<tr>
<td>FW primer for Ty3 Northern probe: 5'-GCTTTATGGATCAAATCCCC-3'</td>
<td>This paper</td>
<td>N/A</td>
<td></td>
</tr>
<tr>
<td>RV primer for Ty3 Northern probe: 5'-AGCATATATCGGAAGTGGTGGA-3'</td>
<td>This paper</td>
<td>N/A</td>
<td></td>
</tr>
<tr>
<td>FW primer for CA Northern probe: 5'-GCTTTATGGATCAAATCCCC-3'</td>
<td>This paper</td>
<td>N/A</td>
<td></td>
</tr>
<tr>
<td>RV primer for CA Northern probe: 5'-ACCGATGATAGTGTCTCC-3'</td>
<td>This paper</td>
<td>N/A</td>
<td></td>
</tr>
<tr>
<td>FW primer for PR Northern probe: 5'-TATATCTGGCAGTGGATGG-3'</td>
<td>This paper</td>
<td>N/A</td>
<td></td>
</tr>
<tr>
<td>RV primer for PR Northern probe: 5'-GACAAACATTTGGAATTTTTCC-3'</td>
<td>This paper</td>
<td>N/A</td>
<td></td>
</tr>
<tr>
<td>FW primer for RT Northern probe: 5'-GATAACAGTTGTTGATGATCC-3'</td>
<td>This paper</td>
<td>N/A</td>
<td></td>
</tr>
<tr>
<td>RV primer for RT Northern probe: 5'-GGAGCAATTTGGAATCAG-3'</td>
<td>This paper</td>
<td>N/A</td>
<td></td>
</tr>
<tr>
<td>REAGENT or RESOURCE</td>
<td>SOURCE</td>
<td>IDENTIFIER</td>
<td></td>
</tr>
<tr>
<td>----------------------------------------</td>
<td>-------------------------</td>
<td>------------</td>
<td></td>
</tr>
<tr>
<td>FW primer for IN Northern probe:</td>
<td>This paper</td>
<td>N/A</td>
<td></td>
</tr>
<tr>
<td>5'-GACGCCTCAAAAGACGGC-3'</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RV primer for IN Northern probe:</td>
<td>This paper</td>
<td>N/A</td>
<td></td>
</tr>
<tr>
<td>5'-TTCCAAGTGTTCTAGTACGG-3'</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Recombinant DNA

<table>
<thead>
<tr>
<th>Construct:</th>
<th>(Benjamin et al., 2003)</th>
<th>N/A</th>
</tr>
</thead>
<tbody>
<tr>
<td>ura3::pGPD-GL4(848).ER::URA3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>pGAL-NDT80::TRP1</td>
<td>(Benjamin et al., 2003)</td>
<td>N/A</td>
</tr>
<tr>
<td>CLB3-3HA:kanMX6</td>
<td>(Carlile and Amon, 2008)</td>
<td>N/A</td>
</tr>
<tr>
<td>RIM4-3V5::His3MX6</td>
<td>(Berchowitz et al., 2013)</td>
<td>N/A</td>
</tr>
<tr>
<td>RIM4-AID*-6FLAG::HYGR</td>
<td>this paper</td>
<td>N/A</td>
</tr>
<tr>
<td>rim4Δ138-3V5::His3MX6</td>
<td>(Berchowitz et al., 2015)</td>
<td>N/A</td>
</tr>
<tr>
<td>rim4Δ271-3V5::His3MX6</td>
<td>(Berchowitz et al., 2015)</td>
<td>N/A</td>
</tr>
<tr>
<td>pdr5::TRP1</td>
<td>(Torres et al., 2007)</td>
<td>N/A</td>
</tr>
<tr>
<td>rpn6-1::His3MX6</td>
<td>(Carpenter et al., 2018)</td>
<td>N/A</td>
</tr>
<tr>
<td>ty3Δ</td>
<td>this paper</td>
<td>N/A</td>
</tr>
<tr>
<td>MSE-His3MX6-Ty3</td>
<td>this paper</td>
<td>N/A</td>
</tr>
<tr>
<td>kanMX6-Ty3</td>
<td>this paper</td>
<td>N/A</td>
</tr>
<tr>
<td>kanMX6-mse(-269)Δ::AAAAA+TT(-163)-mse(-79)Δ::AAAAA-Ty3</td>
<td>this paper</td>
<td>N/A</td>
</tr>
<tr>
<td>his3::pGPD1-OsTIR1::His3MX6</td>
<td>this paper</td>
<td>N/A</td>
</tr>
<tr>
<td>his3::pRIM4-OsTIR1::His3MX6</td>
<td>this paper</td>
<td>N/A</td>
</tr>
<tr>
<td>SUM1-3V5-IAA7-kanMX</td>
<td>this paper</td>
<td>N/A</td>
</tr>
<tr>
<td>(to build RIM4-AID*-6FLAG::HYGR)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>pFA6a-kanMX6</td>
<td>(Longtine et al., 1998)</td>
<td><a href="https://www.addgene.org/39296/">https://www.addgene.org/39296/</a></td>
</tr>
<tr>
<td>(to build kanMX6-Ty3 and kanMX6-mse(-269)Δ::AAAAA-TT(-163)-mse(-79)Δ::AAAAA-Ty3)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>pLZL2422</td>
<td>(Larsen et al., 2008)</td>
<td>N/A</td>
</tr>
<tr>
<td>(to produce Gag3 protein to purify α-VLP from crude serum)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(to build ty3Δ::URA3MX (popin) prior to popout)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>pFA6a-His3MX6</td>
<td>(Longtine et al., 1998)</td>
<td><a href="https://www.addgene.org/41596/">https://www.addgene.org/41596/</a></td>
</tr>
<tr>
<td>(to build MSE-His3MX6-Ty3)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>pUB976</td>
<td>(Sawyer et al., 2019)</td>
<td>N/A</td>
</tr>
<tr>
<td>(to make his3::pRIM4-OsTIR1::His3MX6 and his3::pGPD1-OsTIR1::His3MX6)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Software and Algorithms

- **R v3.6.1**
  - R Core Team
  - [https://www.R-project.org/](https://www.R-project.org/)

- **R studio v1.1.383**
  - R Core Team
  - [http://www.rstudio.com/](http://www.rstudio.com/)

- **STAR v020201**
  - (Dobin et al., 2013)
  - [https://github.com/alexdobin/STAR](https://github.com/alexdobin/STAR)
<table>
<thead>
<tr>
<th>REAGENT or RESOURCE</th>
<th>SOURCE</th>
<th>IDENTIFIER</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ving software v beta1.1</td>
<td>(Descrimes et al., 2015)</td>
<td><a href="http://vm-gb.curie.fr/ving/">http://vm-gb.curie.fr/ving/</a></td>
</tr>
<tr>
<td>Bowtie2 v2.3.4 (yeast analysis) and v2.2.6 (mammals analysis)</td>
<td>(Langmead and Salzberg, 2012)</td>
<td><a href="https://anaconda.org/bioconda/bowtie2/files">https://anaconda.org/bioconda/bowtie2/files</a></td>
</tr>
<tr>
<td>SAMtools v1.6</td>
<td>(Li and Durbin, 2009)</td>
<td><a href="http://www.htslib.org/">http://www.htslib.org/</a></td>
</tr>
<tr>
<td>Picard tool MarkDuplicates v 2.23.3 (yeast analysis) and v2.17.11 (mammals analysis)</td>
<td>Broad Institute</td>
<td><a href="https://broadinstitute.github.io/picard/">https://broadinstitute.github.io/picard/</a></td>
</tr>
<tr>
<td>Deeptools bamcoverage</td>
<td>(Ramírez et al., 2016)</td>
<td><a href="https://deeptools.readthedocs.io/en/develop/content/tools/bamCoverage.html">https://deeptools.readthedocs.io/en/develop/content/tools/bamCoverage.html</a></td>
</tr>
<tr>
<td>Yeast ChIP normalization custom script</td>
<td>Available upon request</td>
<td>N/A</td>
</tr>
<tr>
<td>bpeaks v1.2</td>
<td>(Merhej et al., 2014)</td>
<td><a href="https://cran.r-project.org/web/packages/bPeaks/index.html">https://cran.r-project.org/web/packages/bPeaks/index.html</a></td>
</tr>
<tr>
<td>IGV v2.8.9</td>
<td>(Thorvalsdóttir et al., 2013)</td>
<td><a href="http://software.broadinstitute.org/software/igv/">http://software.broadinstitute.org/software/igv/</a></td>
</tr>
<tr>
<td>Fiji software</td>
<td>(Schindelin et al., 2012)</td>
<td><a href="https://fiji.sc/">https://fiji.sc/</a></td>
</tr>
<tr>
<td>TrimGalore 0.4.1</td>
<td>The Babraham Institute</td>
<td><a href="https://github.com/FelixKrueger/TrimGalore">https://github.com/FelixKrueger/TrimGalore</a></td>
</tr>
<tr>
<td>Tophat v2.1.0</td>
<td>(Kim et al., 2013)</td>
<td><a href="https://ccb.jhu.edu/software/tophat/index.shtml">https://ccb.jhu.edu/software/tophat/index.shtml</a></td>
</tr>
<tr>
<td>EdgeR</td>
<td>(Robinson et al., 2010)</td>
<td><a href="http://bioconductor.org/packages/release/html/edgeR.html">http://bioconductor.org/packages/release/html/edgeR.html</a></td>
</tr>
<tr>
<td>Htseq-count v0.11.2</td>
<td>(Anders et al., 2015)</td>
<td><a href="https://htseq.readthedocs.io/en/release_0.11.1/count.html">https://htseq.readthedocs.io/en/release_0.11.1/count.html</a></td>
</tr>
<tr>
<td>bedtools v2.26.0</td>
<td>(Quinlan and Hall, 2010)</td>
<td><a href="https://github.com/arq5x/bedtools2/releases/tag/v2.26.0">https://github.com/arq5x/bedtools2/releases/tag/v2.26.0</a></td>
</tr>
<tr>
<td>MACS v2.1.1</td>
<td>(Zhang et al., 2008)</td>
<td><a href="https://github.com/macs3-project/MACS">https://github.com/macs3-project/MACS</a></td>
</tr>
<tr>
<td>UCSC liftOver tool</td>
<td>(Haeussler et al., 2019)</td>
<td><a href="https://genome.ucsc.edu/cpibin/hgLiftOver">https://genome.ucsc.edu/cpibin/hgLiftOver</a></td>
</tr>
</tbody>
</table>

**Other**

<table>
<thead>
<tr>
<th>REAGENT or RESOURCE</th>
<th>SOURCE</th>
<th>IDENTIFIER</th>
</tr>
</thead>
<tbody>
<tr>
<td>DeltaVision Elite Microscope</td>
<td>GE Healthcare</td>
<td>Prod #: 29065728</td>
</tr>
<tr>
<td>FastPrep-24</td>
<td>MP biomedicals</td>
<td>Cat# 6004500</td>
</tr>
<tr>
<td>Criterion Vertical Electrophoresis Cell</td>
<td>Bio-Rad</td>
<td>Cat# 1655601</td>
</tr>
<tr>
<td>Trans-Blot Turbo Transfer System</td>
<td>Bio-Rad</td>
<td>Cat# 1704150</td>
</tr>
</tbody>
</table>
Figure 1

A

Ty3 mRNA
Northern probes:

Gag3 protein

Gag3-Pol3 polyprotein

Ty3 (estimate: 5kb)
18S shade (1.8 kb)
-25S shade (3.4 kb)

Ty3 (5052 bp)

B

Genomic position on chr. XIV

C

Ty3/Ty3

ty3Δ/ty3Δ

starvation
S, G2
MI
MII
spores

probe Ty3

Methylene blue

E

strain:

SK1 lab strain
OS673 wild strain

F

counts

Ty3/Ty3

ty3Δ/ty3Δ

G

% cells

SK1
OS673

1 nucleus
2 nuclei
3/4 nuclei

Ty1 19 0.02
Ty2 5 39.67
Ty3 1 2.98
Ty4 0 0.00
Ty5 1 0.00
Figure 2

A Wild type

Sporulation

Time (h)

0 30 45 60 15 30 45 60 75 90

Probe Ty3

Methylene blue

NDT80-IN

Time (h)

0 30 45 60 15 30 45 60 75 90

Probe Ty3

Methylene blue

C Ndt80 binding sites (CACAAA):

mutated binding sites (AAAAAA):

Scale: 500 bp

D NDT80-IN

WT

Mutated binding sites

Time (h)

0 6 7 8 9 10

Probe Ty3

Methylene blue

E NDT80

NDT80-3V5

Time (h)

0 3 5 6

IB: α-V5

IB: α-Pgk1

F SK1 chr XIV

Loci

YNL033W

TY3

YNL035C

G SK1 chr VIII

Loci

YHR123W

YHR124W

tRNA

IB: α-V5

IB: α-Pgk1
Figure 3

A

- Ndt80 target genes in SK1 (n=260)
- non-Ndt80 target genes in SK1 (n=5291)

B

- All Ty3/Gypsy in 13 genomes (n=518)
  - full-length and solo-LTRs

Distribution Ndt80 motif:

- 1000 bootstraps
- mean±SD
- outlier boundaries

- observed
**Figure 4**

**A**

<table>
<thead>
<tr>
<th>time (h)</th>
<th>0</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
</tr>
</thead>
<tbody>
<tr>
<td>probe Ty3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Methylene blue</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IB: α-VLP</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>37 kDa</td>
</tr>
<tr>
<td>IB: α-Pgk1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>37 kDa</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>NDT80-IN</th>
<th>Ty3</th>
<th>Ty3Δ</th>
<th>Ty3</th>
<th>Ty3</th>
<th>Ty3</th>
<th>Ty3</th>
</tr>
</thead>
<tbody>
<tr>
<td>MATa</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MATa</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**B**

<table>
<thead>
<tr>
<th>strain</th>
<th>meiosis (2n)</th>
<th>mitosis (n)</th>
</tr>
</thead>
<tbody>
<tr>
<td>NDT80-IN</td>
<td>Ty3 ty3Δ</td>
<td>Ty3 ty3Δ</td>
</tr>
<tr>
<td>time (h)</td>
<td>8 9 10 11</td>
<td>3 4 5 6</td>
</tr>
<tr>
<td>NDT80 induction</td>
<td>2 3 4 5</td>
<td>2 3 4 5</td>
</tr>
<tr>
<td>probe Ty3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Methylene blue</td>
<td></td>
<td></td>
</tr>
<tr>
<td>IB: α-VLP</td>
<td></td>
<td></td>
</tr>
<tr>
<td>IB: α-Pgk1</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Ty3 mRNA</th>
<th>- 3.4 kb</th>
<th>- 1.8 kb</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ty3 mRNA</td>
<td>- 3.4 kb</td>
<td>- 1.8 kb</td>
</tr>
</tbody>
</table>

- 1.8 kb
- 3.4 kb
- 1.8 kb
- 3.4 kb
- 1.8 kb
- 3.4 kb
- 1.8 kb
- 3.4 kb

<table>
<thead>
<tr>
<th>Gag3 / Pgk1</th>
<th>0 4 8 12</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ty3 / 25S</td>
<td>0 4 8 12</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Gag3 / Pgk1</th>
<th>0 4 8 12</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ty3 / 25S</td>
<td>0 4 8 12</td>
</tr>
<tr>
<td>strain</td>
<td>NDT80-IN NDT80-IN NDT80-IN NDT80-IN</td>
</tr>
<tr>
<td>--------</td>
<td>-----------------------------------</td>
</tr>
<tr>
<td>RIM4 Ty3</td>
<td></td>
</tr>
<tr>
<td>Rim4-deg Ty3</td>
<td></td>
</tr>
<tr>
<td>RIM4 ty3Δ</td>
<td></td>
</tr>
<tr>
<td>NDT80-IN</td>
<td></td>
</tr>
<tr>
<td>NDT80-IN</td>
<td></td>
</tr>
</tbody>
</table>

- **Ty3 mRNA**
  - 3.4 kb
  - 1.8 kb

- Methylene blue
  - <25S rRNA
  - <18S rRNA

- IB: α-VLP
  - 37 kDa

- IB: α-Pgk1
  - 37 kDa

- IB: α-FLAG
  - 100 kDa

- IB: α-v5
  - 100 kDa

- Gag3 / Pgk1
  - Ty3 / 25S
Figure 6

A. RNA abundance (FPKM) for different cell types and two genotypes in mouse spermatocyte RNAseq.

B. Moap1 transcripts and Gypsy-like region in M. musculus chr12.

C. Pnma1 transcripts and Gypsy-like region in M. musculus chr12.

D. RNA abundance (CPM) for Moap1 and Pnma1 in mouse spermatocyte RNAseq.

E. RNA abundance (CPM) for Moap1 and Pnma1 in mouse P14 testis RNAseq.

F. RNA abundance (CPM) for MOAP1 and PNMA1 in human fetal ovary α-DAZL RIP.
SUPPLEMENTAL FIGURE LEGENDS

Figure S1: Ty3 mRNA is full-length and is not required for meiotic progression, related to Figure 1

(A) Heterozygous Ty3/ty3Δ diploid (B838, blue) and homozygous ty3Δ/ty3Δ diploid (B837, yellow) do not exhibit spore viability defect compared to wild type Ty3/Ty3 (B47, purple). Diploid strains were sporulated for the experiment shown in Figure 1C-D, and sporulation cultures were kept over week-end at 30°C with shaking, and a sample was collected for tetrad dissection. Spore viability was assessed as the number of colony-forming spores after 2 days on YPD at 30°C, relative to the number of dissected spores (n is indicated above bars).

(B) Homozygous ty3Δ/ty3Δ diploid (B837, yellow) do not exhibit growth defect compared to wild type Ty3/Ty3 (B47, purple). Mitotic growth curves were determined from OD₆₀₀ measurements of cells grown in YPD using a Tecan plate reader with shaking (n = 4, error bars indicate SEM).

(C) The Ty3 locus produces full-length mRNA. RNA samples from either wild type diploid SK1 meiosis (B47, purple) or NDT80-IN diploid SK1 meiosis (B48, blue) were analyzed using the ‘CA’, ‘PR’, ‘RT’, and ‘IN’ probes (see Figure 1). In each case, the top band corresponding to the full-length mRNA is observed.

Figure S2: Ty3 is transcriptionally activated by Ndt80, related to Figure 2

(A) Ty3 full-length mRNA is not expressed when NDT80 is not induced in meiosis. Meiosis was induced in two wild type strains (B371, pink and B47, purple) and NDT80-IN (B48) cells. NDT80-IN cells were either released from the G2 block by addition of 1 µM β-estradiol at 6 h (blue) or not induced (gray). Shown is Northern blot analysis of Ty3 mRNA with annotation same as Figure 1C.

(B) Meiotic progression and Ty3 full-length mRNA quantification in wild-type, induced, and uninduced NDT80. Dashed lines/left axis: Kinetics of meiotic divisions determined by α-tubulin immunofluorescence analysis (quantification of cells with one or two division spindles). Solid lines/right axis: Quantification of full-length Ty3 mRNA.

Shown are representative results of 2 biological replicates.
Figure S3: Meiotic Ty3 transcription is not governed by elements within the 5' LTR, related to Figure 2

(A) Diagram of endogenous and interrupted Ty3 upstream context. Blue triangles indicate Ndt80 binding sites (-79 and -269 bp upstream Ty3 in the endogenous context, shifted to -1,446 and -1,636 bp upon His3MX insertion). (B) Sporulation was induced in NDT80-IN strains containing the wild type Ty3 configuration (B48, blue), or with an integrated His3MX cassette that separates Ty3 from its upstream genetic context (B1436, orange) as in A. Ty3 and rRNA RNA levels are shown.

Figure S4: Post-transcriptional repression of Ty3 is specific to meiosis, related to Figure 4

(A) The absence of Gag3 protein in meiosis is not due to RNA degradation or proteasomal degradation. Meiosis was induced in four experimental conditions: wild type (B48), proteasome defective strain rpn6-1 (B207), wild type + MG-132, and pdr5Δ (B171) + MG-132. Proteasome inhibitor MG-132 was added at 6 h to 50 µM. Meiotic cells were released from the G2 block by addition of 1 µM β-estradiol at 6 h. Mitotic induction of NDT80 (with and without Sum1-deg) was conducted as in Figure 3B. Ty3 mRNA (black arrow) and rRNA (white arrows) levels and Gag3 and Pgk1 (loading control) protein levels are shown. Shown are representative results of 2 biological replicates.

(B) Sum1 depletion in the context of ectopic NDT80 expression in mitosis (refers to Figure 3B, but with all the time points). To express Ty3 in both meiosis and mitosis, we expressed NDT80 in sporulating diploid cells (wild type B48, ty3Δ B839) and in haploid mitotically dividing cells with and without the Sum1 transcriptional repressor (wild type B477, Sum1-deg B829, and ty3Δ B827). Gag3, Sum1 (tagged with 3V5) and Pgk1 (loading control) protein levels are shown. Asterisks show unspecific bands (e.g. band in α-VLP present in Ty3Δ). Shown are representative results of ≥5 biological replicates.

(C) Ty3 dimer formation upon Sum1 depletion in the context of ectopic NDT80 expression in mitosis compared to meiosis (refers to Figure 3B and S3B, but with selection of time points). To express Ty3 in both meiosis and mitosis, we expressed NDT80 in sporulating diploid cells (wild type B48, ty3Δ B839) and in haploid mitotically dividing cells with and without the Sum1 transcriptional repressor (wild type B477, Sum1-deg B829, and ty3Δ B827). In this case, samples were not denatured prior to gel loading (see methods) allow observation of the Ty3
RNA dimer formed within the Virus-Like-Particle. Ty3 dimer (pink arrow), Ty3 mRNA (black arrow), and rRNA (white arrows) are shown. Shown are representative results of 2 biological replicates.

**Figure S5:** Gag3 protein accumulates in rim4 truncation mutants and upon Rim4 depletion in meiosis, related to Figure 5

Meiosis was induced in four NDT80-IN strains: wild type (B48), Rim4-deg (B1022, Rim4 depletion induced by auxin addition to 1 mM at t = 5.5 h), rim4Δ138C (A33845), and rim4Δ271C (A33848). Meiotic cells were released from the G2 block by addition of 1 µM β-estradiol at 6 h. Mitotic induction of NDT80 with Sum1-deg was conducted as in Figure 3B. Gag3, Rim4 (wild type and truncations mutants are tagged with 3V5, Rim4-deg is tagged with FLAG), and Pgk1 (loading control) protein levels are shown. Asterisks show unspecific bands (e.g. band in α-FLAG present in no-tag samples). Shown are representative results of 2 biological replicates.

**Figure S6:** Mammalian Gypsy-derived genes are activated by meiotic transcription factors and regulated by meiotic RNA-binding proteins, related to Figure 6

(A) RNAseq from indicated cell types from adult mouse testes was mapped to the mouse genome, and Gypsy family derived elements of the same nature were aggregated together (at the level of RepName using Repeatmasker annotations). The FPKM, fragments per kilobase of transcript per million mapped reads for each Gypsy family elements are shown. Biological replicates = 3 for each stage.

(B) RNAseq data from indicated tissue (Marin et al., 2017) and cell types (Lesch et al., 2016) from opossum was analyzed. The FPKM, fragments per kilobase of transcript per million mapped reads for the M-PNMA gene is shown. Biological replicates = 2 for each stage.

(C, D) Plots show the distance between each individual MamGyp-int elements (C), or the transcriptional start site of a transcript encompassing a given MamGyp-int element (D), and the nearest Stra8 ChIP-seq peak and Mybl1 ChIP-seq peaks. Moap1 and Pnma1 MamGyp-int elements are highlighted in blue and orange.

(E - G) Analysis of retrotransposon sequences in anti-DAZL RNA immunoprecipitation (RIP) from human foetal ovaries (biological replicates = 3). (E) Reads were aggregated at the level of RepName using Repeatmasker annotations, and assessed for differential abundance in paired anti-DAZL and IgG RIP. Retrotransposons belonging to the Gypsy family are plotted. Asterisk
indicates FDR-adjusted p-value < 0.01 determined by Student’s t-test. CPM, counts per million mapped reads. (F) Five retrotransposons enriched in Human Fetal Ovary α-DAZL RIP, including MamGyp-int group (in orange, see panel G). (G) Human Fetal Ovary α-DAZL RIP enrichment levels for retrotransposons shown in F (in orange) for MamGyp-int and other LTR retrotransposons.
Table S1: Laboratory *S. cerevisiae* strains used in this study, related to Figures 1-5 and Figures S1-S5

Table S2: Wild *S. cerevisiae* strains used in this study, related to Figure 1

For these strains, the number of Ty3 insertions in the genome has been estimated by depth of coverage at a single Ty3 locus compared to the rest of the genome. Sporulation efficiency was determined by number of asci formed 24h and 72h after sporulation induction. Presence of full-length Ty3 mRNA was detected by Northern blot as in Figure 1. Standardized names and some relevant information from (Peter et al., 2018) were included.
Figure S1

A

B

C

Ty3 mRNA

Northern probes:

GAG3

POL3

5052 bp

Ty3/Ty3

ty3Δ/ty3Δ

0 2 5 1 4 6

Methylene

blue

0.0

1.0

1.5

Time (min)

OD₆₀₀

0 200 400 600 800 1000 1200

A

C

probe CA

probe PR

probe RT

probe IN

0 200 400 600 800 1000

0.0

0.5

1.0

1.5

0 2 5 1 4 6

0 2 5 1 4 6

0 2 5 1 4 6

0 2 5 1 4 6

Ty3 in meiosis (h)

wild type

NDT80-IN

Methylene blue
Figure S2

A

- wt RIM4-3v5
- wt
- NDT80-IN induced at 6h
- NDT80-IN uninduced

<table>
<thead>
<tr>
<th>Time (h)</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
</tr>
</thead>
<tbody>
<tr>
<td>Probe Ty3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Methylene blue</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

B

- Cells in division
- Ty3 mRNA levels

- Wild type
- NDT80-IN
- NDT80-IN uninduced

% cells in meiotic divisions

Ty3 mRNA (a.u.)

Time in meiosis (h)
Figure S3

A
Endogenous upstream genetic context:

\[ \text{LTR}_o \text{Ty3} \]

Altered upstream genetic context:

\[ \text{LTR}_o \text{Ty3} \]

scale: 500 bp

Ndt80 binding motif

B

strain
time (h)
0 6 7 8 9 10

probe Ty3

Methylene blue

scale:

- 1.8 kb
- 3.4 kb

\[ \text{Ty3 mRNA} \]
\[ \text{18S rRNA} \]
\[ \text{25S rRNA} \]
### Figure S4

#### Part A

<table>
<thead>
<tr>
<th>Strain</th>
<th>NDT80-IN</th>
<th>NDT80-IN rpn6-1</th>
<th>NDT80-IN</th>
<th>NDT80-IN pdr5Δ</th>
</tr>
</thead>
<tbody>
<tr>
<td>Time (h)</td>
<td>0 7 8 9 10</td>
<td>0 7 8 9 10</td>
<td>0 7 8 9 10</td>
<td>0 7 8 9 10</td>
</tr>
<tr>
<td>Time post degron induction (h)</td>
<td>- - - - -</td>
<td>1 2 3 4</td>
<td>- - - - -</td>
<td>- - - - -</td>
</tr>
<tr>
<td>Time post NDT80 induction (h)</td>
<td>- - - - -</td>
<td>- - - - -</td>
<td>1 2 3 4</td>
<td>- - - - -</td>
</tr>
<tr>
<td>Time post MG-132 addition (h)</td>
<td>- - - - -</td>
<td>- - - - -</td>
<td>- - - - -</td>
<td>- - - - -</td>
</tr>
</tbody>
</table>

#### Probe Ty3

- Ty3 mRNA: 3.4 kb, 1.8 kb
- Ty3 RNA dimer

#### Methylene blue

- <25S rRNA
- <18S rRNA

#### IB: α-VLP

- 37 kDa

#### IB: α-Pgk1

- 50 kDa, 37 kDa

#### Part B

<table>
<thead>
<tr>
<th>Strain</th>
<th>NDT80-IN</th>
<th>NDT80-IN Sum1-deg</th>
<th>NDT80-IN ty3Δ</th>
</tr>
</thead>
<tbody>
<tr>
<td>Time (h)</td>
<td>0 1 2 4 5 6</td>
<td>0 1 2 4 5 6</td>
<td>0 1 2 4 5 6</td>
</tr>
<tr>
<td>Time post degron induction (h)</td>
<td>- - - - -</td>
<td>1 2 4 5 6</td>
<td>- - - - -</td>
</tr>
<tr>
<td>Time post NDT80 induction (h)</td>
<td>- - - - -</td>
<td>- 1 2 4 5</td>
<td>- - - - -</td>
</tr>
</tbody>
</table>

#### α-VLP

- 37 kDa

#### α-v5

- 250 kDa

#### α-Pgk1

- 150 kDa, 37 kDa

#### Part C

<table>
<thead>
<tr>
<th>Strain</th>
<th>NDT80-IN Ty3</th>
<th>NDT80-IN ty3Δ</th>
<th>NDT80-IN Sum1-deg Ty3</th>
<th>NDT80-IN α-factor +</th>
</tr>
</thead>
<tbody>
<tr>
<td>Time (h)</td>
<td>0 7 8 9 10 11 12</td>
<td>1 2 3 4 5 6</td>
<td>1 2 3 4 5 6</td>
<td>- - - - -</td>
</tr>
<tr>
<td>Time post degron induction (h)</td>
<td>- - - - -</td>
<td>7 8 9 10 11 12</td>
<td>1 2 3 4 5 6</td>
<td>- 2 3 4 5 6</td>
</tr>
<tr>
<td>Time post NDT80 induction (h)</td>
<td>- - - - -</td>
<td>- 7 8 9 10 11 12</td>
<td>- - - - -</td>
<td>- - - - -</td>
</tr>
</tbody>
</table>

#### Probe Ty3

- Ty3 RNA dimer
- Ty3 mRNA: 3.4 kb, 1.8 kb

#### Methylene blue

- <25S rRNA
- <18S rRNA
Click here to access/download

Supplemental Videos and Spreadsheets
Ty3 table S1.xlsx