RIBOSOMAL PROTEINS AND RIBOSOME BIOGENESIS FACTORS AFFECT

MOVEMENT AND COPY NUMBER CONTROL OF THE TY1 RETROTRANSPOSON

IN *SACCHAROMYCES CEREVISIAE*

by

HYO WON AHN

(Under the Direction of David J. Garfinkel)

ABSTRACT

Retrotransposons are mobile genetic elements that replicate via an RNA intermediate and are similar to retroviruses. The long terminal repeat retrotransposon Ty1 is the most abundant transposon in the widely used model organism *Saccharomyces cerevisiae*. Ty1 contains two genes; *GAG*, which encodes the capsid and *POL*, which encodes protease (PR), integrase (IN), and reverse transcriptase (RT). Retrotransposon movement impacts the host genome by insertional mutagenesis or longer-range effects on host gene expression. Chromosomal rearrangements can also occur by recombination between retrotransposons. Host organisms have evolved different ways to control the level of retrotransposition such as DNA methylation and RNA interference, but those pathways are absent in *S. cerevisiae*. For Ty1, a novel

copy number control (CNC) mechanism blocks Ty1 movement by inhibiting virus-like particle (VLP) assembly, function and protein processing. Interestingly, CNC is mediated by the self-encoded restriction factor p22, which is derived from *GAG* and encoded by Ty1 internally-initiated (Ty1i) RNA. While p22 is a key player in CNC, cellular genes or environmental factors that affect p22 have not been studied. We hypothesized that a subset of Ty1 cofactors affect CNC, and these genes can be identified by CNC-specific secondary screens of Ty1 cofactor mutants. We identified rRNA processing/ribosome assembly factor (*loc1Δ*) and ribosomal subunit proteins (*rps0bΔ* and *rpl7aΔ*) that exhibit a severe decrease in Ty1 mobility in strains undergoing CNC. Further examination revealed that the level of p22 increases in these mutants. Modest increases in the level of Ty1i RNA relative to Ty1 mRNA and the half-life of p22 also affect CNC in a *loc1Δ* mutant. Furthermore, a *loc1Δ* mutant exhibits defects in Gag multimerization, Gag complex assembly, and packaging of Ty1 mRNA in VLPs. Strikingly, cytoplasmic foci containing Ty1 products that are sites for VLP assembly are abolished in a *loc1Δ* mutant. In summary, my work emphasizes the importance of ribosome biogenesis in mediating host control of Ty1, and may identify a novel stress response that alters VLP assembly.

INDEX WORDS: transposable elements, retrotransposon, retroelement, retrovirus, budding yeast, Saccharomyces, Ty1, LTR, restriction factor, copy number control, ribosome biogenesis, ribosomal protein, Loc1

RIBOSOMAL PROTEINS AND RIBOSOME BIOGENESIS FACTORS AFFECT

MOVEMENT AND COPY NUMBER CONTROL OF THE TY1 RETROTRANSPOSON

IN *SACCHAROMYCES CEREVISIAE*

by

HYO WON AHN

B.S., Konkuk University, South Korea, 2005

M.S., Texas A&M University, 2008

A Dissertation Submitted to the Graduate Faculty of The University of Georgia in Partial

Fulfillment of the Requirements for the Degree

DOCTOR OF PHILOSOPHY

ATHENS, GEORGIA

2018

RIBOSOMAL PROTEINS AND RIBOSOME BIOGENESIS FACTORS AFFECT

MOVEMENT AND COPY NUMBER CONTROL OF THE TY1 RETROTRANSPOSON

IN *SACCHAROMYCES CEREVISIAE*

by

HYO WON AHN

Major Professor: David J. Garfinkel

Committee: Walter Schmidt
                    Takahiro Ito
                    Michelle Momany

Electronic Version Approved:

Suzanne Barbour
Dean of the Graduate School
The University of Georgia
December 2018

DEDICATION

      This dissertation is dedicated to my dear mother Gil-Bok Choi and brother Seok-Joon Ahn.

ACKNOWLEDGEMENTS

TABLE OF CONTENTS

Suresh *et al*. 2015. Ribosomal protein and biogenesis factors affect

multiple steps during movement of the *Saccharomyces cerevisiae*

Ty1 retrotransposon. Mobile DNA. 2015 Dec 8;6:22

CHAPTER 1

INTRODUCTION AND LITERATURE REVIEW

In the late 1940s, the groundbreaking work of Barbara McClintock amazed the scientific community. She found a DNA segment that can move from one position to another location in the maize genome; these mobile DNA elements were later termed transposons. Insertion of the DNA transposon Activator (*Ac*) - Dissociation (*Ds*) was capable of causing chromosome breakage which results in a kernel color variegation phenotype [1]. Although it took decades for McClintock's discoveries to be fully accepted, by the 1970s scientists found transposons in other organisms, and made progress in understanding transposition of bacterial transposons at the molecular level. Now we appreciate that these genetic elements are found in almost every organism [2]. Although the pioneering work of McClintock eventually resulted in her winning the Nobel Prize in 1983, transposon research has been underestimated due to the belief that one gene should have a function, and the function of transposons was not readily clear besides "jumping". Indeed, transposons can be viewed as clear-cut examples of "selfish or junk DNA" [3]. Now, with increased understanding and advancement in technologies, especially DNA sequencing, the world of transposons has shed light on many aspects of biology, varying from genome evolution to disease. There are two broad classes of transposable elements categorized by their mode of transposition; DNA transposons that use a "cut & paste" mechanism, and retrotransposons that use a "copy & paste"

mechanism involving an RNA replication intermediate. Retrotransposons are of particular interest due to similarities with retroviruses and evidence of retrotransposon activity in eukaryotic host organisms, including retrotransposon activity implicated in human health.

<u>Retrotransposons in humans: LINE-1 and HERV</u>

Retrotransposons are genetic elements that replicate in the host organism involving reverse transcription of an RNA intermediate to a cDNA copy and insertion of the cDNA at a new site in the genome [4]. Retrotransposons can be classified as autonomous or non-autonomous [5]. Autonomous elements carry critical factors such as enzymes necessary for their replication cycle. Conversely, non-autonomous elements rely on at least one factor from other autonomous elements for their replication. For example, human endogenous retroviruses (HERVs) and long interspersed elements (LINEs) are autonomous elements, and short interspersed elements (SINEs) are non-autonomous elements. Retrotransposons can be also classified as long terminal repeat (LTR) and non-LTR retrotransposons [6]. HERVs are LTR retrotransposons and LINEs, SINEs, and processed pseudogenes are non-LTR retrotransposons. Not only differentiated by the presence of an LTR, they also differ in their replication mechanisms. LTR retrotransposons replicate in a fashion similar to retroviruses, while non-LTR retrotransposons such as LINE-1 are mobilized by target-primed reverse transcription (TPRT). In this section, the human LINE-1 element and HERVs are selected from non-LTR transposons and LTR retrotransposons, respectively, and described further.

A mammalian non-LTR retrotransposon, LINE-1 (often further abbreviated as L1) comprises 17% of the human genome [7]. The majority of LINE-1 elements in humans

are not functional due to 5′ truncation, inversion/deletion, or point mutations that result in defective LINE-1-encoded proteins [7, 8]. However, a subset of LINE-1 elements are capable of transposing and are referred to "hot" LINE-1 elements, which contribute to most of L1 retrotransposition events in human [9]. Autonomous LINE-1 elements are ~6kb in length [10, 11], and contain two open reading frames (ORF1 and ORF2) bracketed by a 5′ UTR, and a 3′ UTR with a poly (A) tail (Figure 1.1) [10]. Human ORF1 encodes a 40 kDa RNA binding protein (ORF1p) with nucleic acid chaperone activity [12-15]. Human ORF2 encodes a 150 kDa protein (ORF2p) [16, 17] with endonuclease (EN) [18] and reverse transcriptase (RT) activities [19]. ORF1p and ORF2p are both required for retrotransposition of LINE-1 [18, 20]. The LINE-1 5′ UTR also contains a primate-specific ORF in the antisense orientation named ORF0 [21]. LINE-1 mRNA is transcribed from an active chromosomal LINE-1 element and exported to the cytoplasm (Figure 1.1). After translation of ORF1p and ORF2p from LINE-1 mRNA template, these two proteins bind LINE-1 mRNA and form cytoplasmic foci containing LINE-1 ribonucleoprotein particles (RNP). The LINE-1 RNP minimally consists of LINE-1 mRNA, multiple ORF1p trimers, and at least one molecule of ORF2p. Interestingly, LINE-1 proteins preferentially act *in cis* on the encoding transcript [17, 22]. LINE-1 RNPs can be also formed *in trans* by ORF1p and/or ORF2p bound to cellular RNAs such as coding gene mRNAs and SINE RNAs. The LINE-1 RNP enters the nucleus via an unknown mechanism. LINE-1 uses TPRT which is similar to the process used in the non-LTR retrotransposon R2Bm of *Bombyx mori* [23]. LINE-1 EN, encoded in ORF2p, makes a single-strand nick in genomic DNA at the target sequence (5′ - TTTT/A- 3′). This step exposes a 3′ -OH group that serves as the primer for reverse transcription.

First-strand cDNA synthesis occurs using LINE-1 mRNA as template and LINE-1 RT encoded in ORF2p [18, 24]. The remaining steps of TPRT have not yet been elucidated (top-strand cleavage, second-strand LINE-1 cDNA synthesis, and repair of the DNA ends), but ultimately lead to the insertion of the new copy of LINE-1 element into the host genome.

Endogenous retroviruses are present in the genome of all vertebrates studied so far and thus are considered remnants of ancestral infections that entered the germline [25, 26]. HERVs comprise approximately 8% of human genome with around 700,000 loci [7, 27-29]. As the name reflects, HERVs resemble simple retroviruses in terms of genomic organization. When a viral genome is incorporated into the host genome and is able to replicate within the host genome, it is called a provirus. The organization of the HERV provirus, in general, shares similarities to retroviruses (Figure 1.2). Two LTRs flank four major ORFs: *gag* (encoding matrix and capsid (retroviral core protein)), *pro* (protease), *pol* (reverse transcriptase and integrase), and *env* (envelope which is a surface protein responsible for infection). HERVs generally contain a nonfunctional *env* or lack *env* completely; therefore, these elements do not have the ability to escape and infect another cell [30] because HERVs have undergone extensive mutations and deletions upon entry to the human genome. However, younger elements, especially the HERV-K (HML-2) family, are better preserved [31]. No one has yet reported the identification of an HERV element that can replicate [32]. However, reconstructed HERV-K proviruses based on phylogenetic analyses are infectious in cultured human cells [33, 34]. Unlike in human, other mammalian endogenous retroviruses are active in the host genome, such as Koala retrovirus subgroup A and B (KoRV-A and -B) that is linked to the occurrence

of lymphomas, porcine endogenous retroviruses (PERV) that complicates

xenotransplantation, and the mouse Intracisternal A particle (IAP) that is linked to

response to xenobiotics and carcinogenesis [35-37].

Effects of retrotransposons on the host

Derepression of retrotransposon activities, insertion of retrotransposons to new

positions in the host genome, expression of retrotransposon-derived products,

endogenization of retroviral sequence, and domestication of retroviral/retrotransposon

sequences cause varied outcomes on the host. Here, I present representative positive

(utilization of retrotransposon-derived sequences for new cellular function) and negative

(human disease) impacts of retrotransposons on the host with examples from mammals.

Ancient retroviral sequences that entered the germline of the host genome and were

inherited in the host population are named endogenous viral elements (EVEs). Host

organisms have co-opted EVEs to provide biological functions for host defense,

embryonic development, and more recently for immune response and neuronal function

[38-41].

Host organisms evolved retroelement-derived restriction factors to protect

themselves from retroviruses through direct interaction between retroviral and cellular or

domesticated proteins. Mainly co-option of envelope or Gag derived sequences has

been reported [39]. For example, Friend-virus-susceptibility-1 (Fv1) gene is derived from

the Gag region of an ancient MERVL element, and blocks replication of murine

leukemia virus (MLV) [42, 43]. The gag protein of enJS56A1, one of the endogenous

Jaagsiekte sheep retrovirus (enJSRV) proviruses, acts as a restriction factor for JSRV,

which causes lung cancer [44].

Domestication of retrotransposon-derived sequence has recently gained attention in mammalian development. Aside from contradictory evidence supporting the idea that retrotransposons are activated or repressed during germ cell development and implantation, transcripts and protein products from retrotransposons are expressed in human embryos and embryonic stem cells [45, 46]. Elevated transcription of HERV-H in embryonic stem cells affects pluripotency of human embryonic stem cells [47]. Expression of HERV-K transcripts as well as Gag and Rec (an accessory protein) during embryogenesis implicates their roles during early human development [48]. LINE-1 RNA is expressed in mouse embryonic stem cells, and LINE-1 ORF1-derived LINE-1 type Transposase Domain-containing 1 (L1TD1/ECAT11) gene in human embryonic stem cells play important roles in maintaining pluripotency [49-52]. HERV *env* encoded Syncytin proteins have essential functions during human placentation, fertilization of human embryos, and in reproductive organs (reviewed in [53]). The primate-specific MER41 family of endogenous retrovirus serves as interferon gamma-inducible binding sites that recruit the STAT1 transcription factor. MER41 regulates innate immune pathways by promoting expression of nearby genes such as AIM2 which activates inflammatory pathways [41]. The neuronal gene Arc derived from Gag of LTR-retrotransposons stores long-term information storage and mediates synaptic plasticity in mammalian brain [40].

Even though retrotransposons can positively impact the host genome, these elements also confer negative effects. Due to the nature of the "copy & paste" mechanism of their replication scheme, retrotransposons can cause insertional mutagenesis and genome rearrangement in the host genome. LINE-1 -mediated

6

germline-retrotransposition events are found in approximately every 1 in 250 deleterious human mutations [54]. LINE-1 retrotransposition has been implicated in at least 25 cases of human disease. Non-autonomous retrotransposons can also cause human disease. LINE-1 mediated retrotransposition of Alu and SVA SINE-elements is responsible for over 60 and at least 10 human disease-producing mutations, respectively [32, 55-59]. In a protein coding gene, mutations that alter gene expression can occur in exons, introns, or regulatory regions [60]. Also, retrotransposition can cause genome rearrangement in both the retrotransposon itself and genomic DNA. It could be harmless, if insertion happens at preexisting DNA lesions through process called EN-independent (ENi) retrotransposition [61, 62]. However, mispairing and unequal recombination of LINE-1 and Alu element sequences can result in disease by gene deletion or generating copy number variations [60, 63]. After the initial discovery by Kazazian and colleagues that LINE-1 causes hemophilia A due to mutagenic insertion affecting the coagulation factor VIII gene, relationships between retrotransposons and human disease have been continuously reported [64, 65]. Consequently, germ cells have evolved a variety of genome defense mechanisms such as DNA methylation and RNA interference via piRNAs to restrict retrotransposon activity and maintain genome stability across generations [66]. Elevated retrotransposon expression due to the loss of genome defense mechanisms is correlated with aberrant proliferation of male germ cells [67] and defects in spermatogenesis [68, 69] and oogenesis [70, 71] (also reviewed in [72]).

Retrotransposon activity results in somatically acquired insertions in cancer genomes and can be clinically significant mutagens [73]. While shown in other cancer

types, somatic LINE-1 and Alu insertions are frequently observed in the cancers of epithelial cell origin (colorectal, prostate, and ovarian), and colorectal tumors show the highest frequency of somatic LINE-1 insertions [74]. In addition, tumors of the gastrointestinal tract, namely esophageal squamous carcinomas and adenocarcinomas, gastric and small bowel tumors, hepatocellular carcinomas and PDACs, also show somatic LINE-1 retrotransposition [73]. However, it is unclear whether all insertions contribute to cancerous phenotypes. The most compelling examples of a causal relationship are when insertions affect previously identified oncogenes or tumor suppressor genes. For example, somatic insertion of a human LINE-1 element in the APC tumor suppressor gene disrupts function of the gene and contributes to colorectal cancer [75].

Somatic LINE-1 insertions may also contribute to additional diseases, although causal relationships need to be determined in many instances. The human LINE-1 - EGFP reporter system generates *de novo* insertions in neurons and their precursor cells [76], suggesting that endogenous LINE-1 insertion events may also in these cell types *in vivo*. LINE-1 and HERV are involved in neurological diseases such as multiple sclerosis, ataxia telangiectasia (reviewed in [77, 78]) and schizophrenia [79]. Moreover, increased DNA copies of a multiple sclerosis associated retrovirus (MSRV), a HERV-K family retrotransposon, are present in the genome of multiple sclerosis patients [80]. Innate immune activation and inflammatory stimuli promote transcription of endogenous retroviruses in mouse and human immune cells. For example, mouse B cell activation promotes murine leukemia virus (MLVs) transcription [81, 82]. Mutations in genes involved in nucleic acid metabolism such as three-prime repair exonuclease 1 (TREX1)

and SAM domain and HD domain 1 (SAMHD1) are linked to inflammatory disorder Aicardi-Goutières syndrome (AGS) and autoimmune disease systemic lupus erythematosus (SLE) patients [83]. The protein products of endogenous retroviruses serve as the antigen and stimulate host immune systems. Retroelement-encoded proteins may also be the putative antigens in chronic infection, autoimmunity, or adaptive immunity related to cancer [84].

Host mechanisms that restrict retrotransposon movement

Since retrotransposons can alter the host genome by insertional mutagenesis or genome rearrangement, hosts have evolved ways to deal with these "frenemies within" [85]. A variety of strategies target the inserted DNA or nucleic acid product of the retrotransposon. In this section, widely used pathways used by host organisms to repress retrotransposons will be discussed.

DNA methylation and certain histone modifications repress transcription of different classes of transposable elements. Cytosine methylation is an important DNA modification in plants, the filamentous fungus *Neurospora crassa*, and mammals that is correlated with repression of transposable elements [86, 87]. For example, CpG islands in the 5´ UTR of mammalian LINE-1 are heavily methylated in somatic cells and LINE-1 expression is repressed [88-91]. The global demethylation of genomic DNA reactivates transcription of transposable elements [92-94]. DNA cytosine-5-methyltransferase 3 (DNMT3) enzymes DNMT3A, DNMT3B, and DNMT3L are responsible for methylation of unmethylated CpGs in mammals [5]. In addition, histone H3 methylated at Lys9 (H3K9me2/me3) is the mark of transcriptional repression and enriched on retrotransposons. H3K9 methyltransferases EHMT2/G9a and SUV39H have been

9

implicated in the repression of retrotransposons [95-98].

Apolipoprotein B mRNA editing enzyme, catalytic polypeptide-like 3 (APOBEC3) enzymes are cytidine deaminases. Retroviruses including HIV encode viral infectivity factor (Vif), which inhibits APOBEC3 by promoting its degradation. In Vif-deficient HIV-1, APOBEC3G is packaged inside the virion and interferes with reverse transcription, deaminates cytosine in the nascent first-strand HIV cDNA to form uracil. Uracil substitutions in the cDNA cause dG to dA mutations to accumulate during second strand synthesis [99-101]. APOBEC3 family enzymes have been also shown to inhibit variety of retroviruses as well as retrotransposons, including Ty1 [102-105].

RNA interference (RNAi) also plays a role in retrotransposon control in both somatic cells and the germline. RNAi acts at the post-transcriptional level by causing degradation of retrotransposon transcripts. The RNAi pathway has the capability to inhibit any retrotransposon that generates a dsRNA that can serve as a substrate for the RNAi factors such as Dicer and the RNA-induced silencing complex (RISC) [6]. Microprocessor (a complex of Drosha and DGCR8) is shown to bind transcripts of LINE-1, Alu, and SVA, and cleave LINE-1 RNA *in vitro* [106]. LINE-1 mobility increases in DICER1 or AGO2 knockdown cell culture [107, 108]. In addition, a well-known RNAi system providing transposon control in the germline is based on piwi-interacting RNAs (piRNA). piRNAs are small RNAs of 24-31nts. Most piRNAs are expressed from genomic regions called piRNA clusters, which contain sequences derived from transposable elements [109-111]. PIWI proteins loaded with piRNAs, which are antisense to transposable element derived RNAs, are capable of cleaving target transposable element derived RNAs [111, 112].

Ty retrotransposons

The budding yeast *Saccharomyces cerevisiae* is well-established model organism for biomedical research and harbors five LTR retrotransposon families (Ty1-5). Most Ty (Transposon yeast) retrotransposons generally belong to the pseudoviridae group except Ty3, which belongs to the metaviridae group [113]. Two major differences in the genomic organization of pseudoviridae (Ty1) and metaviridae (Ty3) are: the order of integrase and reverse transcriptase enzymes in *POL*, and Ty3 *GAG3* ORF encodes capsid (CA) and nucleocapsid (NC) proteins compared to Ty1, which encodes a single Gag protein [114]. Ty1 is the most abundant retrotransposon with 32 copies per haploid genome in the reference strain S288c. Among these 32 Ty1 elements, 3 copies belong to the subfamily of Ty1 termed Ty1'. Ty1' differs from Ty1 by divergence in their *GAG* sequences. Ty2 (with 13 copies in the genome) is similar to Ty1, especially the *POL* and LTR sequences [115]. The majority of transposable elements in the *S. cerevisiae* are Ty1 and Ty2 elements. S288c contains 1-3 copies of Ty3, Ty4 and Ty5. Ty3 and Ty4 are active but Ty4 exhibits very low transposition activity due to generation of transcripts lacking the U3-R region, which is essential for reverse transcription [116]. The Ty5 element from *S. cerevisiae* is transpositionally inactive [117], however, Ty5 from *S. paradoxus* is active. This section will focus on Ty1, as it is the most abundant transposable element in many *S. cerevisiae* strains and is under copy number control (CNC). Due to the shared similarity with retroviral replication and genomic organization, Ty1 has served as an important tool to understand retrotransposons and retroviruses. Hereafter, I will refer to the reference element Ty1-H3 when describing Ty1 nucleotide coordinates [118].

Ty1

Ty1 elements are about 5.9 kb in length and bracketed by 334 bp LTRs (Figure 1.3) [114]. Ty1 contains two open reading frames: *GAG* encodes the capsid protein of virus-like particles (VLPs), and *POL* encodes a polyprotein that contains protease (PR), integrase (IN) and reverse transcriptase (RT). *POL* is in the +1 frame relative to *GAG* and overlaps the last 38 bps of *GAG*. Ty1 LTRs are composed of U3, R, and U5 domains, which are important for transcription, reverse transcription and integration. The U3 (240 nts) and U5 (38 nts) are unique to the 5′ and 3′ end of the Ty1 RNA, respectively, while R (56 nts) is repeated at both ends of the processed transcript. The last three nucleotides of the R region of the 5' LTR encode the first codon of Gag.

Ty1 replication

The Ty1 replication cycle (Figure 1.4) begins with transcription of the element. Ty1 mRNA is transcribed by RNA polymerase II from the start of the 5′ R region (nucleotide 241) to the end of the 3′ R region (nucleotide 5880) [119-121], yielding a 5.6 kb transcript (Figure 1.3). The Ty1 promoter is an approximately 1 kb region encompassing the 5´ LTR and part of *GAG*, and containing transcription factor binding sites and two TATA boxes located in the 5´ LTR [122]. Surprisingly, about 0.1- 0.8% of total cellular RNAs and 10% of polyadenylated mRNAs are Ty1 transcripts [123, 124]. Ty1 RNA has a longer half-life than the average yeast mRNAs, rather than increased transcriptional activity [125-127]. Only about 15% of total cellular Ty1 RNA is polyadenylated, suggesting that deadenylated Ty1 RNA is not rapidly degraded [128]. Ty1 mRNA is exported to the cytoplasm via the Mex67-dependent pathway and Gag facilitates Ty1 RNA export and localization to retrosomes [128, 129].

Ty1 mRNA is translated into two protein products: Gag-p49 and Gag-Pol-p199 (Figure 1.3). Pol is generated by a programmed +1 frameshifting event at a heptanucleotide sequence CUU-AGG-C located in the 38-nucleotide overlap between *GAG* and *POL* (Figure 1.3). A ribosome that is decoding the CUU codon and has a peptidyl-tRNA-Leu bound to the codon in the ribosomal P site pauses because of low availability of the AGG-decoding tRNA-Arg-CCU that should anneal with the A site codon. During the pause, peptidyl-tRNA-Leu-UAG also recognizes the overlapping Leu codon (UUA) in the +1 frame. After peptidyl-tRNA slippage, translational elongation resumes in the new +1 reading frame, leading to the expression of the Gag-Pol fusion protein. Gag-p49 and Gag-Pol-p199 are generated in a ratio of about 20:1 [130-133].

Ty1 mRNA and translated Ty1 proteins together form a complex called a retrosome, which are assembly sites for Ty1 VLPs [105, 128, 134]. Retrosomes were initially reported as T-bodies, and were defined as colocalized foci of Ty1 mRNA and Gag [128]. In the S288c strain background, about 30-50% of cells contain retrosomes at 20 °C but rarely at 30 °C. Ty1 retrosomes are distinct from P-bodies, which are cytoplasmic foci involved in mRNA turnover, although P-body components are required for retrosome formation [114, 128, 134]. After synthesis, Gag translocates to the endoplasmic reticulum (ER), then retrotranslocates back to the cytoplasm by an unknown mechanism. Retrosomes are nucleated after Gag exits the ER. Upon Gag's return to the cytoplasm, Gag is postulated to bind the translating Ty1 mRNA/ signal recognition particle (SRP) /ribosome nascent chain complex [135]. In addition, Gag also associates with Ty1 RNA as it exits the nucleus [129].

Ty1 VLPs are particles formed with Gag and Gag-Pol that specifically encapsidate

Ty1 RNA through Gag's nucleic acid chaperone activity [114]. Ty1 VLPs are icosahedral

particles with a diameter of about 40 μm. Some VLPs contain an electron-dense center

indicative of nucleic acid. Surprisingly, the shell is porous to proteins with molecular

mass of ~13 kDa (e.g. RNaseA), but proteins larger than ~30 kDa (e.g. Benzonase)

cannot [136, 137]. In addition, nucleic acids can gain access to RT and IN within the

particle. Gag, the structural protein for the VLPs, is both necessary and sufficient for

assembly as VLPs are formed in *E. coli* when Gag is exogenously expressed [138, 139].

Epitope mapping suggests that the N-terminus of Gag faces the surface of the particle,

while the C-terminus faces the core [140]. In this way, the Pol portion of Gag-Pol is

located inside of the particle [114, 141]. During VLP assembly, Ty1 RNA binds to the

nucleic acid chaperone domain of Gag, which is the functional equivalent of retroviral

nucleocapsid, and is packaged into the VLP. Although Ty1 RNA packaging is specific,

cellular transcripts from *TRP1* and *HIS3* can be also packaged into the VLP [142]. Many

of the cellular transcripts that are packaged reflect their abundance in the cytoplasm.

Ty1 VLPs contain RNA in a dimeric form [143]. Nucleotides 237-380 contains *cis*-acting

sequences required for Ty1 RNA packaging, however their precise location and

structure remains to be determined [114, 142]. After assembly, VLPs undergo

maturation catalyzed by Ty1 PR. Ty1 PR cleaves p199 Gag-Pol to release p49-Gag, PR,

IN and RT, and cleaves p49-Gag to p45-Gag. Cellular proteases cannot substitute for

the specific cleavage events mediated by Ty1 PR, as determined by genetic analyses

[144]. Reverse transcription takes place in the mature VLP, where tRNAi[Met] acts as a

primer to copy Ty1 RNA to cDNA [143, 145, 146]. A primer binding sequence (PBS) in

the 5´ of Ty1 RNA (nts 95-104) binds to 3´ end of tRNA$_i^{Met}$. Ty1 RT and IN interact in the

VLPs and this interaction is critical for proper reverse transcription [147, 148]. Similar to

retroviruses, reverse transcription within Ty1 VLPs yields linear double-stranded cDNA.

Ty1 cDNA, IN and perhaps other Ty1 or cellular proteins form the pre-integration

complex (PIC) or intasome, which may be released from the VLP by an unknown

mechanism. The PIC is imported into the nucleus by a bipartite nuclear localization

signal at the C-terminus of IN through the classical nuclear import pathway mediated by

importin-α [149-151]. The integration machinery recognizes dinucleotides 5´ - TG…..CA

- 3´ at the termini of Ty1 cDNA [152]. IN also cleaves target DNA on both strands

separated by five nucleotides. Thus, a characteristic of Ty1 cDNA integration are 5 bp of

target site duplications flanking newly integrated cDNA [114].

Genomic regions upstream of genes transcribed by RNA polymerase III (RNA pol III)

such as tRNA genes are favored sites for Ty1 integration [153, 154]. Ty1 integration

takes place within 700 bp upstream of tRNA genes with a canonical spacing of 80

nucleotides [153, 155]. Since sequence upstream of tRNA genes generally lacks

protein-coding genes, this preference is thought to be a strategy to minimize the

negative impact of retrotransposition to the host genome [156, 157]. It was discovered

early on that Ty1 does not recognize specific DNA sequences, suggesting that

chromatin or DNA binding proteins may facilitate integration. Ty1 likely targets a

nucleosomal surface at the H2A/H2B interface [154, 158], which explains the periodicity.

RNA polymerase III factor TFIIIB and the Bdp1 subunit, and the ISW2 chromatin-

remodeling complex influence Ty1 integration at preferred sites [159]. Recent studies

reported a direct interaction between Ty1 IN with RNA polymerase III. The C-terminus

of IN interacts with RNA pol III subunits such as Rpc40 and the Rpc53/57 subcomplex, and this interaction is critical for Ty1 integration upstream of RNA polymerase III– transcribed genes [160, 161]. Esp1, a protease that cleaves the cohesin complex, also interacts with Ty1 IN and mediates insertion of Ty1 elements [162].

Ty1 copy number control

While retrotransposons and retrovirus-like elements comprise almost half the human genome, only about 3% of the 12 Mb *S. cerevisiae* genome is composed of Ty retrotransposons [7, 115]. A survey of natural isolates suggests that *S. cerevisiae* does not carry more than 40 Ty1 elements in the genome, although the capability of carrying more Ty1 copies has been reported using experimentally repopulated strains [163-165]. In addition, strains lacking Ty1 elements have been reported, although all strains analyzed to date contain solo-LTRs, indicative of previous Ty1 insertions [166]. Together, these observations suggest that Ty1 retrotransposition may be under tight control in *Saccharomyces*, and copy number is a dynamic process based on Ty1 gain through retrotransposition verses loss by LTR-LTR recombination [167]. To maintain genome integrity, many eukaryotes have evolved ways to silence retrotransposons, such as DNA methylation, the APOBEC3 cytidine deaminase family, RNA interference, and innate immunity pathways [6, 168, 169]. Interestingly, none of these silencing pathways are present in the *S. cerevisiae sensu strictu* group (including *S. cerevisiae* and *S. paradoxus*), suggesting there is a novel way that the budding yeast controls Ty1 retrotransposition.

A genomic copy number dependent effect on Ty1 mobility, termed copy number control (CNC) was first proposed by the Garfinkel lab in 2003 [170]. When a plasmid

containing a Ty1 element marked with the retromobility indicator gene *his3-AI* (Figure

1.5 and [171]) was expressed in the *S. paradoxus* strain that lacks Ty1 elements, a

significantly higher Ty1 transposition frequency was observed compared to the *S.

cerevisiae* laboratory strain S288c which carries 32 Ty1 elements. Moreover, when

additional Ty1 elements were introduced to a Ty1-less *S. paradoxus* strain,

retrotransposition decreased in a copy number dependent manner. Interestingly, the

minimal Ty1 sequence required for CNC was contained within *GAG* plus additional

sequences in the 5′ LTR [170]. After the initial discovery, considerable efforts were

made to determine what factors mediate CNC.

Initially, transcripts expressed from the antisense strand (AS) of the CNC region of

Ty1 were identified. Ty1 AS RNAs were characterized by northern blot hybridization and

mapped by rapid-amplification of cDNA ends (RACE) mapping [172]. Multiple

transcripts with different 5′ ends and identical 3′ ends were identified, and evidence was

presented suggesting that AS RNAs help repress Ty1 transcription. Further work

suggested that Ty1 AS transcripts are involved in CNC [173]. Ty1 mRNA expression is

not necessary for AS RNA transcription. AS RNAs increase as genomic Ty1 copy

number increases and AS RNAs are enriched in VLP fractions. Increases in AS RNA

are correlated with a loss of IN and a decrease in Ty1 reverse transcription. Since AS

RNAs cofractionate with VLPs but are not packaged in the VLPs, they were

hypothesized to contribute to forming defective VLPs, perhaps by interacting with Gag

and interrupting VLP assembly [173, 174]. However, expression of individual or multiple

AS transcripts independently of Ty1 fail to inhibit Ty1 mobility, raising the possibility that

other factors were involved in CNC.

17

In 2015, the Garfinkel lab demonstrated that a novel protein containing the C-terminal half of Gag (denoted p22) underlies CNC. Co-expression of two galactose-inducible plasmids – one plasmid containing Ty1*his3-AI* and the other plasmid containing p22 showed a trans-dominant negative effect and decreased Ty1 mobility 35,000-fold [175]. In addition, loss of p22 abrogated CNC. Thus, p22 action is both necessary and sufficient for CNC and the proposed model for how CNC occurs will be described here (also reviewed in [176]).

An internal sense transcript termed Ty1i RNA with 5´ initiation sites located at position 1000 is the template for p22 (Figure 1.3). Ty1i RNA is expressed in standard lab strains, and its level increases in *spt3*Δ and *xrn1*Δ mutants [175]. Spt3 is a component of the SAGA chromatin-remodeling complex and is required for full-length Ty1 mRNA transcription. Xrn1 is a 5-3´ exonuclease required for RNA degradation in P-bodies (cytoplasmic foci involved in RNA turnover) [177-179]. A recent study described a role for the Mediator complex in Ty1i RNA expression. Mediator transduces a signal from a transcription factor to the transcriptional machinery, and affects Ty1i RNA expression by influencing the balance between initiation from two promoters in Ty1 [180]. Two closely located start codons, AUG1 (position 1038) and AUG2 (position 1068), are present downstream of the Ty1i initiation site. Either AUG1 or AUG2 can be used for p22 translation via cap-dependent process [181]. p22 is processed by Ty1 PR at the same PR cleavage site utilized by the full-length Gag protein to p18 (Figure 1.3), and p22 and p18 are both potent inhibitors of Ty1 transposition [181]. The relative amount of p18 to p22 is correlated with the level of Ty1 PR and association of p22 with assembling VLPs [182].

The mechanism of p22-mediated restriction requires an interaction with Gag, the capsid protein for Ty1 VLPs (Figure 1.6). p22 interacts with full-length Gag *in vitro* [175]. To learn more about p22/VLP interactions, mutations that restore retrotransposition in the presence of p22 and thus conferred resistance to copy number control (CNC^R) were characterized. Missense mutations were mapped in *GAG* but more than half of mutations were outside of p22 coding region and clustered in the "CNC^R domain". Interestingly, CNC^R mutations confer resistance by excluding p22 from assembling particles [182].

Current work suggests that p22 affects retrosome formation, VLP assembly, maturation and function. The earliest step in Ty1 replication affected by p22 is retrosome formation. Retrosomes are disrupted and the number of puncta (indicative of defective retrosome formation) increases in a strain overexpressing p22. Malformed retrosomes correlate to defects in VLP assembly and function. Also, p22 colocalizes with Gag and cofractionates with VLPs determined by sucrose sedimentation analysis [175]. However, p22 cofractionates with soluble proteins in the absence of Gag. Cells undergoing CNC (CNC$^+$) contain malformed or open particles [175]. In CNC$^+$ VLPs, defects in Ty1 protein processing occur during the VLP maturation, as shown by a build-up of Pol precursors, reduced mature RT, and most significantly, loss of mature IN [173]. In addition to the critical role IN plays in integration of Ty1 cDNA, the absence of mature IN impacts replication because both IN and RT are required for reverse transcription [147, 148]. Therefore, Ty1 retrotransposition is blocked due to low levels of reverse transcription [173].

The mode of action for p22 in Ty1 CNC resembles retroviral restriction factors encoded by retroelement *GAG* genes domesticated from retroviruses or LTR-retrotransposons, such as Fv1 in mice and enJS56A1 in sheep [176]. These restriction factors are derived from host retrotransposons or endogenous retroviruses, and thus represent domestication events repurposed to function as part of the innate immunity system. The replication step that restriction factors target and the strategy used by Fv1 and enJS56A1 are different from each other. Fv1 inhibits MLV infection after reverse transcription, but prior to integration. enJS56A1 blocks trafficking of JSRV viral particles to the plasma membrane. p22 affects the assembly and function of Ty1 VLPs [183]. Another difference is that the Fv1 and enJS56A1 are from domesticated retrotransposons and they target the exogenous virus, while p22 is self-encoded and targets Ty1 itself. This raises the interesting question of why p22 was not domesticated, which needs further investigation.

Ty1 host factors

Retrotransposons are capable of making only a few proteins critical for replication relative to their more complex life cycle, therefore, they are under the influence of host factors that modulate their life cycle [184]. Host genes could act to inhibit (restriction factors) or help (cofactors) replication of the retrotransposon. Elucidating the relationship between the retrotransposons and host factors is important to better characterize the element, and *S. cerevisiae* is a powerful model to study host factors as more is known about its genetic repertoire than any other eukaryote. Moreover, what is discovered for LTR-retrotransposons, such as yeast Ty elements, may also benefit the retrovirus field due to the similarity to retroviruses that LTR-retrotransposons possess.

Host factors that modulate Ty1 and Ty3 retrotransposition involve fundamental cellular processes, and some of them, such as DNA repair and RNA processing, have been shown to affect both retrotransposon and retroviral replication [184]. Studying host factors of retrotransposons in the widely used model organism the budding yeast *Saccharomyces cerevisiae* gives the additional benefit of its simple genetic assays and genomic tools, such as the collection of yeast gene deletion strains [114]. To date, several genome-wide screens [185-188] or transposon mutagenesis [189] have reported ~500 cofactor and 110 restriction factor genes. These results are not likely to be a complete list because essential genes were not interrogated for their effect on Ty1 mobility. Host factors modulating Ty1 replication also take part in important cellular pathways. About ~80% of Ty1 restriction factors are involved in nuclear processes including DNA repair, transcription, chromatin structure and function, recombination and the cell cycle [185]. Many Ty1 cofactors are involved in transcription, chromatin modification, histone deacetylation, ribosome biogenesis and function, mRNA catabolic processes and RNA transport [186-188]. Thus, understanding the process of Ty1 transposition can inform other areas of research.

There are core host cellular processes that the retrotransposon requires for replication, such as synthesizing RNA and protein products, mRNA export, cDNA import to the nucleus and integration into the genome. For example, components of the SAGA chromatin-remodeling complex are involved in Ty1 mRNA transcription and Spt3 is required for Ty1 mRNA transcription [178, 179]. The Mex67-dependent mRNA export pathway is primarily responsible for exporting Ty1 mRNA to the cytoplasm [128]. Many nuclear pore complex (NPC) proteins are required for Ty1 mobility [190]. The NPC

nuclear basket components, including Nup1 and Nup2, are critical for Ty1 insertion at preferred sites of integration upstream of tRNA genes. Also, an evolutionary relationship between Ty1 retrotransposition and nuclear pore complex proteins such as Nup82 and Nup84 has been reported [190, 191]. A subset of ribosomal proteins and related factors were identified as Ty1 cofactors, such as the 40S ribosome subunit biogenesis factor Bud22, which is required for +1 frameshifting and Gag-p45 accumulation [187]. P-body components (Xrn1, Lsm1, Pat1, Dhh1) are important for retrosome formation [134]. Recent evidence indicates that the ER translocon and SRP machinery is required for retrosome formation and retrotransposition [135]. In addition, Dfg10, a polyprenol reductase that catalyzes conversion of polyprenol to dolichol (precursor for N-glycosylation in the ER), is required for Gag stability.

Cellular processes important for the host survival affect Ty1 as well, including genes maintaining genome stability and signaling pathway genes responding to environmental stress. Many Ty1 restriction genes have roles in maintaining genome integrity, and DNA repair, replication, recombination, and telomerase maintenance genes are involved. Although some of the genome integrity mutants may affect Ty1 transposition directly, many mutants stimulate transposition indirectly by activating the S-phase checkpoint pathways [192]. Telomere erosion activates Ty1 retrotransposition, as demonstrated in the strain that lacks Est2, a reverse transcriptase subunit of the telomerase holoenzyme. Rad9, a central component of the DNA-damage checkpoint pathway, is required for the activation of Ty1 in *est2Δ* [193]. Ty1 host factors are also involved in pathways responding to environmental cues such as mating pheromone, invasive growth MAPK pathway, and osmolarity. Ty1 restriction factors are components of the

high osmolarity glycerol (HOG) pathway mediating the hyperosmotic response (*AGP2, SSK2, PBS2, HOG1*), transcription factors responding to oxidative stress (*YAP1, STB5*), general stress response (*MSN4*), and a kinase involved in pheromone signaling and invasive growth pathways (*FUS3*) [185, 194, 195].

As described above, retrotransposons make use of fundamental host cellular processes, and the level of host regulation on retrotransposons is complex and diverse [184]. Due to these complexities, studying host factors that respond differently to each replication stage of the retrotransposon is critical to better understand the retrotransposon. There are also many transposon-, or species-specific host factors affecting each replication step in unique ways. Therefore, studying host factors modulating Ty1 CNC will help to better understand the mechanism of how p22 restricts Ty1. What is found in our study may additionally benefit studies of restriction factors affecting other retrotransposons or retroviruses by discovering a new set of conserved genes protecting hosts from retroelements.

REFERENCES

1.   McClintock, B., *The origin and behavior of mutable loci in maize.* Proc Natl Acad Sci U S A, 1950. **36**(6): p. 344-55.

2.   Fedoroff, N.V., *McClintock's challenge in the 21st century.* Proc Natl Acad Sci U S A, 2012. **109**(50): p. 20200-3.

3.   Dawkins, R., *The selfish gene*. New ed. 1989, Oxford ; New York: Oxford University Press. xi, 352 p.

4.   Finnegan, D.J., *Eukaryotic transposable elements and genome evolution.* Trends Genet, 1989. **5**(4): p. 103-7.

5.   Goodier, J.L., *Restricting retrotransposons: a review.* Mob DNA, 2016. **7**: p. 16.

6.   Levin, H.L. and J.V. Moran, *Dynamic interactions between transposable elements and their hosts.* Nat Rev Genet, 2011. **12**(9): p. 615-27.

7.   Lander, E.S., et al., *Initial sequencing and analysis of the human genome.* Nature, 2001. **409**(6822): p. 860-921.

8.   Grimaldi, G., J. Skowronski, and M.F. Singer, *Defining the beginning and end of KpnI family segments.* EMBO J, 1984. **3**(8): p. 1753-9.

9.   Brouha, B., et al., *Hot L1s account for the bulk of retrotransposition in the human population.* Proc Natl Acad Sci U S A, 2003. **100**(9): p. 5280-5.

10.   Dombroski, B.A., et al., *Isolation of an active human transposable element.* Science, 1991. **254**(5039): p. 1805-8.

11.   Scott, A.F., et al., *Origin of the human L1 elements: proposed progenitor genes deduced from a consensus DNA sequence.* Genomics, 1987. **1**(2): p. 113-25.

12.   Martin, S.L. and F.D. Bushman, *Nucleic acid chaperone activity of the ORF1 protein from the mouse LINE-1 retrotransposon.* Mol Cell Biol, 2001. **21**(2): p. 467-75.

13.   Holmes, S.E., M.F. Singer, and G.D. Swergold, *Studies on p40, the leucine zipper motif-containing protein encoded by the first open reading frame of an active human LINE-1 transposable element.* J Biol Chem, 1992. **267**(28): p. 19765-8.

14.   Hohjoh, H. and M.F. Singer, *Cytoplasmic ribonucleoprotein complexes containing human LINE-1 protein and RNA.* EMBO J, 1996. **15**(3): p. 630-9.

15.   Khazina, E., et al., *Trimeric structure and flexibility of the L1ORF1 protein in human L1 retrotransposition.* Nat Struct Mol Biol, 2011. **18**(9): p. 1006-14.

16.   Ergun, S., et al., *Cell type-specific expression of LINE-1 open reading frames 1 and 2 in fetal and adult human tissues.* J Biol Chem, 2004. **279**(26): p. 27753-63.

17.   Doucet, A.J., et al., *Characterization of LINE-1 ribonucleoprotein particles.* PLoS Genet, 2010. **6**(10).

18.   Feng, Q., et al., *Human L1 retrotransposon encodes a conserved endonuclease required for retrotransposition.* Cell, 1996. **87**(5): p. 905-16.

19.   Mathias, S.L., et al., *Reverse transcriptase encoded by a human transposable element.* Science, 1991. **254**(5039): p. 1808-10.

20.   Moran, J.V., et al., *High frequency retrotransposition in cultured mammalian cells.* Cell, 1996. **87**(5): p. 917-27.

21.   Denli, A.M., et al., *Primate-specific ORF0 contributes to retrotransposon-mediated diversity.* Cell, 2015. **163**(3): p. 583-93.

22.   Alisch, R.S., et al., *Unconventional translation of mammalian LINE-1 retrotransposons.* Genes Dev, 2006. **20**(2): p. 210-24.

23.   Luan, D.D., et al., *Reverse transcription of R2Bm RNA is primed by a nick at the chromosomal target site: a mechanism for non-LTR retrotransposition.* Cell, 1993. **72**(4): p. 595-605.

24.   Cost, G.J., et al., *Human L1 element target-primed reverse transcription in vitro.* EMBO J, 2002. **21**(21): p. 5899-910.

25.   Hayward, A., C.K. Cornwallis, and P. Jern, *Pan-vertebrate comparative genomics unmasks retrovirus macroevolution.* Proc Natl Acad Sci U S A, 2015. **112**(2): p. 464-9.

26.   Dewannieux, M. and T. Heidmann, *Endogenous retroviruses: acquisition, amplification and taming of genome invaders.* Curr Opin Virol, 2013. **3**(6): p. 646-56.

27.   Belshaw, R., et al., *Long-term reinfection of the human genome by endogenous retroviruses.* Proc Natl Acad Sci U S A, 2004. **101**(14): p. 4894-9.

28.   Magiorkinis, G., D. Blanco-Melo, and R. Belshaw, *The decline of human endogenous retroviruses: extinction and survival.* Retrovirology, 2015. **12**: p. 8.

29.   Seifarth, W., et al., *Comprehensive analysis of human endogenous retrovirus transcriptional activity in human tissues with a retrovirus-specific microarray.* J Virol, 2005. **79**(1): p. 341-52.

30.   Bannert, N. and R. Kurth, *The evolutionary dynamics of human endogenous retroviral families.* Annu Rev Genomics Hum Genet, 2006. **7**: p. 149-73.

31.   Lower, R., J. Lower, and R. Kurth, *The viruses in all of us: characteristics and biological significance of human endogenous retrovirus sequences.* Proc Natl Acad Sci U S A, 1996. **93**(11): p. 5177-84.

32.   Richardson, S.R., et al., *The Influence of LINE-1 and SINE Retrotransposons on Mammalian Genomes.* Microbiol Spectr, 2015. **3**(2): p. MDNA3-0061-2014.

33.   Dewannieux, M., et al., *Identification of an infectious progenitor for the multiple-copy HERV-K human endogenous retroelements.* Genome Res, 2006. **16**(12): p. 1548-56.

34.   Lee, Y.N. and P.D. Bieniasz, *Reconstitution of an infectious human endogenous retrovirus.* PLoS Pathog, 2007. **3**(1): p. e10.

35.   Xu, W. and M.V. Eiden, *Koala Retroviruses: Evolution and Disease Dynamics.* Annu Rev Virol, 2015. **2**(1): p. 119-34.

36.   Denner, J., *How Active Are Porcine Endogenous Retroviruses (PERVs)?* Viruses, 2016. **8**(8).

37.   Qin, C., et al., *Intracisternal A particle genes: Distribution in the mouse genome, active subtypes, and potential roles as species-specific mediators of susceptibility to cancer.* Mol Carcinog, 2010. **49**(1): p. 54-67.

38.   Malfavon-Borja, R. and C. Feschotte, *Fighting fire with fire: endogenous retrovirus envelopes as restriction factors.* J Virol, 2015. **89**(8): p. 4047-50.

39.   Frank, J.A. and C. Feschotte, *Co-option of endogenous viral sequences for host cell function.* Curr Opin Virol, 2017. **25**: p. 81-89.

40.   Pastuzyn, E.D., et al., *The Neuronal Gene Arc Encodes a Repurposed Retrotransposon Gag Protein that Mediates Intercellular RNA Transfer.* Cell, 2018. **172**(1-2): p. 275-288 e18.

41.   Chuong, E.B., N.C. Elde, and C. Feschotte, *Regulatory evolution of innate immunity through co-option of endogenous retroviruses.* Science, 2016. **351**(6277): p. 1083-7.

42.   Benit, L., et al., *Cloning of a new murine endogenous retrovirus, MuERV-L, with strong similarity to the human HERV-L element and with a gag coding sequence closely related to the Fv1 restriction gene.* J Virol, 1997. **71**(7): p. 5652-7.

43.   Best, S., et al., *Positional cloning of the mouse retrovirus restriction gene Fv1.* Nature, 1996. **382**(6594): p. 826-9.

44.   Arnaud, F., P.R. Murcia, and M. Palmarini, *Mechanisms of late restriction induced by an endogenous retrovirus.* J Virol, 2007. **81**(20): p. 11441-51.

45.   Meyer, T.J., et al., *Endogenous Retroviruses: With Us and against Us.* Front Chem, 2017. **5**: p. 23.

46.   Romer, C., et al., *How to tame an endogenous retrovirus: HERVH and the evolution of human pluripotency.* Curr Opin Virol, 2017. **25**: p. 49-58.

47.   Wang, J., et al., *Primate-specific endogenous retrovirus-driven transcription defines naive-like stem cells.* Nature, 2014. **516**(7531): p. 405-9.

48.   Grow, E.J., et al., *Intrinsic retroviral reactivation in human preimplantation embryos and pluripotent cells.* Nature, 2015. **522**(7555): p. 221-5.

49.   Percharde, M., et al., *A LINE1-Nucleolin Partnership Regulates Early Development and ESC Identity.* Cell, 2018. **174**(2): p. 391-405 e19.

50.   Narva, E., et al., *RNA-binding protein L1TD1 interacts with LIN28 via RNA and is required for human embryonic stem cell self-renewal and cancer cell proliferation.* Stem Cells, 2012. **30**(3): p. 452-60.

51.   Iwabuchi, K.A., et al., *ECAT11/L1td1 is enriched in ESCs and rapidly activated during iPSC generation, but it is dispensable for the maintenance and induction of pluripotency.* PLoS One, 2011. **6**(5): p. e20461.

52.   McLaughlin, R.N., Jr., et al., *Positive selection and multiple losses of the LINE-1-derived L1TD1 gene in mammals suggest a dual role in genome defense and pluripotency.* PLoS Genet, 2014. **10**(9): p. e1004531.

53.   Soygur, B. and L. Sati, *The role of syncytins in human reproduction and reproductive organ cancers.* Reproduction, 2016. **152**(5): p. R167-78.

54.   Wimmer, K., et al., *The NF1 gene contains hotspots for L1 endonuclease-dependent de novo insertion.* PLoS Genet, 2011. **7**(11): p. e1002371.

55.   Hancks, D.C. and H.H. Kazazian, Jr., *Roles for retrotransposon insertions in human disease.* Mob DNA, 2016. **7**: p. 9.

56.   van der Klift, H.M., et al., *Insertion of an SVA element, a nonautonomous retrotransposon, in PMS2 intron 7 as a novel cause of Lynch syndrome.* Hum Mutat, 2012. **33**(7): p. 1051-5.

57.   Vogt, J., et al., *SVA retrotransposon insertion-associated deletion represents a novel mutational mechanism underlying large genomic copy number changes with non-recurrent breakpoints.* Genome Biol, 2014. **15**(6): p. R80.

58.   Wallace, M.R., et al., *A de novo Alu insertion results in neurofibromatosis type 1.* Nature, 1991. **353**(6347): p. 864-6.

59.   Hassoun, H., et al., *A novel mobile element inserted in the alpha spectrin gene: spectrin dayton. A truncated alpha spectrin associated with hereditary elliptocytosis.* J Clin Invest, 1994. **94**(2): p. 643-8.

60.   Beck, C.R., et al., *LINE-1 elements in structural variation and disease.* Annu Rev Genomics Hum Genet, 2011. **12**: p. 187-215.

61.   Morrish, T.A., et al., *Endonuclease-independent LINE-1 retrotransposition at mammalian telomeres.* Nature, 2007. **446**(7132): p. 208-12.

62.   Morrish, T.A., et al., *DNA repair mediated by endonuclease-independent LINE-1 retrotransposition.* Nat Genet, 2002. **31**(2): p. 159-65.

63.   Konkel, M.K. and M.A. Batzer, *A mobile threat to genome stability: The impact of non-LTR retrotransposons upon the human genome.* Semin Cancer Biol, 2010. **20**(4): p. 211-21.

64.   Kazazian, H.H., Jr. and J.V. Moran, *Mobile DNA in Health and Disease.* N Engl J Med, 2017. **377**(4): p. 361-370.

65.   Kazazian, H.H., Jr., et al., *Haemophilia A resulting from de novo insertion of L1 sequences represents a novel mechanism for mutation in man.* Nature, 1988. **332**(6160): p. 164-6.

66.   Garcia-Perez, J.L., T.J. Widmann, and I.R. Adams, *The impact of transposable elements on mammalian development.* Development, 2016. **143**(22): p. 4101-4114.

67.   Galli, U.M., et al., *Human endogenous retrovirus rec interferes with germ cell development in mice and may cause carcinoma in situ, the predecessor lesion of germ cell tumors.* Oncogene, 2005. **24**(19): p. 3223-8.

68.   Newkirk, S.J., et al., *Intact piRNA pathway prevents L1 mobilization in male meiosis.* Proc Natl Acad Sci U S A, 2017. **114**(28): p. E5635-E5644.

69.   Geister, K.A., et al., *LINE-1 Mediated Insertion into Poc1a (Protein of Centriole 1 A) Causes Growth Insufficiency and Male Infertility in Mice.* PLoS Genet, 2015. **11**(10): p. e1005569.

70.   Malki, S., et al., *A role for retrotransposon LINE-1 in fetal oocyte attrition in mice.* Dev Cell, 2014. **29**(5): p. 521-533.

71.   Su, Y.Q., et al., *MARF1 regulates essential oogenic processes in mice.* Science, 2012. **335**(6075): p. 1496-9.

72.	Crichton, J.H., et al., *Defending the genome from the enemy within: mechanisms of retrotransposon suppression in the mouse germline.* Cell Mol Life Sci, 2014. **71**(9): p. 1581-605.

73.	Burns, K.H., *Transposable elements in cancer.* Nat Rev Cancer, 2017. **17**(7): p. 415-424.

74.	Lee, E., et al., *Landscape of somatic retrotransposition in human cancers.* Science, 2012. **337**(6097): p. 967-71.

75.	Scott, E.C., et al., *A hot L1 retrotransposon evades somatic repression and initiates human colorectal cancer.* Genome Res, 2016. **26**(6): p. 745-55.

76.	Macia, A., et al., *Engineered LINE-1 retrotransposition in nondividing human neurons.* Genome Res, 2017. **27**(3): p. 335-348.

77.	Antony, J.M., et al., *Human endogenous retroviruses and multiple sclerosis: innocent bystanders or disease determinants?* Biochim Biophys Acta, 2011. **1812**(2): p. 162-76.

78.	Thomas, C.A., A.C. Paquola, and A.R. Muotri, *LINE-1 retrotransposition in the nervous system.* Annu Rev Cell Dev Biol, 2012. **28**: p. 555-73.

79.	Bundo, M., et al., *Increased l1 retrotransposition in the neuronal genome in schizophrenia.* Neuron, 2014. **81**(2): p. 306-13.

80.	Garcia-Montojo, M., et al., *The DNA copy number of human endogenous retrovirus-W (MSRV-type) is increased in multiple sclerosis patients and is influenced by gender and disease severity.* PLoS One, 2013. **8**(1): p. e53623.

81.   Moroni, C., et al., *Normal B-cell activation involves endogenous retroviral antigen expression: implications for leukemogenesis.* Cold Spring Harb Symp Quant Biol, 1980. **44 Pt 2**: p. 1205-10.

82.   Stoye, J.P. and C. Moroni, *Endogenous retrovirus expression in stimulated murine lymphocytes. Identification of a new locus controlling mitogen induction of a defective virus.* J Exp Med, 1983. **157**(5): p. 1660-74.

83.   Crow, Y.J. and J. Rehwinkel, *Aicardi-Goutieres syndrome and related phenotypes: linking nucleic acid metabolism with autoimmunity.* Hum Mol Genet, 2009. **18**(R2): p. R130-6.

84.   Kassiotis, G. and J.P. Stoye, *Immune responses to endogenous retroelements: taking the bad with the good.* Nat Rev Immunol, 2016. **16**(4): p. 207-19.

85.   Rowley, P.A., *The frenemies within: viruses, retrotransposons and plasmids that naturally infect Saccharomyces yeasts.* Yeast, 2017. **34**(7): p. 279-292.

86.   Yoder, J.A., C.P. Walsh, and T.H. Bestor, *Cytosine methylation and the ecology of intragenomic parasites.* Trends Genet, 1997. **13**(8): p. 335-40.

87.   Selker, E.U., et al., *The methylated component of the Neurospora crassa genome.* Nature, 2003. **422**(6934): p. 893-7.

88.   Crowther, P.J., et al., *Revised genomic consensus for the hypermethylated CpG island region of the human L1 transposon and integration sites of full length L1 elements from recombinant clones made using methylation-tolerant host strains.* Nucleic Acids Res, 1991. **19**(9): p. 2395-401.

89.   Thayer, R.E., M.F. Singer, and T.G. Fanning, *Undermethylation of specific LINE-1 sequences in human cells producing a LINE-1-encoded protein.* Gene, 1993. **133**(2): p. 273-7.

90.   Hata, K. and Y. Sakaki, *Identification of critical CpG sites for repression of L1 transcription by DNA methylation.* Gene, 1997. **189**(2): p. 227-34.

91.   Woodcock, D.M., et al., *Asymmetric methylation in the hypermethylated CpG promoter region of the human L1 retrotransposon.* J Biol Chem, 1997. **272**(12): p. 7810-6.

92.   Goll, M.G. and T.H. Bestor, *Eukaryotic cytosine methyltransferases.* Annu Rev Biochem, 2005. **74**: p. 481-514.

93.   Schaefer, C.B., et al., *Epigenetic decisions in mammalian germ cells.* Science, 2007. **316**(5823): p. 398-9.

94.   Maksakova, I.A., D.L. Mager, and D. Reiss, *Keeping active endogenous retroviral-like elements in check: the epigenetic perspective.* Cell Mol Life Sci, 2008. **65**(21): p. 3329-47.

95.   Bulut-Karslioglu, A., et al., *Suv39h-dependent H3K9me3 marks intact retrotransposons and silences LINE elements in mouse embryonic stem cells.* Mol Cell, 2014. **55**(2): p. 277-90.

96.   Martens, J.H., et al., *The profile of repeat-associated histone lysine methylation states in the mouse epigenome.* EMBO J, 2005. **24**(4): p. 800-12.

97.   Hunter, R.G., et al., *Acute stress and hippocampal histone H3 lysine 9 trimethylation, a retrotransposon silencing response.* Proc Natl Acad Sci U S A, 2012. **109**(43): p. 17657-62.

98.   Di Giacomo, M., et al., *G9a co-suppresses LINE1 elements in spermatogonia.* Epigenetics Chromatin, 2014. **7**: p. 24.

99.   Vieira, V.C. and M.A. Soares, *The role of cytidine deaminases on innate immune responses against human viral infections.* Biomed Res Int, 2013. **2013**: p. 683095.

100.   Sheehy, A.M., et al., *Isolation of a human gene that inhibits HIV-1 infection and is suppressed by the viral Vif protein.* Nature, 2002. **418**(6898): p. 646-50.

101.   Harris, R.S. and M.T. Liddament, *Retroviral restriction by APOBEC proteins.* Nat Rev Immunol, 2004. **4**(11): p. 868-77.

102.   Chiu, Y.L. and W.C. Greene, *The APOBEC3 cytidine deaminases: an innate defensive network opposing exogenous retroviruses and endogenous retroelements.* Annu Rev Immunol, 2008. **26**: p. 317-53.

103.   Koito, A. and T. Ikeda, *Intrinsic immunity against retrotransposons by APOBEC cytidine deaminases.* Front Microbiol, 2013. **4**: p. 28.

104.   Arias, J.F., et al., *Retroelements versus APOBEC3 family members: No great escape from the magnificent seven.* Front Microbiol, 2012. **3**: p. 275.

105.   Dutko, J.A., et al., *5' to 3' mRNA decay factors colocalize with Ty1 Gag and human APOBEC3G and promote Ty1 retrotransposition.* J Virol, 2010. **84**(10): p. 5052-66.

106.   Heras, S.R., et al., *The Microprocessor controls the activity of mammalian retrotransposons.* Nat Struct Mol Biol, 2013. **20**(10): p. 1173-81.

107.   Chen, L., et al., *Naturally occurring endo-siRNA silences LINE-1 retrotransposons in human cells through DNA methylation.* Epigenetics, 2012. **7**(7): p. 758-71.

108.   Yang, N. and H.H. Kazazian, Jr., *L1 retrotransposition is suppressed by endogenously encoded small interfering RNAs in human cultured cells.* Nat Struct Mol Biol, 2006. **13**(9): p. 763-71.

109.   Aravin, A., et al., *A novel class of small RNAs bind to MILI protein in mouse testes.* Nature, 2006. **442**(7099): p. 203-7.

110.   Girard, A., et al., *A germline-specific class of small RNAs binds mammalian Piwi proteins.* Nature, 2006. **442**(7099): p. 199-202.

111.   Brennecke, J., et al., *Discrete small RNA-generating loci as master regulators of transposon activity in Drosophila.* Cell, 2007. **128**(6): p. 1089-103.

112.   Gunawardane, L.S., et al., *A slicer-mediated mechanism for repeat-associated siRNA 5' end formation in Drosophila.* Science, 2007. **315**(5818): p. 1587-90.

113.   Lesage, P. and A.L. Todeschini, *Happy together: the life and times of Ty retrotransposons and their hosts.* Cytogenet Genome Res, 2005. **110**(1-4): p. 70-90.

114.   Curcio, M.J., S. Lutz, and P. Lesage, *The Ty1 LTR-retrotransposon of budding yeast, Saccharomyces cerevisiae.* Microbiol Spectr, 2015. **3**(2): p. 1-35.

115.   Kim, J.M., et al., *Transposable elements and genome organization: a comprehensive survey of retrotransposons revealed by the complete Saccharomyces cerevisiae genome sequence.* Genome Res, 1998. **8**(5): p. 464-78.

116. Hug, A.M. and H. Feldmann, *Yeast retrotransposon Ty4: the majority of the rare transcripts lack a U3-R sequence.* Nucleic Acids Res, 1996. **24**(12): p. 2338-46.

117. Voytas, D.F. and J.D. Boeke, *Yeast retrotransposon revealed.* Nature, 1992. **358**(6389): p. 717.

118. Boeke, J.D., et al., *Ty elements transpose through an RNA intermediate.* Cell, 1985. **40**(3): p. 491-500.

119. Mules, E.H., O. Uzun, and A. Gabriel, *In vivo Ty1 reverse transcription can generate replication intermediates with untidy ends.* J Virol, 1998. **72**(8): p. 6490-503.

120. Elder, R.T., E.Y. Loh, and R.W. Davis, *RNA from the yeast transposable element Ty1 has both ends in the direct repeats, a structure similar to retrovirus RNA.* Proc Natl Acad Sci U S A, 1983. **80**(9): p. 2432-6.

121. Hou, W., R. Russnak, and T. Platt, *Poly(A) site selection in the yeast Ty retroelement requires an upstream region and sequence-specific titratable factor(s) in vitro.* EMBO J, 1994. **13**(2): p. 446-52.

122. Servant, G., C. Pennetier, and P. Lesage, *Remodeling yeast gene transcription by activating the Ty1 long terminal repeat retrotransposon under severe adenine deficiency.* Mol Cell Biol, 2008. **28**(17): p. 5543-54.

123. Elder, R.T., et al., *Studies on the transposable element Ty1 of yeast. I. RNA homologous to Ty1. II. Recombination and expression of Ty1 and adjacent sequences.* Cold Spring Harb Symp Quant Biol, 1981. **45 Pt 2**: p. 581-91.

124. Curcio, M.J., et al., *Ty RNA levels determine the spectrum of retrotransposition events that activate gene expression in Saccharomyces cerevisiae.* Mol Gen Genet, 1990. **220**(2): p. 213-21.

125. Huang, Q., et al., *Retrotransposon Ty1 RNA contains a 5'-terminal long-range pseudoknot required for efficient reverse transcription.* RNA, 2013. **19**(3): p. 320-32.

126. Morillon, A., et al., *Differential effects of chromatin and Gcn4 on the 50-fold range of expression among individual yeast Ty1 retrotransposons.* Mol Cell Biol, 2002. **22**(7): p. 2078-88.

127. Munchel, S.E., et al., *Dynamic profiling of mRNA turnover reveals gene-specific and system-wide regulation of mRNA decay.* Mol Biol Cell, 2011. **22**(15): p. 2787-95.

128. Malagon, F. and T.H. Jensen, *The T body, a new cytoplasmic RNA granule in Saccharomyces cerevisiae.* Mol Cell Biol, 2008. **28**(19): p. 6022-32.

129. Checkley, M.A., et al., *Ty1 Gag enhances the stability and nuclear export of Ty1 mRNA.* Traffic, 2013. **14**(1): p. 57-69.

130. Belcourt, M.F. and P.J. Farabaugh, *Ribosomal frameshifting in the yeast retrotransposon Ty: tRNAs induce slippage on a 7 nucleotide minimal site.* Cell, 1990. **62**(2): p. 339-52.

131. Kawakami, K., et al., *A rare tRNA-Arg(CCU) that regulates Ty1 element ribosomal frameshifting is essential for Ty1 retrotransposition in Saccharomyces cerevisiae.* Genetics, 1993. **135**(2): p. 309-20.

132. Sundararajan, A., et al., *Near-cognate peptidyl-tRNAs promote +1 programmed translational frameshifting in yeast.* Mol Cell, 1999. **4**(6): p. 1005-15.

133. Weissenbach, J., et al., *Yeast tRNALeu (anticodon U--A--G) translates all six leucine codons in extracts from interferon treated cells.* FEBS Lett, 1977. **82**(1): p. 71-6.

134. Checkley, M.A., et al., *P-body components are required for Ty1 retrotransposition during assembly of retrotransposition-competent virus-like particles.* Mol Cell Biol, 2010. **30**(2): p. 382-98.

135. Doh, J.H., S. Lutz, and M.J. Curcio, *Co-translational localization of an LTR-retrotransposon RNA to the endoplasmic reticulum nucleates virus-like particle assembly sites.* PLoS Genet, 2014. **10**(3): p. e1004219.

136. HA, A.L.-K., et al., *Yeast Ty retrotransposons assemble into virus-like particles whose T-numbers depend on the C-terminal length of the capsid protein.* J Mol Biol, 1999. **292**(1): p. 65-73.

137. Burns, N.R., et al., *Symmetry, flexibility and permeability in the structure of yeast retrotransposon virus-like particles.* EMBO J, 1992. **11**(3): p. 1155-64.

138. Luschnig, C. and A. Bachmair, *RNA packaging of yeast retrotransposon Ty1 in the heterologous host, Escherichia coli.* Biol Chem, 1997. **378**(1): p. 39-46.

139. Luschnig, C., et al., *The gag homologue of retrotransposon Ty1 assembles into spherical particles in Escherichia coli.* Eur J Biochem, 1995. **228**(3): p. 739-44.

140. Brookman, J.L., et al., *Analysis of TYA protein regions necessary for formation of the Ty1 virus-like particle structure.* Virology, 1995. **212**(1): p. 69-76.

141. Roth, J.F., *The yeast Ty virus-like particles.* Yeast, 2000. **16**(9): p. 785-95.

142. Xu, H. and J.D. Boeke, *Localization of sequences required in cis for yeast Ty1 element transposition near the long terminal repeats: analysis of mini-Ty1 elements.* Mol Cell Biol, 1990. **10**(6): p. 2695-702.

143. Feng, Y.X., et al., *The genomic RNA in Ty1 virus-like particles is dimeric.* J Virol, 2000. **74**(22): p. 10819-21.

144. Youngren, S.D., et al., *Functional organization of the retrotransposon Ty from Saccharomyces cerevisiae: Ty protease is required for transposition.* Mol Cell Biol, 1988. **8**(4): p. 1421-31.

145. Chapman, K.B., A.S. Bystrom, and J.D. Boeke, *Initiator methionine tRNA is essential for Ty1 transposition in yeast.* Proc Natl Acad Sci U S A, 1992. **89**(8): p. 3236-40.

146. Wilhelm, M., et al., *Yeast Ty1 retrotransposon: the minus-strand primer binding site and a cis-acting domain of the Ty1 RNA are both important for packaging of primer tRNA inside virus-like particles.* Nucleic Acids Res, 1994. **22**(22): p. 4560-5.

147. Wilhelm, M., M. Boutabout, and F.X. Wilhelm, *Expression of an active form of recombinant Ty1 reverse transcriptase in Escherichia coli: a fusion protein containing the C-terminal region of the Ty1 integrase linked to the reverse transcriptase-RNase H domain exhibits polymerase and RNase H activities.* Biochem J, 2000. **348 Pt 2**: p. 337-42.

148. Wilhelm, M. and F.X. Wilhelm, *Cooperation between reverse transcriptase and integrase during reverse transcription and formation of the preintegrative complex of Ty1.* Eukaryot Cell, 2006. **5**(10): p. 1760-9.

149. Kenna, M.A., et al., *Invading the yeast nucleus: a nuclear localization signal at the C terminus of Ty1 integrase is required for transposition in vivo.* Mol Cell Biol, 1998. **18**(2): p. 1115-24.

150. McLane, L.M., et al., *The Ty1 integrase protein can exploit the classical nuclear protein import machinery for entry into the nucleus.* Nucleic Acids Res, 2008. **36**(13): p. 4317-26.

151. Moore, S.P., L.A. Rinckel, and D.J. Garfinkel, *A Ty1 integrase nuclear localization signal required for retrotransposition.* Mol Cell Biol, 1998. **18**(2): p. 1105-14.

152. Sharon, G., T.J. Burkett, and D.J. Garfinkel, *Efficient homologous recombination of Ty1 element cDNA when integration is blocked.* Mol Cell Biol, 1994. **14**(10): p. 6540-51.

153. Devine, S.E. and J.D. Boeke, *Integration of the yeast retrotransposon Ty1 is targeted to regions upstream of genes transcribed by RNA polymerase III.* Genes Dev, 1996. **10**(5): p. 620-33.

154. Baller, J.A., et al., *A nucleosomal surface defines an integration hotspot for the Saccharomyces cerevisiae Ty1 retrotransposon.* Genome Res, 2012. **22**(4): p. 704-13.

155. Ji, H., et al., *Hotspots for unselected Ty1 transposition events on yeast chromosome III are near tRNA genes and LTR sequences.* Cell, 1993. **73**(5): p. 1007-18.

156.   Blanc, V.M. and J. Adams, *Ty1 insertions in intergenic regions of the genome of Saccharomyces cerevisiae transcribed by RNA polymerase III have no detectable selective effect.* FEMS Yeast Res, 2004. **4**(4-5): p. 487-91.

157.   Boeke, J.D. and S.E. Devine, *Yeast retrotransposons: finding a nice quiet neighborhood.* Cell, 1998. **93**(7): p. 1087-9.

158.   Mularoni, L., et al., *Retrotransposon Ty1 integration targets specifically positioned asymmetric nucleosomal DNA segments in tRNA hotspots.* Genome Res, 2012. **22**(4): p. 693-703.

159.   Bachman, N., et al., *TFIIIB subunit Bdp1p is required for periodic integration of the Ty1 retrotransposon and targeting of Isw2p to S. cerevisiae tDNAs.* Genes Dev, 2005. **19**(8): p. 955-64.

160.   Bridier-Nahmias, A., et al., *Retrotransposons. An RNA polymerase III subunit determines sites of retrotransposon integration.* Science, 2015. **348**(6234): p. 585-8.

161.   Cheung, S., et al., *Ty1 integrase interacts with RNA polymerase III-specific subcomplexes to promote insertion of Ty1 elements upstream of polymerase (Pol) III-transcribed genes.* J Biol Chem, 2016. **291**(12): p. 6396-411.

162.   Ho, K.L., et al., *A role for the budding yeast separase, Esp1, in Ty1 element retrotransposition.* PLoS Genet, 2015. **11**(3): p. e1005109.

163.   Scheifele, L.Z., et al., *Retrotransposon overdose and genome integrity.* Proc Natl Acad Sci U S A, 2009. **106**(33): p. 13927-32.

164. Carr, M., D. Bensasson, and C.M. Bergman, *Evolutionary genomics of transposable elements in Saccharomyces cerevisiae.* PLoS One, 2012. **7**(11): p. e50978.

165. Wilke, C.M. and J. Adams, *Fitness effects of Ty transposition in Saccharomyces cerevisiae.* Genetics, 1992. **131**(1): p. 31-42.

166. Moore, S.P., et al., *Analysis of a Ty1-less variant of Saccharomyces paradoxus: the gain and loss of Ty1 elements.* Yeast, 2004. **21**(8): p. 649-660.

167. Garfinkel, D.J., et al., *Ty1 copy number dynamics in Saccharomyces.* Genetics, 2005. **169**(4): p. 1845-57.

168. Beauregard, A., M.J. Curcio, and M. Belfort, *The take and give between retrotransposable elements and their hosts.* Annu Rev Genet, 2008. **42**: p. 587-617.

169. Volkman, H.E. and D.B. Stetson, *The enemy within: endogenous retroelements and autoimmune disease.* Nat Immunol, 2014. **15**(5): p. 415-22.

170. Garfinkel, D.J., et al., *Post-transcriptional cosuppression of Ty1 retrotransposition.* Genetics, 2003. **165**(1): p. 83-99.

171. Curcio, M.J. and D.J. Garfinkel, *Single-step selection for Ty1 element retrotransposition.* Proc Natl Acad Sci U S A, 1991. **88**(3): p. 936-40.

172. Berretta, J., M. Pinskaya, and A. Morillon, *A cryptic unstable transcript mediates transcriptional trans-silencing of the Ty1 retrotransposon in S. cerevisiae.* Genes Dev, 2008. **22**(5): p. 615-26.

173. Matsuda, E. and D.J. Garfinkel, *Posttranslational interference of Ty1 retrotransposition by antisense RNAs.* Proc Natl Acad Sci U S A, 2009. **106**(37): p. 15657-62.

174. Purzycka, K.J., et al., *Exploring Ty1 retrotransposon RNA structure within virus-like particles.* Nucleic Acids Res, 2013. **41**(1): p. 463-73.

175. Saha, A., et al., *A trans-dominant form of Gag restricts Ty1 retrotransposition and mediates copy number control.* J Virol, 2015. **89**(7): p. 3922-38.

176. Garfinkel, D.J., et al., *A self-encoded capsid derivative restricts Ty1 retrotransposition in Saccharomyces.* Curr Genet, 2016. **62**(2): p. 321-9.

177. Parker, R., *RNA degradation in Saccharomyces cerevisae.* Genetics, 2012. **191**(3): p. 671-702.

178. Winston, F., K.J. Durbin, and G.R. Fink, *The SPT3 gene is required for normal transcription of Ty elements in S. cerevisiae.* Cell, 1984. **39**(3 Pt 2): p. 675-82.

179. Pollard, K.J. and C.L. Peterson, *Role for ADA/GCN5 products in antagonizing chromatin-mediated transcriptional repression.* Mol Cell Biol, 1997. **17**(11): p. 6212-22.

180. Salinero, A.C., et al., *The Mediator co-activator complex regulates Ty1 retromobility by controlling the balance between Ty1i and Ty1 promoters.* PLoS Genet, 2018. **14**(2): p. e1007232.

181. Nishida, Y., et al., *Ty1 retrovirus-like element Gag contains overlapping restriction factor and nucleic acid chaperone functions.* Nucleic Acids Res, 2015. **43**(15): p. 7414-31.

182. Tucker, J.M., et al., *The Ty1 retrotransposon restriction factor p22 targets Gag.* PLoS Genet, 2015. **11**(10): p. e1005571.

183. Tucker, J.M. and D.J. Garfinkel, *Ty1 escapes restriction by the self-encoded factor p22 through mutations in capsid.* Mob Genet Elements, 2016. **6**(2): p. e1154639.

184. Maxwell, P.H. and M.J. Curcio, *Host factors that control long terminal repeat retrotransposons in Saccharomyces cerevisiae: implications for regulation of mammalian retroviruses.* Eukaryot Cell, 2007. **6**(7): p. 1069-80.

185. Nyswaner, K.M., et al., *Chromatin-associated genes protect the yeast genome from Ty1 insertional mutagenesis.* Genetics, 2008. **178**(1): p. 197-214.

186. Griffith, J.L., et al., *Functional genomics reveals relationships between the retrovirus-like Ty1 element and its host Saccharomyces cerevisiae.* Genetics, 2003. **164**(3): p. 867-79.

187. Dakshinamurthy, A., et al., *BUD22 affects Ty1 retrotransposition and ribosome biogenesis in Saccharomyces cerevisiae.* Genetics, 2010. **185**(4): p. 1193-205.

188. Risler, J.K., et al., *Host co-factors of the retrovirus-like transposon Ty1.* Mob DNA, 2012. **3**(1): p. 12.

189. Scholes, D.T., et al., *Multiple regulators of Ty1 transposition in Saccharomyces cerevisiae have conserved roles in genome maintenance.* Genetics, 2001. **159**(4): p. 1449-65.

190. Manhas, S., L. Ma, and V. Measday, *The yeast Ty1 retrotransposon requires components of the nuclear pore complex for transcription and genomic integration.* Nucleic Acids Res, 2018. **46**(7): p. 3552-3578.

191. Rowley, P.A., et al., *Control of yeast retrotransposons mediated through nucleoporin evolution.* PLoS Genet, 2018. **14**(4): p. e1007325.

192. Curcio, M.J., et al., *S-phase checkpoint pathways stimulate the mobility of the retrovirus-like transposon Ty1.* Mol Cell Biol, 2007. **27**(24): p. 8874-85.

193. Scholes, D.T., et al., *Activation of a LTR-retrotransposon by telomere erosion.* Proc Natl Acad Sci U S A, 2003. **100**(26): p. 15736-41.

194. Conte, D., Jr., et al., *Posttranslational regulation of Ty1 retrotransposition by mitogen-activated protein kinase Fus3.* Mol Cell Biol, 1998. **18**(5): p. 2502-13.

195. Conte, D., Jr. and M.J. Curcio, *Fus3 controls Ty1 transpositional dormancy through the invasive growth MAPK pathway.* Mol Microbiol, 2000. **35**(2): p. 415-27.

A

B
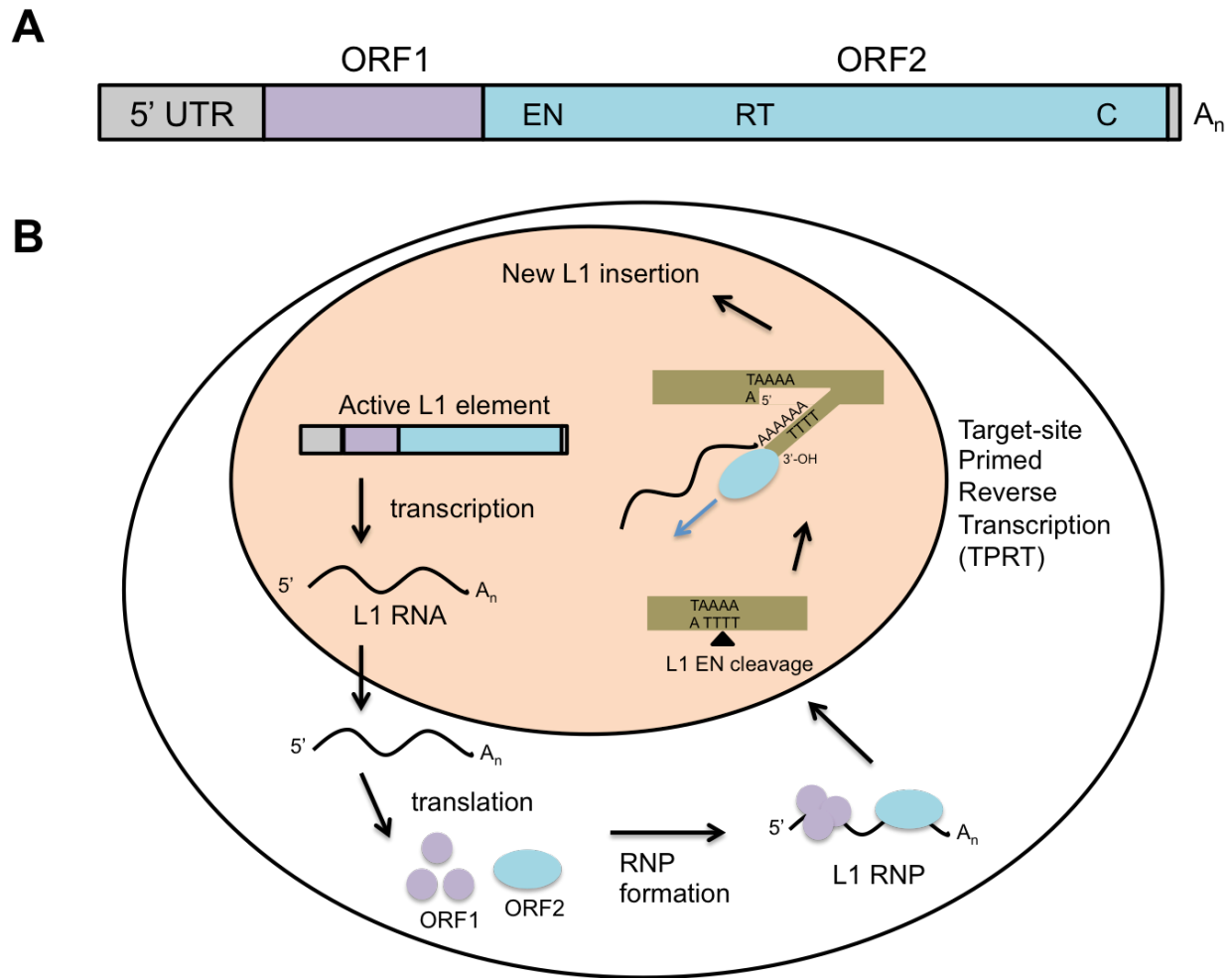
Figure 1.1 Retrotransposition cycle of LINE-1 element.

(A) LINE-1 element. 5′ Untranslated regions (UTRs), open reading frame 1 (ORF1),

open reading frame 2 (ORF2) encoding endonuclease (EN) and reverse transcriptase

(RT), and cysteine-rich domains (C); poly (A) tract (indicated as A_n, downstream of 3′

UTR) are shown. (B) Replication cycle of LINE-1. LINE-1 RNA is transcribed from an

active LINE-1 element, and then exported to the cytoplasm. Upon translation in the

cytoplasm, ORF1p and ORF2p (violet circle and light blue oval, respectively) are

generated. ORF1p and ORF2p bind back to their encoding RNA (*cis*-preference) to form a ribonucleoprotein particle (RNP) complex. Minimally LINE-1 RNA, ORF1p trimers and at least one ORF2p compose LINE-1 RNP complex. The LINE-1 RNP enters the nucleus and gets inserted to the host genome by a mechanism termed target-site primed reverse transcription (TPRT). The EN activity of ORF2p makes a single-strand nick at the genomic DNA target at a consensus site (5′ -TTTT/A- 3′). RT activity of ORF2p uses the exposed 3′ -OH group to initiate first-strand LINE-1 cDNA synthesis using the bound LINE-1 RNA as a template. After undergoing the rest of the steps of TPRT, this leads to the insertion of a new LINE-1 copy.
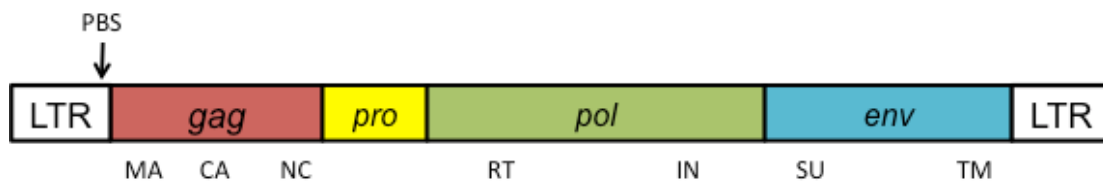
Figure 1.2 Genomic organization of a HERV provirus element.

Two LTRs flank four major open reading frames: *gag*, *pol*, *pro*, and *env*. The primer binding site (PBS), commonly used to classify HERVs, is located upstream of the *gag* reading frame. The Gag polyprotein encodes viral structural components, matrix (MA), capsid (CA), and nucleocapsid (NC). Pol consists of reverse transcriptase (RT) and integrase (IN). The *env* gene encodes surface protein and consists of the surface (SU) and transmembrane (TM) domain that is required for cell entry.
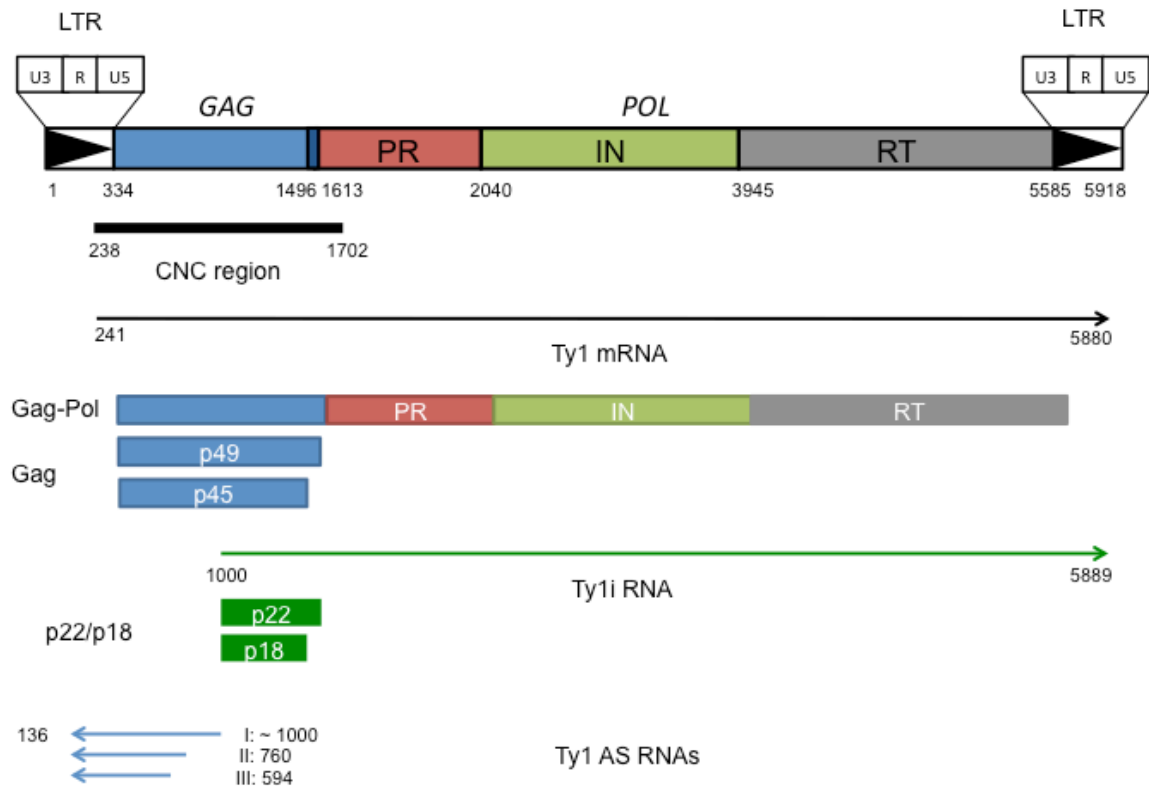
Figure 1.3 Genomic organization of Ty1 element and its gene products.

The LTRs flank two overlapping ORFs: *GAG* and *POL*. *GAG* encodes the capsid protein

Gag. *POL* encodes the enzymes protease (PR), integrase (IN), and reverse

transcriptase (RT). The CNC region (black bar), RNA transcripts (Ty1 mRNA, AS RNAs,

and Ty1i mRNA), translation products from *GAG* (Gag-p49) and *POL* (Gag-Pol-P199),

and processing of *GAG* derived proteins by Ty1 PR (p49 to p45, p22 to p18), are

shown. Numbers indicate nucleotide positions of the Ty1-H3 (accession number

M18706.1).
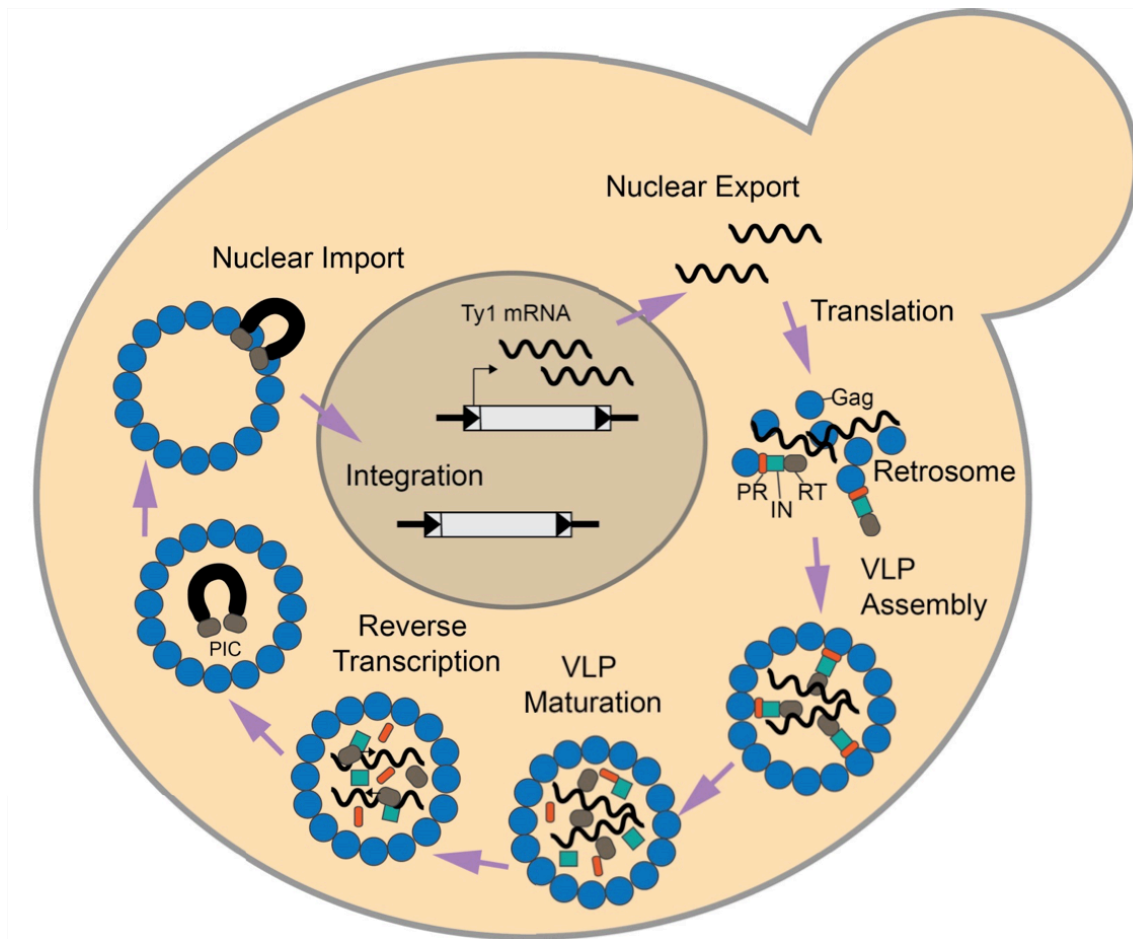
Fig 1.4 Replication cycle of Ty1.

Ty1 mRNA is transcribed by RNA polymerase II and exported to the cytoplasm. Ty1 mRNA is translated into two protein products, Gag-p49 and Gag-Pol-p199. Pol is generated by a programmed +1 frameshifting event near the end of *GAG*. Ty1 mRNA and translated proteins together form a complex called the retrosome. Retrosomes are the assembly sites for Ty1 virus-like particle (VLP). After assembly, VLPs undergo maturation catalyzed by Ty1 PR. As a result, mature Gag-p45, PR, IN and RT are released. Reverse transcription takes place in the mature VLP, using Ty1 mRNA as template and tRNAi^{Met} as a primer. Ty1 cDNA, IN and perhaps other Ty1 proteins form the pre-integration complex (PIC) or intasome, which exits the VLP by an unknown

mechanism. The PIC is imported into the nucleus and inserted into a new genomic

location. Regions upstream of RNA polymerase III transcribed genes (such as tRNA

genes) are favored sites for Ty1 integration. Taken from [176].
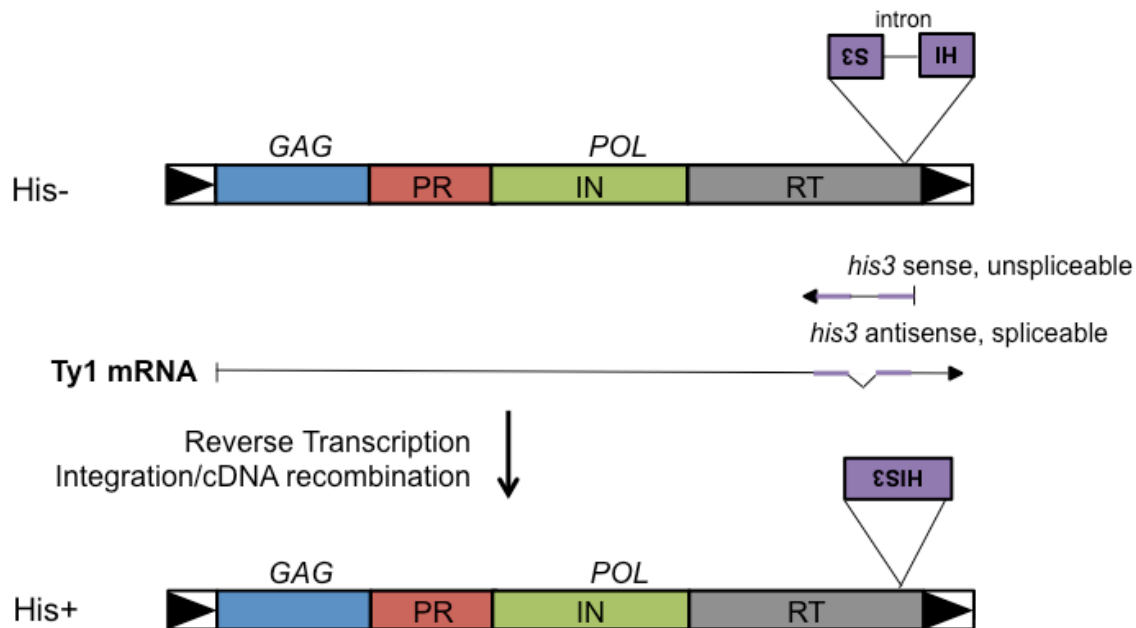
Fig. 1.5 Ty1*his3-AI* retromobility indicator system.

A Ty1 element tagged with the indicator gene *his3-AI* (Ty1*his3-AI*) is used to measure

Ty1 mobility. The *HIS3* gene (including its promoter) is located in the antisense

orientation with respect to Ty1 transcription, and the *HIS3* gene is interrupted by

artificial intron (AI) that is inserted in the sense orientation with respect to Ty1

transcription. Both Ty1 and *his3-AI* transcripts can be transcribed. The AI within the

Ty1*his3-AI* transcript is spliceable, but the AI cannot be spliced from *the his3*-AI

transcript. Neither the Ty1*his3-AI* nor *his3-AI* transcripts lead to synthesis of functional

His3 because Ty1*his3-AI* RNA contains antisense *HIS3* sequences, and *his3-AI* RNA is

disrupted by an unspliceable intron in its coding region. However, spliced Ty1*HIS3* RNA can undergo reverse transcription, and the resulting cDNA can either integrate or recombine with existing Ty1 elements in the host genome. The cells that undergo either integration/cDNA recombination (together referred as Ty1 mobility) are able to grow on medium lacking histidine.
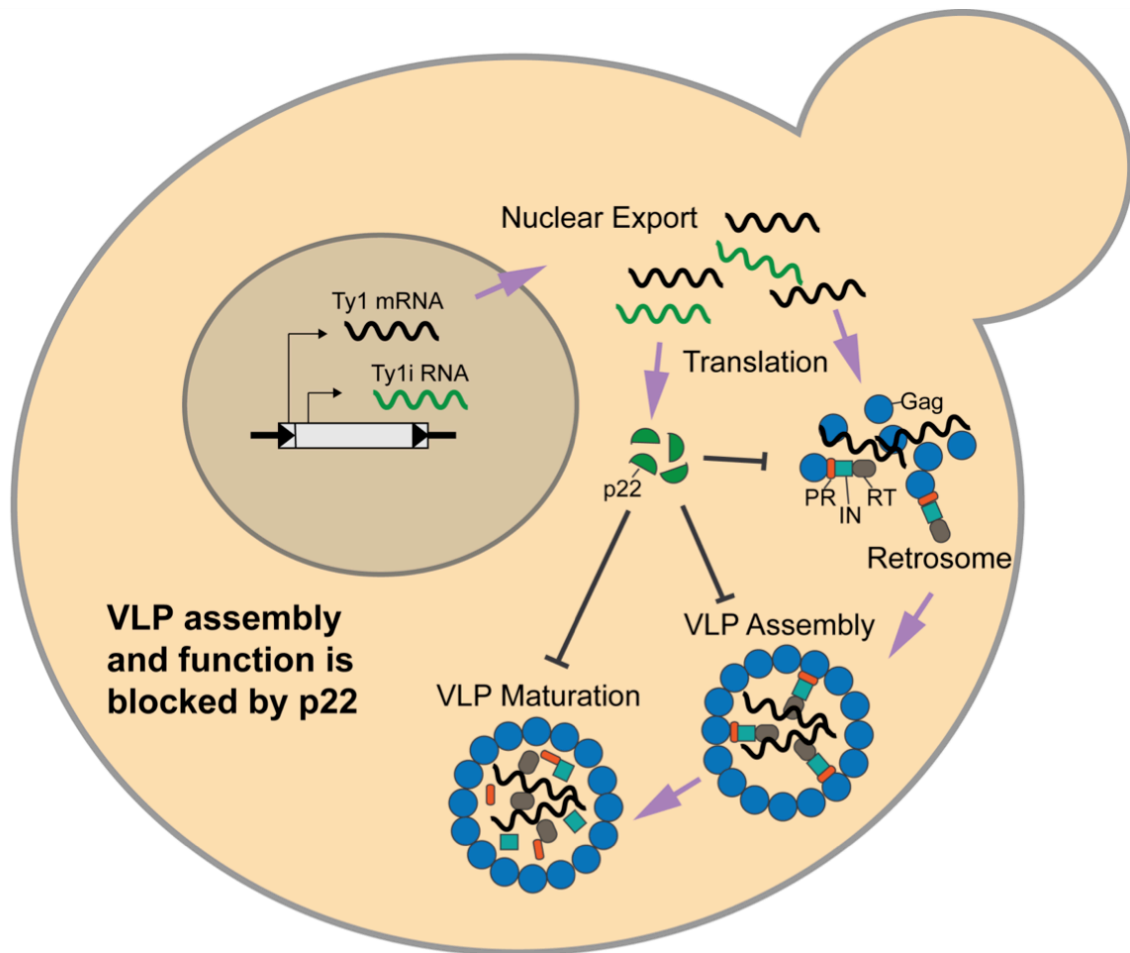
Fig 1.6 Ty1 copy number control mediated by p22.

An abbreviated Ty1 replication cycle with steps affected by p22 (retrosome formation, VLP assembly, maturation and function) is shown. In cells undergoing CNC, Ty1i RNA is transcribed and serves as template for the p22 restriction factor. p22 colocalizes with Gag and disrupts retrosomes. p22 also associates with Ty1 VLPs and alters assembly maturation and function of VLPs. Taken from [176].

CHAPTER 2

RIBOSOME BIOGENESIS MODULATES TY1 COPY NUMBER CONTROL IN

SACCHAROMYCES CEREVISIAE [1]

ABSTRACT

Transposons can impact the host genome by altering gene expression and participating in chromosome rearrangements. Therefore, organisms evolved different ways to minimize the level of transposition. In *Saccharomyces cerevisiae* and its close relative *S. paradoxus*, Ty1 copy number control (CNC) is mediated by the self-encoded restriction factor p22, which is derived from the *GAG* capsid gene and inhibits virus-like particle (VLP) assembly and function. Based on secondary screens of Ty1 cofactors, we identified *LOC1*, a RNA localization/ribosome biogenesis gene that affects Ty1 mobility predominantly in strains harboring Ty1 elements. Ribosomal protein mutants *rps0bΔ* and *rpl7aΔ* displayed similar CNC-specific phenotypes as *loc1Δ*, suggesting that ribosome biogenesis is critical for CNC. The level of Ty1 mRNA and Ty1 internal (Ty1i) transcripts encoding p22 was altered in these mutants, and displayed a trend where the level of Ty1i RNA increased relative to full-length Ty1 mRNA. The level of p22 increased in these mutants, and the half-life of p22 also increased in a *loc1Δ* mutant. Transcriptomic analyses revealed small changes in the level of Ty1 transcripts or efficiency of translation initiation in a *loc1Δ* mutant. Importantly, a *loc1Δ* mutant had defects in assembly of Gag complexes and packaging Ty1 RNA. Our results indicate that defective ribosome biogenesis enhances CNC by increasing the level of p22, and raise the possibility for versatile links between VLP assembly, its cytoplasmic environment, and a novel stress response.

INTRODUCTION

Ty1 is the most abundant retrotransposon in the *Saccharomyces cerevisiae* reference strain at 32 copies per haploid genome, and is an effective model for understanding the mechanism and consequences of retrotransposition [1]. Ty1 shares many similarities with retroviruses, except retrotransposition is not infectious. Ty1 long terminal repeats (LTRs) bracket two overlapping genes; *GAG* encodes the capsid protein of virus-like particles (VLPs), and *POL* encodes a polyprotein that contains protease (PR), integrase (IN) and reverse transcriptase (RT). Once transcribed by RNA polymerase II from the 5′ LTR to the 3′ LTR, Ty1 mRNA is exported to the cytoplasm and translated into two protein products, Gag-p49 and Gag-Pol-p199. Gag-Pol is generated by a programmed +1 frameshifting event near the end of *GAG*. Gag-p49 and Gag-Pol-p199 are present in a ratio of 20:1 [2]. Ty1 mRNA and translated proteins form cytoplasmic granules called the retrosomes, or T-bodies, which are sites where VLPs assemble [3-5]. During or after Ty1 VLP assembly, Gag-p49 and Gag-Pol-p199 undergo maturation catalyzed by PR to form mature Gag-p45, PR, IN and RT. Reverse transcription takes place using packaged dimeric Ty1 RNA as template and a heterodimeric complex of IN and RT [6-8]. Ty1 cDNA and IN likely form a preintegration complex or intasome, which is imported into the nucleus via a nuclear localization sequence present on IN [9-11]. Genomic regions upstream of genes transcribed by RNA polymerase III (such as tRNA genes) are favored sites for Ty1 integration, based on functional interactions between IN and subunits of RNA polymerase III [12-15].

While retrotransposons and retrovirus-like elements comprise almost 50% of the human genome, Ty1, Ty2, Ty3, Ty4, and Ty5 sequences make up ~ 3% of the compact

S. cerevisiae reference genome [16, 17]. Ty1 movement occurs at a low rate of about 1 x 10^{-6} transposition events/Ty1 element/generation [18], yet yeast has the capacity to harbor hundreds of additional Ty1 insertions when cells containing Ty1 fused to the *GAL1* promoter undergo multiple rounds of induction [19]. These observations suggest that cellular or element-encoded products prevent rampant Ty1 retrotransposition in *S. cerevisiae*. Interestingly, none of the silencing pathways used by other eukaryotes to minimize transposon movement such as RNAi or DNA methylation are present in the *S. cerevisiae sensu stricto* group [20-22], implying there may be a novel pathway utilized by *S. cerevisiae* to modulate Ty1 retrotransposition. Copy-number dependent effects on Ty1 mobility, termed copy number control (CNC) [23] were first addressed when a plasmid containing Ty1 tagged with the retromobility indicator *his3-AI* was expressed in a strain of *S. paradoxus* strain that lacks complete Ty1 elements [18, 24]. A dramatically higher level of Ty1 mobility is observed in this Ty1-less strain when compared to the *S. cerevisiae* laboratory strain S288c, which carries 32 Ty1 elements, or when the Ty1-less strain is repopulated with Ty1. Interestingly, the minimal Ty1 sequence required for CNC is located within *GAG* and the 5′ LTR. In cells undergoing CNC, defective VLPs are formed that contain lower levels of mature RT and IN, and reverse transcription products [25, 26].

Although early work suggested that Ty1 antisense transcripts confer CNC [25], extensive biochemical and genetic analyses of the CNC region revealed a self-encoded Ty1 protein termed p22 that is necessary and sufficient for CNC [27]. p22 is encoded by the 3'-half of *GAG* and is produced from subgenomic internally-initiated Ty1(i) transcripts. p22 is processed to p18 by PR cleavage at the same site utilized by the full-

length Gag protein. p22/p18 confers CNC and a trans-dominant negative effect on Ty1 mobility. p22/p18 associates with VLPs, co-immunoprecipitates with Ty1 Gag and colocalizes with Gag in the cytoplasm. p22/p18 disturbs retrosome formation and VLP assembly, and blocks maturation and reverse transcription within the VLPs. In addition, p18 interferes with the nucleic acid chaperone function of Gag-p45 [28], and p22/p18 targets Gag and disrupts Gag-Gag interactions during VLP assembly [29].

Like retroviruses, retrotransposons are capable of making only a few proteins critical for replication; therefore, a variety of host functions modulate their life cycle [30, 31]. Host genes from diverse biological processes affect Ty1 transposition and can stimulate (cofactors) or inhibit (restriction factors) various aspects of the Ty1 life cycle. To date, about 500 cofactor and 110 restriction factor genes have been identified by genome-wide screens using yeast gene deletion mutants [32-35] or transposon mutagenesis [36]. Ty1 cofactors are involved in processes such as transcription, chromatin modification, histone deacetylation, ribosome biogenesis and function, mRNA turnover and transport [33-35]. In contrast, about 80% of Ty1 restriction factors are involved in nuclear processes including DNA repair, transcription, chromatin structure and function, recombination and the cell cycle [32, 35, 36]. We set out to identify host modulators that influence the synthesis or activity of the newly discovered Ty1 restriction factor p22. Such host factors would greatly help us understand the mechanism of CNC.

We reasoned that CNC-specific host factors should stimulate or inhibit CNC in a Ty1 copy number-dependent manner. In other words, a host modulator displaying a CNC-specific phenotype would occur predominantly in CNC^+ cells containing Ty1 elements but not in CNC^- cells lacking Ty1. To identify CNC-specific modulators, we analyzed

previously identified Ty1 cofactor mutants involved in RNA metabolism, which is central

to the process of Ty1 retrotransposition [1], for those that displayed phenotypes strongly

associated with CNC [25, 27]. We then analyzed candidate deletion mutants in

congenic CNC⁻ and CNC⁺ strains that differ only in Ty1 copy number. Surprisingly, we

identified several genes involved in ribosome biogenesis that increase the level p22 and

show CNC-specificity; Ty1 mobility decreased preferentially in CNC⁺ cells when

compared to CNC⁻ cells. In particular, *LOC1*, a gene required for asymmetric

localization of *ASH1* mRNA [37] and rRNA processing [38], facilitates Gag-Gag

interactions and retrosome formation by modulating the level of p22.

MATERIALS AND METHODS

Genetic techniques, media, strains, and plasmids

Strains are listed in Table S2.1. Standard yeast genetic and microbiological

procedures were used in this work [39]. The *S. cerevisiae* CNC⁻ strain DG3453 was

derived from DJ12, which was isolated in Djibouti, Africa and generously provided by

Michel Aigle. A *MATα ho* ascospore (DJ12-1B) was subjected to 5-fluoroorotic acid (5-

FOA) selection [40], and a spontaneous *ura3* mutant was isolated to generate DG3093.

The results of Southern blot analysis [23, 24] suggest there are 10 Ty1 elements in

DJ12-1B and DG3093. To eliminate the Ty1 elements from DG3093, we generated a

Ty1:*URA3* targeting fragment containing Ty1 sequences 1702-3945 [41]. The 2243-

base pair segment was amplified by PCR from DG3093 using primers 1702XhoBglF (5′

- CCCG<u>CTGCAG</u>AGATCTGCTCATCACATACACTC - 3′) and 3945BamR (5′ -

CCGC<u>GGATCC</u>TGCAATCAGGTGAATTCGT - 3′) and subcloned into pSP70

58

(Promega, Madison, WI) digested with XhoI and BglII. *URA3* bracketed by XhoI

restriction sites on pB130 (kindly provided by Gerald Fink) was digested and cloned into

the Ty1 SalI site (nucleotide 2173) of pSP70/Ty1. The resulting plasmid pBDG1374 was

digested with XhoI and EcoRI to release the Ty1:*URA3* targeting fragment. Ura$^+$

transformants containing putative Ty1:*URA3* recombination events with individual Ty1

elements were subjected to 5-FOA selection followed by Southern blot analysis to

detect LTR-LTR recombinants [42]. Changes in Ty1 hybridization patterns suggest that

one Ty1 was eliminated with each round of transformation and 5-FOA selection. After

removing 9 of 10 Ty1 elements, *HIS3* was deleted as described previously using

pBDG652 [23] to generate DG3453. The same Ty1 was present in independent

elimination lineages and could not be lost to generate a Ty1-less strain, raising the

possibility that the remaining Ty1 insertion may be required for the function of an

essential gene. Crosses between DG3093 containing 10 Ty1 elements and strains with

a single Ty1, or crosses between the reference strain BY4741 [43] and the single Ty1

strains displayed 70% spore viability in 18-20 tetrads/cross. Ascosporal progeny

displayed similar variations in growth and expected segregation of genetic markers.

DG3453 was repopulated with approximately 19 Ty1 elements by induction of a pGTy1

plasmid to generate DG3648, as described previously [23]. *MATα* deletion strains [44]

derived from BY4742 [43] were obtained from Invitrogen (Carlsbad, CA). Gene

disruptions in DG3453 and DG3648 were carried out using the *KanMX4* cassette [45].

The *KanMX* disruption cassettes were PCR amplified using template DNA extracted

from deletion mutants with gene-specific primers A and D from the Saccharomyces

Genome Database (https://www.yeastgenome.org/). Correct gene disruptions by

59

KanMX for all of the deletion mutant strains used in this study were confirmed by PCR using locus-specific primers. To analyze Ty1 mobility and protein levels, pBDG606 (pGTy1*his3-AI*/Cen/*URA3*) [34] was introduced to BY4742, DG3453 and DG3648 derived strains.

Ty1*his3-AI* mobility

Quantitative and qualitative Ty1*his3-AI* mobility assays were determined as described previously [18, 23] with minor modifications. For strains expressing pGTy1*his3-AI* (pBDG606), a single colony grown on SC-Ura at 30°C was resuspended in 1 ml of SC-Ura + 2% raffinose and incubated for 24 hr at 30°C with aeration; 200 µl microliters of the raffinose culture was centrifuged, and the cell pellet was resuspended in 1 ml of SC-Ura + 2% galactose in quadruplicate. The cultures were grown for 16 hr at 22°C, washed, diluted, and spread onto SC-Ura and SC-His-Ura plates, which were incubated for 3-4 days at 30°C until colonies formed. The frequency of Ty1*his3-AI* mobility was calculated by dividing the number of His$^+$ Ura$^+$ colonies by the number of Ura$^+$ colonies. For qualitative Ty1*his3-AI* mobility assays, cells patched onto SC-Ura were incubated at 30°C for 2 days. Cells were replica plated onto SC-Ura + 2% galactose medium plates and incubated at 22°C for 16 hr. To detect Ty1*HIS3* mobility events, galactose-induced cells were replica plated to SC-Ura-His followed by incubation for 2-3 days at 30°C.

RNA isolation and northern blot analysis

For strains carrying pBDG606, 10 ml of SC-Ura + 2% raffinose was inoculated with a single colony and the culture was grown at 30°C to an OD$_{600}$ of 2 or up to 24 hr, depending on the strain; 2ml of the raffinose culture was centrifuged and cells were

suspended in 10 ml of SC-Ura + 2% galactose and grown at 22°C for 16 hr. To detect

endogenous Ty1i RNA, a single colony was suspended in 5 ml YEPD and grown at

30°C overnight. The overnight culture was diluted different amounts depending on the

growth of each strain (1:10 for WT, 1:4 for *loc1Δ*, 1:5 for *rps0bΔ*, and 1:8 for *rpl7aΔ*) in a

total of 10 ml fresh YEPD and grown at 22°C for 8 hr. Total RNA was extracted using

the MasterPure yeast RNA purification kit (Epicentre Biotechnologies, Madison, WI) with

modifications as described previously [27]. Poly(A)$^+$ RNA was isolated from 250 μg total

RNA using the NucleoTrap mRNA purification kit (Clontech, Mountain View, CA)

following manufacturer's protocol. Northern blot analysis using ^{32}P-labeled riboprobes

was performed as described previously [27]. Hybridization signals were visualized and

quantified using a STORM 840 phosphorimager and ImageQuant software (GE

Healthcare).

RNA sequencing analyses

RNA sequencing (RNA-seq) and ribosome footprint profiling (Ribo-seq) datasets

were created by S. Komili, D. Muzzey, and FP Roth, and downloaded from GEO

(accession no. GSE34438). Three replicate RNA-seq or Ribo-seq libraries were

analyzed from the wild type parent BY4741 and the isogenic *loc1Δ* derivative (12

libraries in total). Note that deleting *LOC1* confers a transposition defect in BY4741 or

BY4742 [35], which are closely related *MATa* and *MATα* strains, respectively [43].

Sequencing reads were analyzed as follows: reads were trimmed of adaptor

sequences. For Ribo-seq, a size filter of 28 bp was imposed, as 28 bp Ribo-seq reads

tended to have a high proportion of in-frame reads. As Ty1 is present in multiple copies

in the genome, the Ty1-H3 reference element (GenBank M18706.1) was used in the

initial round of mapping. Mapping was performed with STAR RNA-seq aligner (version 2.4.2a) allowing zero mismatches. The remaining (non-Ty1) reads were mapped to the *S. cerevisiae* genome (Release 85; Ensembl). Mapped reads were assigned to genes and plotted with custom scripts. Gene counts are the average of the three replicates for each library, with similar results observed for pairwise library comparisons.

Protein extraction and western blot analysis

Total protein was extracted from strains expressing pGTy1*his3-AI* as previously described [46]. For detecting p22/p18, cells from 2 OD_{600} of culture were processed by trichloroacetic acid (TCA) extraction and immunoblotted as described [27]. Samples were separated on 7.5 or 8% (for detecting IN), 10% (for detecting Gag-isoforms) or 15% (for detecting Gag-isoforms and p22/p18) SDS-PAGE gels. Antibody dilutions were as follows: anti-p18 1:5000, anti-VLP 1:7000, anti-IN 1:2500, anti-Pgk1 1:40000, and anti-TY tag (BB2; UAB Epitope Recognition and Immunoreagent Core, Birmingham AL) 1:20000 [47]. Immunoreactive protein signals were quantified using Quantity One software (Bio-Rad, Hercules, CA).

Stability of p22

p22 stability in wild type and *loc1Δ* strains was determined as described previously [48] with minor modifications; 5 ml of SC-Ura+2% raffinose medium was inoculated with a single colony and grown for ~ 24 hr at 30°C. The raffinose cultures (3 OD_{600} for HWA625 and 6 OD_{600} for HWA626) were added to 50 ml SC-Ura+2% galactose medium and grown at 20°C for to an OD_{600} of 0.6 - 0.8. Thirty OD_{600} of cells per strain were harvested and washed three times in 5 ml of SC-Ura-Met+2% Gal. Click-iT L-HPG (Life Technologies) was added to 5 ml cell suspensions in SC-Ura-Met+2% Gal to a

final concentration of 80 μM. Cultures were incubated at 20°C for 30 min with shaking. Cells were washed twice with 5 ml of SC-Ura-Met+2% Gal and then resuspended in chase medium (SC-Ura-Met+2% Gal +50 mM L-methionine). Equal amounts of cells were harvested at 0, 1, 3 and 6 hr time points. Cell pellets were resuspended in 100 μl HB buffer (25 mM Tris pH 7.5, 125 mM NaCl, 5 mM EDTA, 0.5% IPEGAL) with protease inhibitors (20 μl of 1 mg/ml Aprotonin, Leupeptin, Pepstatin, and 100 μl of 10 mg/ml PMSF) and vortexed at 4°C for 5 min two times with 0.15 g of glass beads. Following addition of 5 μl of 10% SDS, the cell extract was boiled for 5 min and 500 μl ice-cold HB buffer was added. The extract was centrifuged for 20 min at 10,000 g at 4°C and precleared with 50 μl of Protein A agarose (Pierce) slurry (50% beads) washed with HB buffer and incubation at 4°C for 1 hr; 10 μg of anti-p18 antibody was added to precleared lysate and rotated at 4°C for 16 hr, and then 50 μl of Protein A agarose slurry was added to each sample and rotated at 4°C for 2 hr. Samples were washed three times with 500 μl HB buffer and proteins were eluted from the beads with 50 μl of 50 mM Glycine (pH 3.0) equilibrated in 200 μl of 50 mM Tris-Cl (pH 8.0), 1% SDS. Eluted proteins were precipitated by methanol/chloroform. Click-iT Cell Reaction Buffer Kit (Life Technologies) was used to conjugate TAMRA (Life Technologies) to HPG-labeled p22 according to the manufacturer's protocol. Protein samples were separated on 15% SDS-PAGE gels. The fluorescent signal of TAMRA-conjugated p22 was detected using a Typhoon trio scanner with 580 BP30 filter, and p22 bands were quantified with ImageQuant Software (GE Healthcare). We assumed that p22 degradation follows first-order kinetics and determined p22 half-life as described previously [49].

<u>Fluorescence *in situ* hybridization/Immunofluorescence (FISH/IF)</u>

Retrosomes were analyzed by FISH/IF as described previously [5, 27, 50]. To visualize endogenous retrosomes, 5 ml YEPD cultures were inoculated with a single colony and grown for 16 hr at 30°C. The overnight cultures were diluted 40- (HWA215) or 500-fold (BY4742) into fresh YEPD and grown at 20°C for ~24 hr to an OD_{600} of 0.8 to 1.0. To visualize pBDG606- induced retrosomes, 5 ml of SC-Ura+2% raffinose medium was inoculated with a single colony and grown for 24 hr at 30°C. The raffinose cultures were diluted to 20- (HWA169) or 80-fold (HWA15) in SC-Ura+2% galactose medium and grown at 20°C for ~24 hr to an OD_{600} of 0.8 to 1.0. Formaldehyde was added directly to the culture to a final concentration of 4% and allowed to fix for 1 hr. Anti-VLP (rabbit polyclonal, 1:2000) and a cocktail of three GAG digoxigenin (DIG)-labeled antisense oligonucleotides (TyA_380, 5´ - GCCTTCTCACATTCTTCTGTTTTGGAAGCTGAAACGTCTAACGGATCTTG – 3´; TyA_444, 5´ - TTCTCTGGAACAGCTGATGAAGCAGGTGTTGTTGTCTGTTGAGAGTTA - 3´; TyA_545, 5´ - CAACCAGATGGATTGGCTTGGTTTTGGGTCATCATGCACTGCTGTGGGTA – 3´) were used to detect Gag and Ty1 mRNA, respectively. Secondary antibodies used here were anti-rabbit-AF594 (1:200; Life Technologies) and Fluorescein (FITC) conjugated sheep anti-DIG Fab fragment (1:200; Roche Applied Science). Image acquisition was carried out using a ZeissAxio Observer microscope equipped with an AxioCam HSm camera, and images were analyzed with AxioVision v4.6 software (Carl Zeiss Microscopy, Thornwood, NY). The percentage of cells with retrosomes was calculated

64

by dividing the number of cells with colocalized foci (Ty1 mRNA and Gag) by the total number of cells. At least 200 cells were analyzed for each strain.

Sucrose gradient sedimentation

Sedimentation of Gag complexes was analyzed as described [29] with the following modifications. For strains BY4742 and HWA215, a 100-ml culture was grown as described for FISH/IF microscopy. Equal amounts of total protein (7-8 mg in 300–450 µl) were applied to a 7–47% continuous sucrose gradient and centrifuged at 25,000 rpm in a SW 41 swinging bucket rotor (Beckman Coulter, Brea, CA) for 3 hr at 4°C. Nineteen 0.6 ml fractions were collected and an equal volume of each fraction was subjected to western blot analysis to detect Ty1 Gag as described above.

Nuclease protection

Strain BY4742 and HWA215 were grown as described above for FISH/IF microscopy. Nuclease protection using benzonase (EMD Millipore, Billerica, MA) was performed as described previously [28]. Total RNA extraction and northern blot analysis was performed as described above.

Crosslinking Gag complexes

HWA15 and HWA169 were grown as described above for FISH/IF microscopy. Formaldehyde (Fisher Scientific, Waltham, MA) was added directly to the culture medium to a final concentration of 1%. Fixed cells were incubated for 1 hr at 20°C. Proteins were extracted as previously described [51] and separated on NuPAGE 3–8% Tris-Acetate gels (Invitrogen, Carlsbad, CA) for detecting Gag multimers and 10% SDS-PAGE gels for detecting Pgk1. An equal amount of protein extract (7.5 µg) was loaded per lane and electrophoresis was performed at 120 volts for 2 hr. Proteins were

transferred to PVDF membranes and probed with anti-TY tag at a 1:50000 dilution. Gag_1 quantification was performed as described above.

Data availability

The authors state that all data necessary for confirming the conclusions presented in the article are represented fully within the article. All strains and reagents are available upon request.

<div align="center">RESULTS</div>

Identifying CNC-specific Ty1 cofactors

To identify CNC-specific genes, 498 Ty1 cofactors were analyzed using DAVID (http://david.abcc.ncifcrf.gov/), which enables functional annotation of large gene lists [52]. Functional annotation clustering was performed to determine the gene ontology (GO) terms enriched in Ty1 cofactor genes with reduced redundancy (for example, nucleic acid transport and RNA export were grouped in one cluster). Among the top four gene clusters with high enrichment scores (> 4) (Table S2.2) were Ty1 modulators involved in post-transcriptional RNA biogenesis (RNA catabolic processes, RNA transport). To identify CNC-specific cofactors, 51 Ty1 cofactor mutants with defects in RNA metabolism were subjected to two secondary screens (Table 2.1). First, mutants involved in CNC should contain much less mature IN relative to Gag since CNC is associated with a block in the accumulation of mature Pol proteins [23, 25]. Second, CNC-specific mutations should affect Ty1 mobility to a greater extent in CNC^+ strains containing Ty1 elements vs. CNC^- strains that do not.

Selected Ty1 cofactor mutants were assessed for the level of Ty1 Gag and IN by western blot analysis using VLP and IN antisera (Figure S2.1A and Table S2.3). Transposition defects were monitored by a qualitative papillation assay for Ty1*his3-AI* mobility [18] (Table S2.3). To reliably detect mature Ty1 IN but not overwhelm CNC by overexpressing Ty1 (Figure S2.1B) [23], a low-copy centromere-based *GAL1* expression plasmid pGTy1*his3-AI* plasmid (pBDG606) was introduced into the Ty1 cofactor deletion mutants [44]. For example, deletion of *LOC1*, a gene involved in biogenesis of 60S ribosomes and asymmetric localization of *ASH1* RNA [38, 53], resulted in the preferential loss of mature IN relative to Gag following galactose induction of pBDG606 (Figure 2.1). Also, when wild type and *loc1*Δ cells were monitored qualitatively for Ty1*his3-AI* mobility after induction on SC-Ura + Galactose plates, the *loc1*Δ mutant harbored few Ty1*HIS3* mobility events when compared to the wild type parent. Two mutant classes were evident from the western blot analysis (Figure S2.1A) and qualitative mobility analyses (Table S2.3). We identified nine candidates (*kap123*Δ, *loc1*Δ, *mot2*Δ, *mrt4*Δ, *nup170*Δ, *pap2*Δ, *ref2*Δ, *ssn3*Δ, and *xrn1*Δ) with considerably less mature IN relative to Gag and the Pgk1 loading control, and a low level of Ty1*his3-AI* mobility. Also, we identified mutants such as *npl3*Δ and *sto1*Δ that displayed a lower level of both Gag and IN. These mutants were not studied further as they may have less Ty1 mRNA or defects in Ty1 protein synthesis or stability.

We deleted 9 CNC-specific candidate genes in congenic *S. cerevisiae* CNC⁻ and CNC⁺ strains DG3453 (+1 Ty1) and DG3648 (+20 Ty1s), respectively. During strain construction, mutants showing CNC-independent Ty1 mobility defects (*kap123*Δ) or no defect in Ty1 mobility (*ref2*Δ, *nup170*Δ), clonal variability in Ty1 mobility (*pap2*Δ), or a

severe growth defect (*mot2Δ*) were not pursued further. The remaining four mutants (*ssn3Δ*, *mrt4Δ*, *loc1Δ*, and *xrn1Δ*) were assessed for Ty1*his3-AI* mobility (Figure S2.2A) and Ty1 protein level (Figure S2.2B) in the CNC$^-$ and CNC$^+$ backgrounds. Mutants showing a decrease (*ssn3Δ*) or possible increase (*xrn1Δ*) of IN in the CNC$^-$ background or a relatively mild decrease of IN in the CNC$^+$ background (*mrt4Δ*) compared to wild type were not pursued further. Only *loc1Δ* displayed both phenotypes expected for a CNC-specific mutant: a severe decrease in mature IN relative to Gag, and a decrease in Ty1*his3-AI* mobility in CNC$^+$ but not in CNC$^-$ strains. Therefore, we analyzed how *LOC1* modulated Ty1 CNC in greater detail.

Ribosome biogenesis affects Ty1 CNC

LOC1 is important for ribosome biogenesis and asymmetric RNA localization [37, 38]. We hypothesized that if ribosome biogenesis or asymmetric RNA localization were critical for CNC, disrupting additional genes in those pathways would phenocopy the *loc1Δ* mutant. For the *ASH1* mRNA localization pathway, *PUF6* and *SHE2* were selected as representatives.[37]. For ribosome biogenesis, the Ty1 cofactors *RPS0B* and *RPL7A* as well as their paralogs *RPS0A* and *RPL7B* were analyzed for CNC-specificity [34, 35, 54]. Deletion mutants were generated in *S. cerevisiae* CNC$^-$ and CNC$^+$ backgrounds, and pBDG606-induced Ty1 mobility was determined (Table 2.2). The frequency of Ty1 mobility in CNC$^-$ and CNC$^+$ strains was assessed individually and also reported as a CNC$^-$/CNC$^+$ ratio, which estimates the increase in CNC observed in mutant strains when compared with wild type. The *loc1Δ* mutant displayed a 3.4- and a 117.6-fold decrease in Ty1*his3-AI* mobility in the CNC$^-$ and CNC$^+$ backgrounds, respectively. The increase in CNC as estimated by the *loc1Δ* CNC$^-$/CNC$^+$ ratio was over

300-fold, which was ~30-fold higher CNC than observed in the wild type. Interestingly, the ribosomal protein paralog mutants *rps0a*Δ and *rps0b*Δ, and *rpl7a*Δ and *rpl7b*Δ all showed an increase in CNC. However, the mutants identified as Ty1 cofactors (*rps0b*Δ and *rpl7a*Δ) had higher levels of CNC than their paralogs (*rps0a*Δ and *rpl7b*Δ), which were not recovered as cofactors. *PUF6*, previously identified as a Ty1 cofactor [35], displayed a CNC ratio of 69. In contrast, Ty1 mobility and CNC remained at wild type levels in cells lacking *SHE2*, which neither modulates Ty1 transposition nor affects ribosome biogenesis. In addition, deleting other genes (*myo4*Δ, *she3*Δ, and *scp160*Δ) involved in asymmetric localization of *ASH1* mRNA [37] did not affect Ty1*his3-AI* mobility in CNC⁻ or CNC⁺ strains (Figure S2.3). Overall, our results show that defects in ribosome biogenesis trigger an increase in Ty1 CNC while defects in mRNA localization do not.

Ty1 transcript and protein levels in candidate CNC mutants

Total RNA and protein from wild type and *loc1*Δ CNC⁻ and CNC⁺ strains were subjected to northern (Figure 2.2A) and western blot (Figure 2.2B) analyses following pBDG606 induction. The level of Ty1*his3-AI* and Ty1 transcripts decreased modestly in wild type and *loc1*Δ CNC⁺ strains when hybridized with a ³²P-labeled riboprobe derived from *GAG* and normalized to actin (*ACT1*) mRNA. The level of chromosomal Ty1 transcripts from the single copy CNC⁻ strains was below the limit of detection, and the level of pGTy1-induced Ty1*his3-AI* mRNA decreased slightly in the CNC⁻ *loc1*Δ mutant. Western blot analysis showed that the levels of IN precursors and mature IN remained about the same in the *loc1*Δ CNC⁻ mutant. However, mature IN was not detected in the *loc1*Δ CNC⁺ mutant, which validates the results obtained in the BY4742 knockout strain

(Figure 2.1, Figure S2.1A). The level of Gag isoforms was also similar in the CNC$^-$ and

CNC$^+$ loc1Δ mutants. Together, our results suggest that the CNC-specific loss of

mature IN is enhanced in a loc1Δ mutant while Ty1 mRNA level is comparable to wild

type.

To determine if the level of Ty1 Gag and IN proteins was similar in strains defective

in ribosome biogenesis or RNA localization, total protein from wild type, loc1Δ, rps0bΔ,

rpl7aΔ, she2Δ and puf6Δ strains expressing pBDG606 was subjected to western blot

analysis (Figure 2.3). A trend emerged suggesting that mutants showing CNC-

specificity (loc1Δ, rps0bΔ, and rpl7aΔ) (Table 2.2) contain less mature IN only in the

CNC$^+$ background, except the puf6Δ mutant showed a decrease of mature IN in both

CNC$^-$ and CNC$^+$ backgrounds. Normal levels of Gag and mature IN were present in a

CNC$^-$ and CNC$^+$ she2Δ mutant. Therefore, LOC1 likely modulates Ty1 transposition via

its role in ribosome biogenesis and not through RNA localization.

Ty1i RNA and p22/p18 levels are altered in CNC-specific mutants

Suresh et al. (2015) [55] reported that Ty1i RNA and p22/p18 levels increase in

several ribosome biogenesis mutants including rps0bΔ. Therefore, endogenous levels

of Ty1 mRNA, Ty1i RNA, Gag, and p22/p18 were determined in loc1Δ, rps0bΔ and

rpl7aΔ CNC mutants (Figure 2.4). To clearly distinguish Ty1i RNA from Ty1 mRNA [27,

55], poly(A)$^+$ RNA was subjected to northern blot analysis using ^{32}P-labeled riboprobe

containing Ty1 nucleotides 1266-1601 that hybridizes with Ty1 mRNA and Ty1i

transcripts; ACT1 mRNA served as a loading control (Figure 2.4A). Ty1 transcripts were

below the limit of detection in wild type or CNC$^-$ mutants since there is only one Ty1

element present in the CNC$^-$ background. The level of Ty1 mRNA decreased fivefold in

the CNC$^+$ *loc1Δ* background relative to wild type, whereas a similar decrease in Ty1

mRNA was not observed in a BY4742 *loc1Δ* mutant. We did not examine this strain-

specific effect further. Ty1 mRNA levels also varied modestly in the CNC$^+$ *rps0bΔ* and

rpl7aΔ mutants. The level of poly(A)$^+$ Ty1i RNA was altered in the CNC$^+$ *loc1Δ*, *rps0bΔ*,

and BY4742 *loc1Δ* mutants; slightly increased levels were detected in the CNC$^+$ *rps0bΔ*

and BY4742 *loc1Δ* mutants, and a decreased level was detected in the CNC$^+$ *loc1Δ*

mutant. The level of Ty1i RNA remained unchanged in a *rpl7aΔ* mutant. Although there

was no clear pattern of changes for Ty1i and Ty1 poly(A)$^+$ transcripts in the ribosome

biogenesis mutants, the CNC-specific mutants displayed a trend where the ratio of

Ty1i/Ty1 mRNA increased twofold. We also compared the ratio of Ty1i RNA and Ty1

mRNA in each lane without normalizing to the *ACT1* mRNA loading control. In the CNC$^+$

background, Ty1i/Ty1 mRNA ratio for wild type was 0.4, while the ratio for ribosome

biogenesis mutants were 0.9, 0.8, and 0.8 for *loc1Δ*, *rps0bΔ*, and *rpl7aΔ*, respectively.

In BY4742, the Ty1i/Ty1 mRNA ratio for wild type was 0.5, while the ratio for *loc1Δ* was

0.7. Taken together, the transcript ratios support the idea that the enhancement of CNC

correlates with an increase in Ty1i RNA over Ty1 mRNA.

   To determine the level of endogenous Gag and p22/p18 in the CNC$^-$ and CNC$^+$

strains, total protein was subjected to western blot analysis using an antiserum that

detects both Gag and p22/p18 (Figure 2.4B) [27]. In accordance with the levels of Ty1

RNAs (Figure 2.4A), full-length Gag and a low level of p22/18 were detected in CNC$^+$

wild type but not in CNC$^-$ wild type strains. In the CNC$^+$ *loc1Δ* mutant, we observed a

decrease in full-length Gag and a 6.8-fold increase in p22 when compared to the CNC$^+$

wild type. An increase in p22 of 3.5-fold, 4.2-fold, and 19.4-fold was observed in the

CNC[+] *rps0b*Δ and *rpl7a*Δ mutants, and the BY4742 *loc1*Δ mutant, respectively. Also, there was a 5- to 10-fold decrease in the level of p18 in the CNC[+] ribosome biogenesis mutants. Some p18 was detected in the BY4742 *loc1*Δ mutant, although p22 was present in a much higher amount. Recent work suggests that PR-mediated cleavage of p22 to p18 takes place in VLPs [27, 29]. Therefore, the lower level of p18 observed in the ribosome biogenesis mutants is consistent with a defect in VLP assembly or function.

Since Ty1 Gag stability decreased in several Ty1 cofactor mutants defective in ribosome biogenesis and translation [48], we examined p22 stability in wild type and an isogenic *loc1*Δ mutant in the BY4742 background by pulse-chase immunoprecipitation (Figure S2.4). Due to the low level of endogenous p22 and p18 in BY4742, a previously characterized galactose-inducible p22 expression plasmid pBDG1565 (pAUG1p22/*GAL1*/2μ/*URA3*) [28] was used to assess p22 stability. The half-life of p22 increased from 2.6 hr in wild-type cells to 4.1 hr in the isogenic *loc1*Δ mutant. Although a ~60% increase in half-life suggests that p22 may be more stable, this result does not fully account for the increased level of p22 in the CNC[+] or BY4742 *loc1*Δ mutants. We did not detect p18 in strains expressing pBDG1565, presumably because of the low-level of endogenous Ty1 protein processing by PR [56].

Transcriptomic analyses in a *loc1*Δ mutant

To determine if the changes in Ty1 transcript levels are part of a global response to a defect in ribosome biogenesis, we analyzed publicly available RNA-seq and ribosome profiling (Ribo-seq) datasets generated in BY4741 and a *loc1*Δ derivative (GEO accession no. GSE34438). First, RNA-seq datasets were analyzed for transcriptional

changes. It is not possible to identify unique Ty1i reads from Ty1 as Ty1i transcripts

overlap completely with full-length Ty1 mRNA [27]. We split the RNA-seq counts into

two regions: (1) Ty1 exclusive and (2) Ty1 or Ty1i-derived reads (Figure 2.5A and

2.5B). We observed a slight decrease in counts assigned to both regions in the *loc1Δ*

mutant. However, as Ty1 mRNA is more abundant than the Ty1i transcripts (Figure

2.4A) [27, 55], even a large fold change in Ty1i RNA might be undetectable by RNA-

seq. In agreement with this idea, only a modest increase in Ty1i RNA relative to Ty1

mRNA was observed in the BY4742 *loc1Δ* mutant as determined by northern analysis

(Figure 2.4A).

Previous work suggested that Loc1 represses translation of *ASH1* mRNA [57].

Therefore, we analyzed a Ribo-seq dataset generated in BY4741 and an isogenic *loc1Δ*

mutant to determine if a specific increase in translation initiation can account for the

increase of p22 in the *loc1Δ* mutant (Figure 2.5A and 2.5C). The Ribo-seq reads were

split according to region and frame: *GAG* (0-frame), *GAG* or p22 (0-frame), and *POL*

(+1 frame). Ribo-seq coverage in these regions was not altered in the *loc1Δ* mutant

compared to wild type. Thus, there are no remarkable changes in gene expression

profiles to suggest a reason for observed CNC phenotype or increased p22 levels in the

loc1Δ mutant.

There were 67 and 102 genes where RNA-seq transcript levels increased or

decreased more than twofold in a *loc1Δ* mutant, respectively, with a false discovery rate

(FDR)-adjusted P-value (padj) < 0.05 (Table S2.4). Functional annotation clustering was

performed to find which GO terms are enriched among these transcripts (Table S2.4). In

general, expression of genes involved in growth-related processes such as cell wall

structure, glycolysis/gluconeogenesis, ion transport, and amino acid biosynthesis were affected by the absence of *LOC1*. It remains to be determined whether these effects represent primary or secondary consequences of deleting *LOC1*.

<u>Gag-Gag interactions and Ty1 mRNA packaging are altered in a *loc1Δ* mutant</u>

Formaldehyde cross-linking and the sedimentation pattern of Gag complexes was used to investigate the role of *LOC1* in VLP assembly. Formaldehyde treatment of cells captures HIV-1 Gag assembly intermediates at the plasma membrane [58]. However, our application is simplified since retrosome/VLP assembly takes place in the cytoplasm [4, 5, 48]. To help ensure there are comparable levels of Gag prior to crosslinking, we induced pBDG606 expression in wild type BY4742 and an isogenic *loc1Δ* mutant (Figure 2.6). Following galactose induction, yeast cells were treated with formaldehyde and total cell extracts were separated on a 3-8% gradient gel. Gag complexes and a Pgk1 loading control were detected by western blot analysis. In the absence of formaldehyde, Ty1 Gag was mostly present in a monomeric form (Gag_1) (Figure 2.6A). We also detected a small amount of dimeric Gag (Gag_2) and the 199 kDa Gag-Pol precursor (denoted by an asterisk). In wild type cells treated with formaldehyde, there was an ordered accumulation of Gag multimers with Gag_2 and trimeric Gag_3 as the most prominent forms. When the *loc1Δ* mutant was treated with formaldehyde, a comparable amount of Gag_1 was present when compared to wild type; however, there was a lower level of Gag_2 and a marked decrease in higher order multimers, which is also evident in densitometric tracings (Figure 2.6B). There were also similar levels of Gag_1 in wild type and *loc1Δ* cells when compared with Pgk1. Our results suggest that *LOC1* facilitates early steps in VLP assembly as monitored by Gag multimerization.

74

To visualize more complex assembly intermediates, total protein extracted from BY4742 wild type and *loc1Δ* cells was fractionated through 7-47% sucrose gradients, and the sedimentation pattern of endogenous Gag was monitored by western blot analysis (Figure 2.7A). The majority of Gag from wild type cells reproducibly sedimented as larger complexes (fractions 11-19) likely comprised of VLP assembly intermediates, since few endogenous VLPs are detected in the absence of pGTy1 expression [5]. In contrast, Gag from the *loc1Δ* mutant reproducibly sedimented as smaller complexes (fractions 7-10), suggesting that higher order Gag complex formation is also defective in a *loc1Δ* mutant.

To examine if Ty1 mRNA packaging into nuclease-resistant Gag assemblies is defective [26, 28], cell extracts from BY4742 wild type and *loc1Δ* were treated with the nuclease benzonase followed by northern blot analysis (Figure 2.7B). Hybridization signals from the benzonase treated samples were normalized to untreated controls and *ACT1* mRNA served as a control for RNA integrity and benzonase activity. Ty1 mRNA level reproducibly decreased about threefold when extracts from wild type cells were treated with benzonase, whereas Ty1 mRNA decreased more than sixfold when extracts from the *loc1Δ* mutant were treated with the nuclease. These results suggest that less Ty1 mRNA is packaged into nuclease resistant Gag complexes in the absence of *LOC1*.

Retrosomes are not detected in *loc1Δ* cells

BY4742 wild type and *loc1Δ* cells were subjected to FISH/IF to visualize endogenous (Figure 2.8) or pBDG606 induced retrosomes (Figure S2.5). VLP antiserum was used to detect Gag, and a cocktail of DIG-labeled oligonucleotide *GAG*

probes was used to detect full-length Ty1 mRNA [27]. About 34% of wild type cells contained retrosomes, which are defined as distinct foci containing Gag and Ty1 mRNA [3, 5]. Retrosomes were not detected in the *loc1Δ* mutant. We also failed to observe nondistinct punctate staining for both Ty1 mRNA and Gag, which is present in other Ty1 cofactor mutants [4, 5]. Instead, diffuse cytoplasmic staining was detected in the *loc1Δ* mutant. Furthermore, retrosomes were also absent when pBDG606 was expressed in *loc1Δ* cells (Figure S2.5). Taken together, our results indicate that steps leading to assembly of functional VLPs are disrupted in a *loc1Δ* mutant.

DISCUSSION

Genome-wide screens have revealed cellular modulators of Ty1 retrotransposition and some are conserved with cellular factors involved in retroviral replication [30, 31]. Since all the genetic screens for Ty1 cofactors [33-35] were performed in CNC[+] strains containing genomic Ty1 elements and producing the p22 restriction factor [27], a subset of previously identified Ty1 cofactors may affect the level or activity of p22. RNA metabolism is central to the process of retrotransposition; therefore, cellular processes affecting the Ty1 life cycle between RNA transcription and VLP assembly (Ty1 RNA export, translation, and retrosome formation) may contribute to CNC. Here, we utilized a candidate gene approach to identify CNC-specific modulators based on informative secondary screens that monitor preferential loss of mature IN and Ty1 mobility in a CNC[+] strain compared to an isogenic CNC[-] strain. One weakness of a candidate gene approach is that not all possible causative genes may be queried [59]. However, the hundreds of known Ty1 modulators [32-36] provide a rich source of interactions

76

between Ty1 and its host that can be mined by additional phenotypic analyses. Our analysis of CNC-specific modulators revealed an unexpected connection between Ty1 CNC and a large group of host cofactors involved in ribosome biogenesis and function [55].

The secondary screens identified *LOC1* as a CNC-specific Ty1 cofactor since the level of mature Ty1 IN and retromobility decreased much more in CNC$^+$ than in CNC$^-$ strains. *LOC1* encodes a nucleolar protein that binds RNA [53], associates with the pre-60S ribosome complex [60], and is necessary for rRNA processing, 60S assembly and nuclear export of pre-60S subunits [38]. Loc1 also helps localize *ASH1* mRNA to the distal tip of anaphase cells [53] via a direct interaction with She2 [61]. To determine whether *LOC1* modulated Ty1 transposition by altering RNA localization or ribosome biogenesis, we analyzed two other genes implicated in *ASH1* mRNA dynamics, *PUF6* and *SHE2*, for their ability to phenocopy the alteration of CNC observed with *LOC1*. Disrupting *SHE2* leads to complete loss of *ASH1* mRNA localization, and Puf6 interacts with Loc1 in the nucleolus and has a role in translational repression of *ASH1* mRNA [37]. Our data suggest that the *ASH1* mRNA localization pathway is not involved in CNC because deleting a key component of the pathway, *SHE2*, does not affect Ty1 mobility, confer a CNC-specific transposition defect, or result in loss of IN. Furthermore, several other genes implicated in *ASH1* mRNA localization (*MYO4*, *SHE3*, and *SCP160*) do not affect Ty1 CNC (Figure S2. 3).

Like *LOC1*, *PUF6* is involved in both *ASH1* mRNA localization as well as 60S ribosome biogenesis [62]. The *puf6Δ* mutant showed a decrease of IN in both CNC$^-$ and CNC$^+$ strains, but its transposition defect was CNC-specific. Perhaps *PUF6* is required

for the process of Ty1 retrotransposition and is also involved in CNC. Interestingly, Loc1 function may be compromised in the absence of Puf6 since these proteins interact [63, 64]. While their interaction is important for *ASH1* mRNA localization and 60S ribosome assembly, previous studies suggest that Loc1 and Puf6 carry out distinct functions. For example, *puf6Δ*, *loc1Δ*, and a *puf6Δ loc1Δ* double mutant each display different phenotypes with respect to the localization of ribosomal protein paralogs, and assembly of Rpl43 in the 60S ribosome subunit [57, 64].

Analysis of *LOC1* focused our attention on how ribosome biogenesis affects Ty1 transposition and CNC. Remarkably, 71 of 458 host genes identified in multiple genetic screens for Ty1 modulators [32-36] encode ribosomal subunits, ribosome biogenesis factors, and translation factors. Recent work suggests that these genes affect multiple steps during transposition, including translation initiation, programmed frameshifting, protein stability, and subcellular protein localization [55]. A rare tRNA-Arg(CCU) gene *HSX1* specifically involved in programmed frameshifting is also required for Ty1 mobility [2]. *rpl21Δ*, *rpl27aΔ*, *rpl39Δ*, and *rps0b* mutants all show transcriptional patterns where Ty1i RNA increases relative to Ty1 mRNA. Additionally, Ty1 mRNA and Gag localization in retrosomes is absent in a *rpl7aΔ* mutant, further implicating ribosome biogenesis in Ty1 retrotransposition [48].

Ribosome biogenesis and nucleolar functions influence the replication of retroviruses and other retrotransposons. In addition to its primary function in ribosome biogenesis, the nucleolus is involved in ribonucleoprotein remodeling, cell cycle progression, and stress response [65]. Cellular stress and viral infection change the organization, composition, and morphology of the nucleolus [66]. The nucleolus is the

target for viral infection by RNA viruses [67, 68], and nucleolar trafficking of viral proteins occurs with retroviruses such as Rous Sarcoma Virus (RSV), Mouse Mammary Tumor Virus (MMTV), and HIV [69-71]. Ty3 and L1 retrotransposon proteins [72, 73] and Alu RNA [74] display nucleolar localization. Moreover, ribosomal protein Rpl7 affects HIV-1 Gag nucleic acid chaperone activity [75], and Rpl9 helps traffic MMTV Gag proteins [70, 75, 76]. Taken together, our work provides additional evidence linking ribosome biogenesis and retroelement replication.

We expanded the CNC analysis to include the ribosomal protein gene paralogs *RPS0A* and *RPS0B*, and *RPL7A* and *RPL7B*. As is the case for other paralogs, only *RPS0B* and *RPL7A* were recovered as Ty1 cofactors [34, 35]. *rps0b∆* and *rpl7a∆* mutants showed loss of IN and a lower level of Ty1 mobility in CNC$^+$ but not in CNC$^-$ strains, demonstrating that these mutants are CNC-specific (Table 2.2). However, the *rpl7b∆* and *rps0a∆* paralogs also displayed CNC-specificity but at a lower level than *rpl7a∆* and *rps0b∆*. Expression of these ribosomal protein genes may account for these results. The level of *RPL7A* mRNA is fourfold higher than *RPL7B,* and *RPS0B* mRNA is 1.5-fold higher when compared to *RPS0A* [57], which correlates with CNC$^-$/CNC$^+$ ratios of Ty1*his3-AI* mobility (Table 2.2). Based on mRNA expression, Rpl7a may be the predominant ribosomal protein in the cell. Therefore, the absence of Rpl7a caused a more severe CNC-specific transposition defect. However, losing Rpl7b does not affect CNC because there is still a substantial amount of Rpl7a present in a cell. For Rps0a and Rps0b, comparable amounts of protein may be expressed from each gene. The absence of one *RPS0* paralog could impact CNC, as suggested by the mobility defect observed in both CNC$^+$ *rps0a∆* and *rps0b∆* mutants. Our results also agree with recent

79

work regarding dosage-dependent rather than paralog-specific function of Rpl7a and Rpl7b on different cellular phenotypes, including Ty1 life cycle stages [54]. Taken together, any paralog-specific effects of ribosomal proteins on CNC is probably due to expression levels of the duplicated ribosomal protein genes.

Our results did not identify a step during ribosome biogenesis that impacts Ty1 CNC. Although *LOC1* was analyzed in greatest detail as a *loc1Δ* mutation conferred robust CNC-specificity, both small and large subunit ribosomal protein mutations as well as a ribosome assembly factor mutation also affect CNC to varying degrees. Instead of responding to specific steps in ribosome biogenesis, Ty1 CNC may be modulated by ribosome availability. Polysome profiles of several mutants analyzed for Ty1 CNC suggest there is an imbalance of 40S and 60S subunits. A decrease in 40S and an increase in 60S subunits is observed in a *rps0bΔ* mutant [55], a decrease of both 40S and 60S subunits occurs in a *loc1Δ* mutant [64], and a decrease in 60S subunits, 80S monosomes, and polysomes is detected in a *rpl7aΔ* mutant [54]. Defects in functional ribosome production affect cell growth and proliferation [77]. Interestingly, there is also a correlation between ribosome availability, cell growth, and Ty1 transposition. *loc1Δ* and *puf6Δ* mutants grow slowly, contain decreased amounts of 40S and 60S subunits [64], and display a high CNC$^-$/CNC$^+$ ratio (Table 2). Growth of *RPL7* paralog mutants also show a similar relationship with Ty1 CNC. An *rpl7aΔ* mutant grows more slowly, and impacts Ty1 CNC more than a *rpl7bΔ* mutant, which contains more Rpl7a due to the higher expression of *RPL7A* [54]. In addition, the slow growth of ribosomal protein mutants is related to other cellular phenotypes such as enhanced resistance to endoplasmic reticulum stress [78]. Although the magnitude of CNC may correlate with

the level of functional ribosomes, other CNC phenotypes may not be directly related. For example, retrosomes are not detected in *loc1Δ* (Figure 2.8) and *rpl7aΔ* mutants [48], however, a more severe transposition defect is conferred by a *loc1Δ* mutant (Table 2.2). Together, our data support the idea that different steps in the Ty1 life cycle have different thresholds for ribosome levels.

The observed CNC-specific phenotypes in *loc1Δ* and two ribosomal protein mutants likely results from an increased amount of p22 when compared to the isogenic wild type strain (Figure 2.4B). We examined steps in Ty1 gene expression and protein stability to determine what might account for the increase in p22 observed in several ribosome biogenesis mutants. Our data permit the following conclusions:

(1) The level of poly(A)$^+$ Ty1 mRNA and Ty1i RNA is altered in the absence of *LOC1*, *RPS0B,* and *RPL7A*, and displays a trend suggesting that more Ty1i RNA is present relative to Ty1 mRNA (Figure 2.4A). However, the decrease in the level of poly(A)$^+$ Ty1i RNA in the *loc1Δ* and *rpl7aΔ* mutants or the increase in the level of Ty1i RNA observed in other mutants fails to explain the 3.5 to 19.4-fold increase in the level of p22. Perhaps nonpolyadenylated Ty1i transcripts can also be used to synthesize p22 in the ribosome biogenesis mutants [27, 55].

(2) Loc1 is reported to function as a translational repressor [57], therefore, loss of Loc1 may enhance the initiation of p22 translation. However, Ribo-seq analysis of p22 in a *loc1Δ* mutant is comparable to other Ty1 proteins as well as cellular proteins (Figure 2.5C). Although this result raises the possibility that initiation of translation is not enhanced in a *loc1Δ* mutant, it is unclear why the Ribo-seq analysis does not account for the increased level of p22 observed in *loc1Δ* as

well as the *rps0bΔ* and *rpl7aΔ* mutants (Figure 2.4B). Possible reasons for this
discrepancy include: (a) the effects on translation initiation are obscured by the
overlapping nature of RNAs and proteins in this region, (b) the experimental
conditions or strains used for the libraries differ from those used in our work, or
(c) Loc1 may exert its effects through some non-RNA or nontranslational
mechanism.

(3) Loc1 enhances p22 turnover. Pulse-chase immunoprecipitation analysis
suggests that p22 stability increases in a *loc1Δ* mutant (Figure S2.4). However, a
~ 60% increase in p22 stability does not completely explain the increase in p22
level (Figure 2.4B).

What could be causing the transcriptional changes of Ty1 mRNA and Ty1i RNA in
the ribosome biogenesis mutants? Perhaps defects in ribosome biogenesis cause a
form of stress that leads to changes in Ty1 transcript level. A nucleolar stress response
pathway activates p53-dependent as well as p53-independent downstream pathways in
mammalian cells [79]. Nucleolar stress may also occur in yeast [80, 81], though the
signaling pathway has not been defined. We considered the possibility that the
environmental stress response (ESR), which alters expression of a core set of genes
under diverse environmental changes [82] is activated in a *loc1Δ* mutant. However,
analysis of the *loc1Δ* and ESR transcriptomes [83] reveals little overlap with the ESR
(Table S2.4). Interestingly, stress-responsive transcription factors such as Msn2, Tye7,
and Gcn4 affect Ty1 mRNA expression may also be involved in Ty1i RNA transcription.
Msn2 binds to Ty1 elements *in vivo* and is capable of acting as a transcriptional
activator as well as a repressor [84, 85]. Tye7 is a bHLH transcription factor that

activates Ty1-mediated gene expression and regulates Ty1 mRNA and antisense RNAs expression upon adenine starvation [86, 87]. Gcn4 is a transcription factor that responds to amino acid starvation via the TOR pathway and can activate Ty1 transcription by recruiting chromatin remodelers Swi/Snf and SAGA [88, 89]. Further work will be required to understand the relationship between factors responsible for Ty1 mRNA and Ty1i RNA transcription.

Several Ty1 defects in cells lacking *LOC1* or ribosomal proteins can be reconciled by an increase in the level of p22 [27]. These include a marked decrease in mature IN relative to Gag (Figure 2.2B and 2.3), altered Gag-Gag interactions (Figure 2.6 and 2.7), and formation of endogenous (Figure 2.8) or pGTy1-induced retrosomes (Figure S2. 5). Previous work [27] along with data presented here led us to consider retrosome formation in CNC$^-$ and CNC$^+$ cells. In CNC$^-$ cells, retrosomes may not form or are undetectable due to low levels of Ty1 products including p22. The Ty1 copy number dependence of CNC, however, suggests the low level of p22 present in cells containing just a few elements may be less inhibitory [23, 90]. In CNC$^+$ cells, the increase in retrosome abundance may sensitize Ty1 retrotransposition to small changes in p22 levels and disturbances in the retrosome environment or network of interactions.

Recent work on biomolecular condensates [91, 92], also known as membrane-less organelles or quinary structures, have influenced our view of retrosomes and their interactions with other cellular components. We postulate that Ty1 retrosomes are biomolecular condensates as they contain Ty1 RNA and proteins, and likely cellular proteins in an undefined assembly, no fixed membranous structure, and are highly dynamic. Phase separation caused by multivalent macromolecular interactions of

83

nucleic acids and proteins drive the formation of molecular condensates. Indeed, P-bodies and stress granules are biomolecular condensates and required for Ty1 transposition. Retrosomes disassemble in P-body component as well as stress granule component mutants [4, 5, 48]. The highly dynamic nature of retrosomes is also revealed by their disassembly under conditions, such as glucose deprivation, that causes P-bodies to form [5]. Our work favors the idea that normal ribosome biogenesis and functional ribosomes contribute to the retrosome environment required to maintain these condensates in a phase-separated or demixed state. The continuum from retrosome formation to VLP assembly may be similar to the formation of condensates such as stress granules [92] or transcriptional super-enhancers [93] from ribonucleoprotein complexes or "normal" enhancers, respectively. As various chemical and environmental factors or stress conditions influence the structure and function of biomolecular condensates [94], changes in the cellular environment resulting from defects in ribosome biogenesis in combination with increased p22 levels contribute to altered retrosome dynamics and defects in Ty1 mobility.

In summary, our study demonstrates that ribosome biogenesis is critical for Ty1 CNC. The diverse defects in ribosome biogenesis that enhance Ty1 CNC raise the possibility of a cellular response that leads to an increase in the amount of p22 relative to its target Gag. Further work will be required to define a stress response in yeast that affects p22 abundance and retrosome formation, characterize the retrosome interactome both genetically and biochemically, and determine what triggers VLP assembly from retrosomes.

REFERENCES

1.   Curcio, M.J., S. Lutz, and P. Lesage, *The Ty1 LTR-retrotransposon of budding yeast, Saccharomyces cerevisiae.* Microbiol Spectr, 2015. **3**(2): p. 1-35.

2.   Kawakami, K., et al., *A rare tRNA-Arg(CCU) that regulates Ty1 element ribosomal frameshifting is essential for Ty1 retrotransposition in Saccharomyces cerevisiae.* Genetics, 1993. **135**(2): p. 309-20.

3.   Malagon, F. and T.H. Jensen, *The T body, a new cytoplasmic RNA granule in Saccharomyces cerevisiae.* Mol Cell Biol, 2008. **28**(19): p. 6022-32.

4.   Dutko, J.A., et al., *5' to 3' mRNA decay factors colocalize with Ty1 Gag and human APOBEC3G and promote Ty1 retrotransposition.* J Virol, 2010. **84**(10): p. 5052-66.

5.   Checkley, M.A., et al., *P-body components are required for Ty1 retrotransposition during assembly of retrotransposition-competent virus-like particles.* Mol Cell Biol, 2010. **30**(2): p. 382-98.

6.   Feng, Y.X., et al., *The genomic RNA in Ty1 virus-like particles is dimeric.* J Virol, 2000. **74**(22): p. 10819-21.

7.   Chapman, K.B., A.S. Bystrom, and J.D. Boeke, *Initiator methionine tRNA is essential for Ty1 transposition in yeast.* Proc Natl Acad Sci U S A, 1992. **89**(8): p. 3236-40.

8.   Wilhelm, M., et al., *Yeast Ty1 retrotransposon: the minus-strand primer binding site and a cis-acting domain of the Ty1 RNA are both important for packaging of primer tRNA inside virus-like particles.* Nucleic Acids Res, 1994. **22**(22): p. 4560-5.

9.   Moore, S.P., L.A. Rinckel, and D.J. Garfinkel, *A Ty1 integrase nuclear localization signal required for retrotransposition.* Mol Cell Biol, 1998. **18**(2): p. 1105-14.

10.   Kenna, M.A., et al., *Invading the yeast nucleus: a nuclear localization signal at the C terminus of Ty1 integrase is required for transposition in vivo.* Mol Cell Biol, 1998. **18**(2): p. 1115-24.

11.   McLane, L.M., et al., *The Ty1 integrase protein can exploit the classical nuclear protein import machinery for entry into the nucleus.* Nucleic Acids Res, 2008. **36**(13): p. 4317-26.

12. Devine, S.E. and J.D. Boeke, *Integration of the yeast retrotransposon Ty1 is targeted to regions upstream of genes transcribed by RNA polymerase III.* Genes Dev, 1996. **10**(5): p. 620-33.

13. Baller, J.A., et al., *A nucleosomal surface defines an integration hotspot for the Saccharomyces cerevisiae Ty1 retrotransposon.* Genome Res, 2012. **22**(4): p. 704-13.

14. Bridier-Nahmias, A., et al., *Retrotransposons. An RNA polymerase III subunit determines sites of retrotransposon integration.* Science, 2015. **348**(6234): p. 585-8.

15. Cheung, S., et al., *Ty1 integrase interacts with RNA polymerase III-specific subcomplexes to promote insertion of Ty1 elements upstream of polymerase (Pol) III-transcribed genes.* J Biol Chem, 2016. **291**(12): p. 6396-411.

16. Kim, J.M., et al., *Transposable elements and genome organization: a comprehensive survey of retrotransposons revealed by the complete Saccharomyces cerevisiae genome sequence.* Genome Res, 1998. **8**(5): p. 464-78.

17. Lander, E.S., et al., *Initial sequencing and analysis of the human genome.* Nature, 2001. **409**(6822): p. 860-921.

18. Curcio, M.J. and D.J. Garfinkel, *Single-step selection for Ty1 element retrotransposition.* Proc Natl Acad Sci U S A, 1991. **88**(3): p. 936-40.

19. Scheifele, L.Z., et al., *Retrotransposon overdose and genome integrity.* Proc Natl Acad Sci U S A, 2009. **106**(33): p. 13927-32.

20. Drinnenberg, I.A., et al., *RNAi in budding yeast.* Science, 2009. **326**(5952): p. 544-50.

21. Drinnenberg, I.A., G.R. Fink, and D.P. Bartel, *Compatibility with killer explains the rise of RNAi-deficient fungi.* Science, 2011. **333**(6049): p. 1592.

22. Levin, H.L. and J.V. Moran, *Dynamic interactions between transposable elements and their hosts.* Nat Rev Genet, 2011. **12**(9): p. 615-27.

23. Garfinkel, D.J., et al., *Post-transcriptional cosuppression of Ty1 retrotransposition.* Genetics, 2003. **165**(1): p. 83-99.

24. Moore, S.P., et al., *Analysis of a Ty1-less variant of Saccharomyces paradoxus: the gain and loss of Ty1 elements.* Yeast, 2004. **21**(8): p. 649-660.

25. Matsuda, E. and D.J. Garfinkel, *Posttranslational interference of Ty1 retrotransposition by antisense RNAs.* Proc Natl Acad Sci U S A, 2009. **106**(37): p. 15657-62.

26. Purzycka, K.J., et al., *Exploring Ty1 retrotransposon RNA structure within virus-like particles.* Nucleic Acids Res, 2013. **41**(1): p. 463-73.

27. Saha, A., et al., *A trans-dominant form of Gag restricts Ty1 retrotransposition and mediates copy number control.* J Virol, 2015. **89**(7): p. 3922-38.

28. Nishida, Y., et al., *Ty1 retrovirus-like element Gag contains overlapping restriction factor and nucleic acid chaperone functions.* Nucleic Acids Res, 2015. **43**(15): p. 7414-31.

29. Tucker, J.M., et al., *The Ty1 retrotransposon restriction factor p22 targets Gag.* PLoS Genet, 2015. **11**(10): p. e1005571.

30.   Maxwell, P.H. and M.J. Curcio, *Host factors that control long terminal repeat retrotransposons in Saccharomyces cerevisiae: implications for regulation of mammalian retroviruses.* Eukaryot Cell, 2007. **6**(7): p. 1069-80.

31.   Goff, S.P., *Knockdown screens to knockout HIV-1.* Cell, 2008. **135**(3): p. 417-20.

32.   Nyswaner, K.M., et al., *Chromatin-associated genes protect the yeast genome from Ty1 insertional mutagenesis.* Genetics, 2008. **178**(1): p. 197-214.

33.   Griffith, J.L., et al., *Functional genomics reveals relationships between the retrovirus-like Ty1 element and its host Saccharomyces cerevisiae.* Genetics, 2003. **164**(3): p. 867-79.

34.   Dakshinamurthy, A., et al., *BUD22 affects Ty1 retrotransposition and ribosome biogenesis in Saccharomyces cerevisiae.* Genetics, 2010. **185**(4): p. 1193-205.

35.   Risler, J.K., et al., *Host co-factors of the retrovirus-like transposon Ty1.* Mob DNA, 2012. **3**(1): p. 12.

36.   Scholes, D.T., et al., *Multiple regulators of Ty1 transposition in Saccharomyces cerevisiae have conserved roles in genome maintenance.* Genetics, 2001. **159**(4): p. 1449-65.

37.   Heym, R.G. and D. Niessing, *Principles of mRNA transport in yeast.* Cell Mol Life Sci, 2012. **69**(11): p. 1843-53.

38.   Urbinati, C.R., et al., *Loc1p is required for efficient assembly and nuclear export of the 60S ribosomal subunit.* Mol Genet Genomics, 2006. **276**(4): p. 369-77.

39.   Guthrie, C. and G.R. Fink, *Guide to yeast genetics and molecular biology.* Methods Enzymol., 1991. **194**: p. 1-863.

40.   Boeke, J.D., F. LaCroute, and G.R. Fink, *A positive selection for mutants lacking orotidine-5'-phosphate decarboxylase activity in yeast: 5-fluoro-orotic acid resistance.* Mol Gen Genet, 1984. **197**(2): p. 345-6.

41.   Boeke, J.D., et al., *The Saccharomyces cerevisiae genome contains functional and nonfunctional copies of transposon Ty1.* Mol Cell Biol, 1988. **8**(4): p. 1432-42.

42.   Winston, F., et al., *Mutations affecting Ty-mediated expression of the HIS4 gene of Saccharomyces cerevisiae.* Genetics, 1984. **107**(2): p. 179-97.

43.   Brachmann, C.B., et al., *Designer deletion strains derived from Saccharomyces cerevisiae S288c: a useful set of strains and plasmids for PCR-mediated gene disruption and other applications.* Yeast, 1998. **14**(2): p. 115-32.

44.   Giaever, G., et al., *Functional profiling of the Saccharomyces cerevisiae genome.* Nature, 2002. **418**(6896): p. 387-91.

45.   Wach, A., et al., *New heterologous modules for classical or PCR-based gene disruptions in Saccharomyces cerevisiae.* Yeast, 1994. **10**(13): p. 1793-808.

46.   Braiterman, L.T., et al., *In-frame linker insertion mutagenesis of yeast transposon Ty1: phenotypic analysis.* Gene, 1994. **139**(1): p. 19-26.

47.   Bastin, P., et al., *A novel epitope tag system to study protein targeting and organelle biogenesis in Trypanosoma brucei.* Mol Biochem Parasitol, 1996. **77**(2): p. 235-9.

48.   Doh, J.H., S. Lutz, and M.J. Curcio, *Co-translational localization of an LTR-retrotransposon RNA to the endoplasmic reticulum nucleates virus-like particle assembly sites.* PLoS Genet, 2014. **10**(3): p. e1004219.

49.   Belle, A., et al., *Quantification of protein half-lives in the budding yeast proteome.* Proc Natl Acad Sci U S A, 2006. **103**(35): p. 13004-9.

50.   Checkley, M.A., et al., *Ty1 Gag enhances the stability and nuclear export of Ty1 mRNA.* Traffic, 2013. **14**(1): p. 57-69.

51.   Braiterman, L.T. and J.D. Boeke, *In vitro integration of retrotransposon Ty1: a direct physical assay.* Mol Cell Biol, 1994. **14**(9): p. 5719-30.

52.   Huang da, W., B.T. Sherman, and R.A. Lempicki, *Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources.* Nat Protoc, 2009. **4**(1): p. 44-57.

53.   Long, R.M., et al., *An exclusively nuclear RNA-binding protein affects asymmetric localization of ASH1 mRNA and Ash1p in yeast.* J Cell Biol, 2001. **153**(2): p. 307-18.

54.   Palumbo, R.J., et al., *Paralog-specific functions of RPL7A and RPL7B mediated by ribosomal protein or snoRNA dosage in Saccharomyces cerevisiae.* G3 (Bethesda), 2017. **7**(2): p. 591-606.

55.   Suresh, S., et al., *Ribosomal protein and biogenesis factors affect multiple steps during movement of the Saccharomyces cerevisiae Ty1 retrotransposon.* Mob DNA, 2015. **6**: p. 22.

56.   Curcio, M.J. and D.J. Garfinkel, *Posttranslational control of Ty1 retrotransposition occurs at the level of protein processing.* Mol Cell Biol, 1992. **12**(6): p. 2813-25.

57.   Komili, S., et al., *Functional specificity among ribosomal proteins regulates gene expression.* Cell, 2007. **131**(3): p. 557-71.

58. Kutluay, S.B. and P.D. Bieniasz, *Analysis of the initiating events in HIV-1 particle assembly and genome packaging.* PLoS Pathog, 2010. **6**(11): p. e1001200.

59. Zhu, M. and S. Zhao, *Candidate gene identification approach: progress and challenges.* Int J Biol Sci, 2007. **3**(7): p. 420-7.

60. Harnpicharnchai, P., et al., *Composition and functional characterization of yeast 66S ribosome assembly intermediates.* Mol Cell, 2001. **8**(3): p. 505-15.

61. Niedner, A., et al., *Role of Loc1p in assembly and reorganization of nuclear ASH1 messenger ribonucleoprotein particles in yeast.* Proc Natl Acad Sci U S A, 2013. **110**(52): p. E5049-58.

62. Li, Z., et al., *Rational extension of the ribosome biogenesis pathway using network-guided genetics.* PLoS Biol, 2009. **7**(10): p. e1000213.

63. Shen, Z., et al., *Nuclear shuttling of She2p couples ASH1 mRNA localization to its translational repression by recruiting Loc1p and Puf6p.* Mol Biol Cell, 2009. **20**(8): p. 2265-75.

64. Yang, Y.T., et al., *The Roles of Puf6 and Loc1 in 60S biogenesis are interdependent, and both are required for efficient accommodation of Rpl43.* J Biol Chem, 2016. **291**(37): p. 19312-23.

65. Boisvert, F.M., et al., *The multifunctional nucleolus.* Nat Rev Mol Cell Biol, 2007. **8**(7): p. 574-85.

66. Boulon, S., et al., *The nucleolus under stress.* Mol Cell, 2010. **40**(2): p. 216-27.

67. Hiscox, J.A., *RNA viruses: hijacking the dynamic nucleolus.* Nat Rev Microbiol, 2007. **5**(2): p. 119-27.

68.   Salvetti, A. and A. Greco, *Viruses and the nucleolus: The fatal attraction.* Biochim Biophys Acta, 2013.

69.   Lochmann, T.L., et al., *NC-mediated nucleolar localization of retroviral gag proteins.* Virus Res, 2013. **171**(2): p. 304-18.

70.   Beyer, A.R., et al., *Nucleolar trafficking of the mouse mammary tumor virus gag protein induced by interaction with ribosomal protein L9.* J Virol, 2013. **87**(2): p. 1069-82.

71.   Jarboui, M.A., et al., *Nucleolar protein trafficking in response to HIV-1 Tat: rewiring the nucleolus.* PLoS One, 2012. **7**(11): p. e48702.

72.   Lin, S.S., et al., *Integrase mediates nuclear localization of Ty3.* Mol Cell Biol, 2001. **21**(22): p. 7826-38.

73.   Goodier, J.L., et al., *A potential role for the nucleolus in L1 retrotransposition.* Hum Mol Genet, 2004. **13**(10): p. 1041-8.

74.   Caudron-Herger, M., et al., *Alu element-containing RNAs maintain nucleolar structure and function.* EMBO J, 2015. **34**(22): p. 2758-74.

75.   Mekdad, H.E., et al., *Characterization of the interaction between the HIV-1 Gag structural polyprotein and the cellular ribosomal protein L7 and its implication in viral nucleic acid remodeling.* Retrovirology, 2016. **13**(1): p. 54.

76.   Lee, A.S., R. Burdeinick-Kerr, and S.P. Whelan, *A ribosome-specialized translation initiation pathway is required for cap-dependent translation of vesicular stomatitis virus mRNAs.* Proc Natl Acad Sci U S A, 2013. **110**(1): p. 324-9.

77. Woolford, J.L. and S.J. Baserga, *Ribosome biogenesis in the yeast Saccharomyces cerevisiae.* Genetics, 2013. **195**(3): p. 643-681.

78. Steffen, K.K., et al., *Ribosome deficiency protects against ER stress in Saccharomyces cerevisiae.* Genetics, 2012. **191**(1): p. 107-118.

79. James, A., et al., *Nucleolar stress with and without p53.* Nucleus, 2014. **5**(5): p. 402-26.

80. Gomez-Herreros, F., et al., *Balanced production of ribosome components is required for proper G1/S transition in Saccharomyces cerevisiae.* J Biol Chem, 2013. **288**(44): p. 31689-700.

81. Thapa, M., et al., *Repressed synthesis of ribosomal proteins generates protein-specific cell cycle and morphological phenotypes.* Mol Biol Cell, 2013. **24**(23): p. 3620-3633.

82. Gasch, A.P. and M. Werner-Washburne, *The genomics of yeast responses to environmental stress and starvation.* Funct Integr Genomics, 2002. **2**(4-5): p. 181-92.

83. Gasch, A.P., et al., *Genomic expression programs in the response of yeast cells to environmental changes.* Mol Biol Cell, 2000. **11**(12): p. 4241-57.

84. Elfving, N., et al., *A dynamic interplay of nucleosome and Msn2 binding regulates kinetics of gene activation and repression following stress.* Nucleic Acids Res, 2014. **42**(9): p. 5468-82.

85. Schmitt, A.P. and K. McEntee, *Msn2p, a zinc finger DNA-binding protein, is the transcriptional activator of the multistress response in Saccharomyces cerevisiae.* Proc Natl Acad Sci U S A, 1996. **93**(12): p. 5777-82.

86.   Lohning, C. and M. Ciriacy, *The TYE7 gene of Saccharomyces cerevisiae encodes a putative bHLH-LZ transcription factor required for Ty1-mediated gene expression.* Yeast, 1994. **10**(10): p. 1329-39.

87.   Servant, G., et al., *Tye7 regulates yeast Ty1 retrotransposon sense and antisense transcription in response to adenylic nucleotides stress.* Nucleic Acids Res, 2012. **40**(12): p. 5271-82.

88.   Hinnebusch, A.G. and K. Natarajan, *Gcn4p, a master regulator of gene expression, is controlled at multiple levels by diverse signals of starvation and stress.* Eukaryot Cell, 2002. **1**(1): p. 22-32.

89.   Morillon, A., et al., *Differential effects of chromatin and Gcn4 on the 50-fold range of expression among individual yeast Ty1 retrotransposons.* Mol Cell Biol, 2002. **22**(7): p. 2078-88.

90.   Garfinkel, D.J., et al., *A self-encoded capsid derivative restricts Ty1 retrotransposition in Saccharomyces.* Curr Genet, 2016. **62**(2): p. 321-9.

91.   Banani, S.F., et al., *Biomolecular condensates: organizers of cellular biochemistry.* Nat Rev Mol Cell Biol, 2017. **18**(5): p. 285-298.

92.   Riback, J.A., et al., *Stress-triggered phase separation is an adaptive, evolutionarily tuned response.* Cell, 2017. **168**(6): p. 1028-1040 e19.

93.   Hnisz, D., et al., *A phase separation model for transcriptional control.* Cell, 2017. **169**(1): p. 13-23.

94.   Mitrea, D.M. and R.W. Kriwacki, *Phase separation in biology; functional organization of a higher order.* Cell Commun Signal, 2016. **14**: p. 1.

95.   Sharon, G., T.J. Burkett, and D.J. Garfinkel, *Efficient homologous recombination of Ty1 element cDNA when integration is blocked.* Mol Cell Biol, 1994. **14**(10): p. 6540-51.
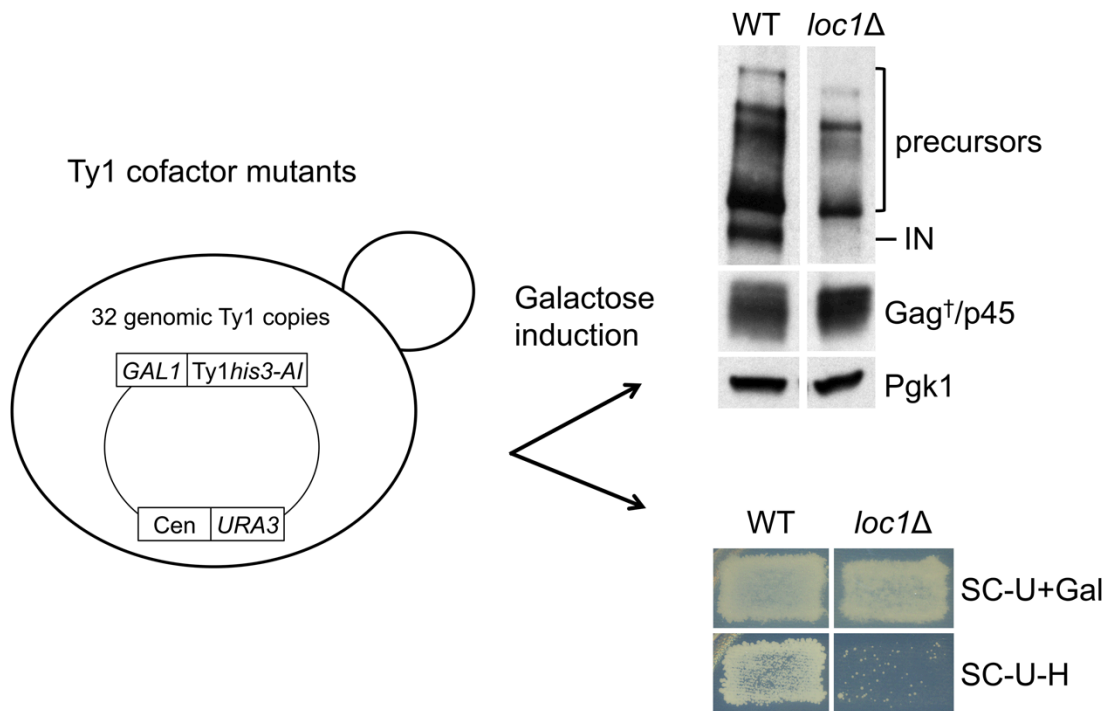
Figure 2.1 Identifying Ty1 cofactor mutants that phenocopy CNC.

Selected Ty1 cofactor mutants containing pGTy1*his3-AI* (pBDG606) the BY4742

background were induced on SC-Uracil (U) + Galactose (Gal) solid and liquid media

(Figure S2.1A). Results from *loc1Δ* are shown as an example. The level of mature IN

and Gag isoforms was assessed by western blot analysis using IN and VLP antisera. IN

and Pol precursors, mature Gag-p45 and Gag isoforms (†) that include Gag-p49 are

noted alongside the blot. Pgk1 (3-phosphoglycerate kinase) served as a loading control.

Cell patches from SC-U+Galactose were replicated to SC-U-Histidine (H) to detect

Ty1*HIS3* mobility events, as monitored by the formation of His$^+$ papillae. Mutants

showing a severe decrease in mature IN but not Gag and a decrease in Ty1 mobility

were selected for further analysis.
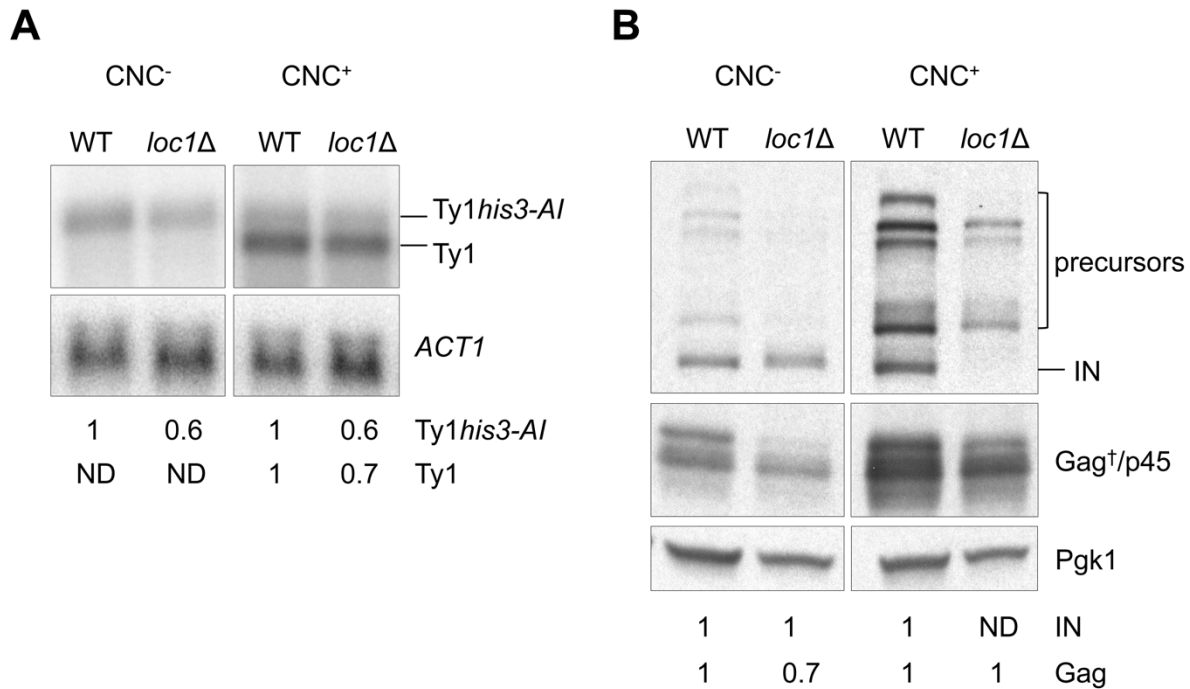
Figure 2.2 Ty1 RNA and protein levels in congenic CNC⁻ and CNC⁺ WT and *loc1*Δ strains expressing pBDG606. (A) Northern blot analysis of total RNA hybridized with a ³²P-labeled riboprobe containing Ty1 nt 238-1702 to detect Ty1*his3-AI* and chromosomal Ty1 transcripts. *ACT1* served as a loading control. Below are relative changes in RNA level compared to WT as determined by phosphorimage analysis. (B) Western blot analysis of Ty1 IN and Gag detected in total cell extracts (40 µg/lane) of CNC⁻ and CNC⁺ *loc1*Δ mutant strains. Refer to the Figure 2.1 legend for a description of Ty1 proteins. Pgk1 served as a loading control. Numbers below each panel indicate relative changes in protein level compared to WT as determined by quantification of protein bands by densitometry. ND: not detected.
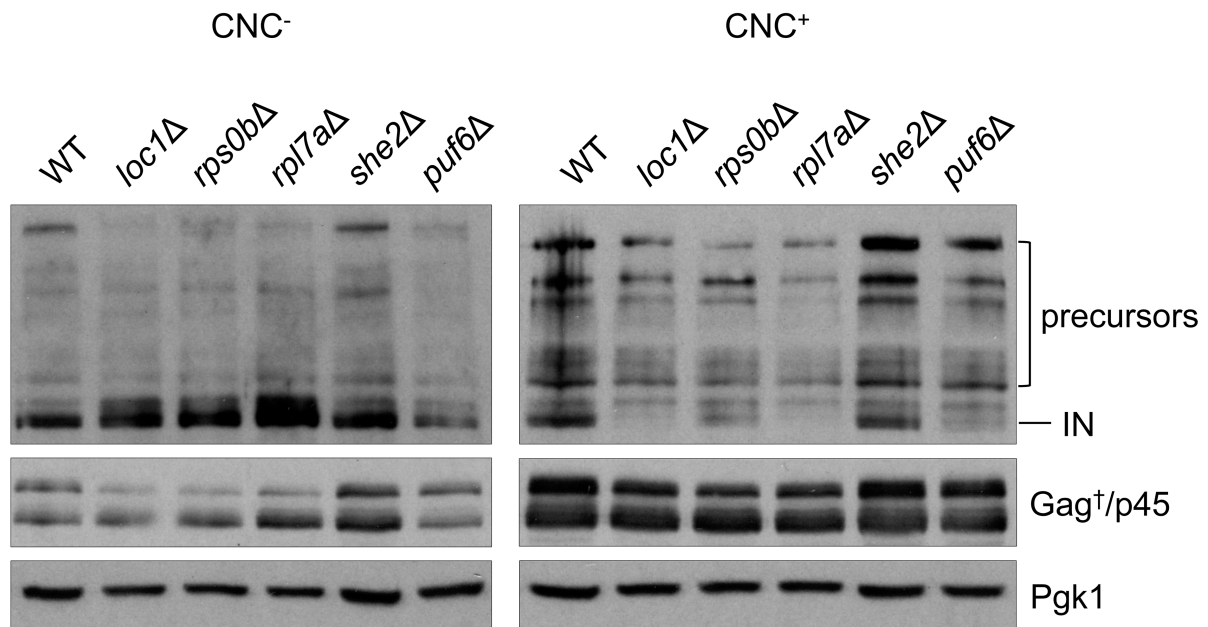
Figure 2.3 Ty1 IN and Gag levels in ribosome biogenesis and RNA localization mutants.

Total cell extracts from CNC⁻ and CNC⁺ wild type and *loc1Δ*, *rps0bΔ*, *rpl7aΔ*, *she2Δ*, and *puf6Δ* mutant strains expressing pBDG606 were subjected to western blot analysis using IN and VLP antisera. Pgk1 served as a loading control. Refer to Figure 2.1 for a description of Ty1 proteins.
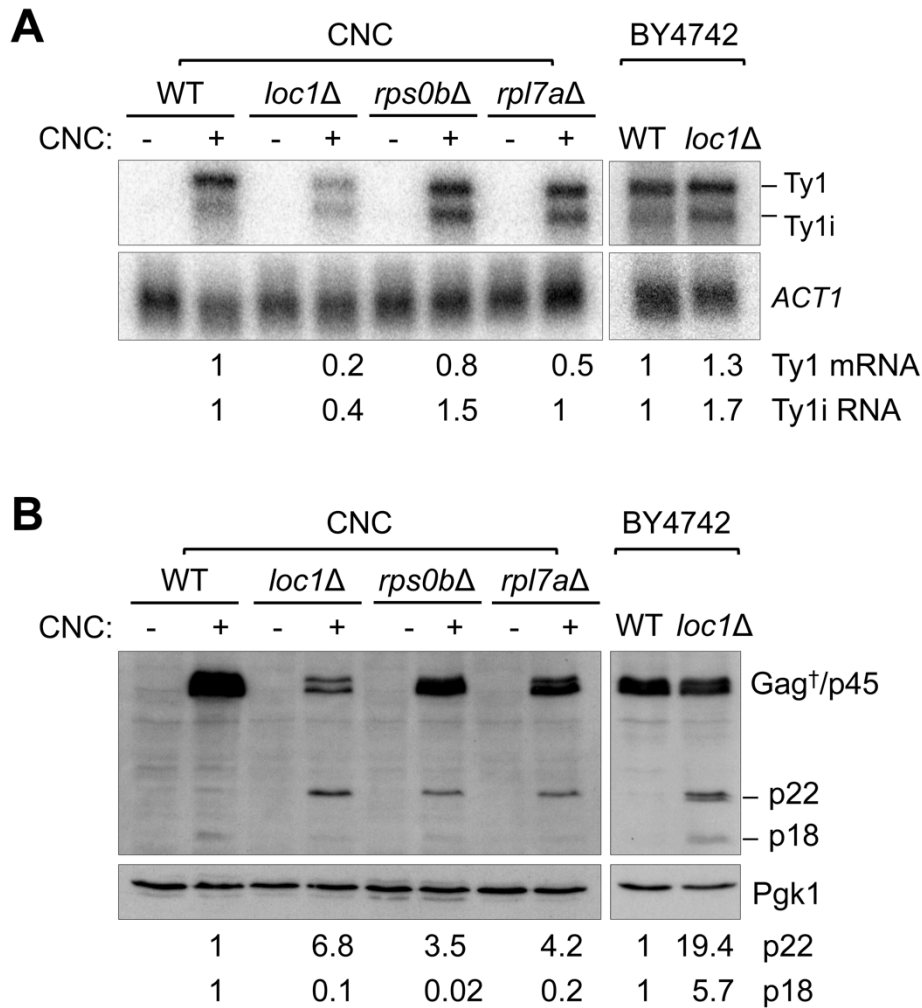
Figure 2.4 Endogenous Ty1 transcript and p22 levels from CNC WT and ribosome biogenesis mutants, and BY4742 wild type and a *loc1Δ* mutant. (A) Northern blot analysis of poly(A)$^+$ RNA hybridized with ^{32}P-labeled riboprobe containing Ty1 nt 1266-1601. This probe detects Ty1 mRNA, and Ty1 internally-initiated (i) RNA that encodes p22/p18. *ACT1* served as a loading control. The relative amount of Ty1 transcripts is shown below. (B) Western blot analysis of Ty1 and p22/p18 from total cell extracts using p18 antiserum, which detects Gag isoforms (†), p45, p22, and p18. Pgk1 served as a loading control. The relative amount of p22 and p18 is shown below.
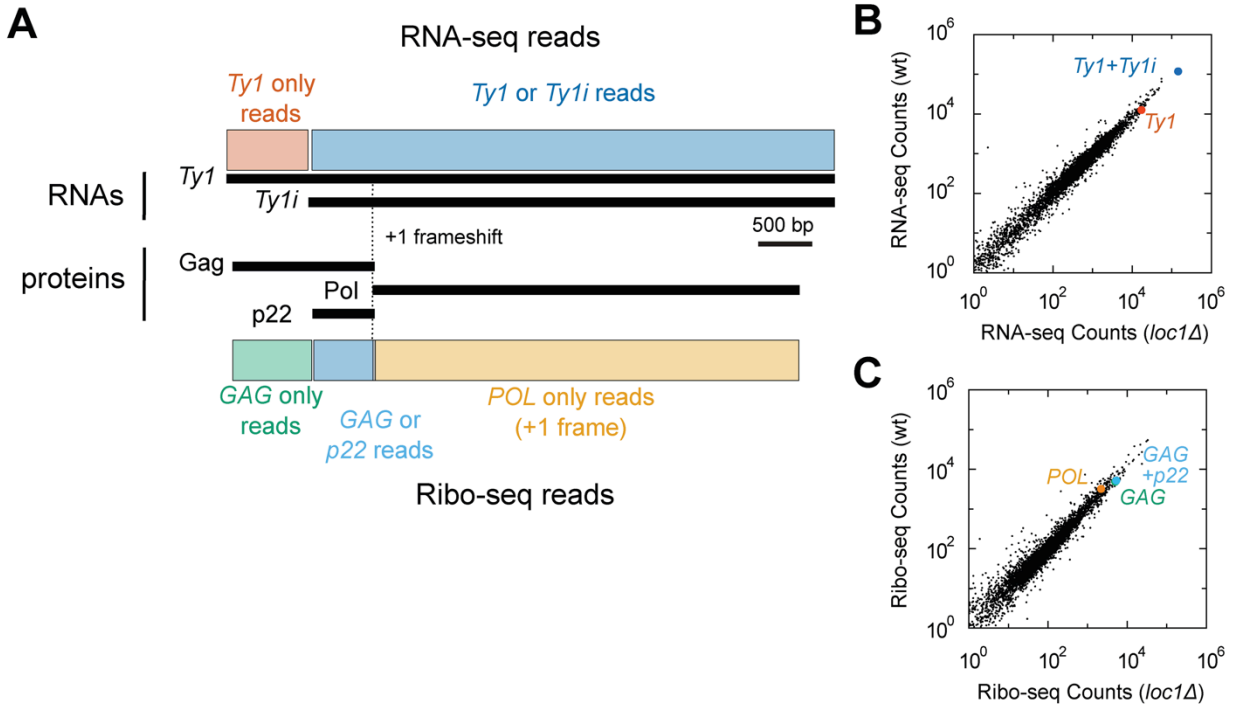
Figure 2.5 Genomic analysis of RNA level and ribosome initiation in wild type and a

loc1Δ mutant. (A) Transcripts and protein produced from Ty1. The element was divided

into regions uniquely assignable to Ty1 (red), or Ty1 and Ty1i transcripts (blue). Three

proteins are presented, two overlapping in the zero frame (Gag and p22), and a third

produced after the +1 frameshift (Pol). These were assigned to Gag (green), Gag or p22

(light blue), or Pol (orange). (B) RNA-seq read counts for all genes (black dots) between

wild type and _loc1Δ_. Each point is the average of three biological replicas for each

strain. Highlighted are RNA read counts uniquely assignable to Ty1 (red), or Ty1 and

Ty1i (blue). (C) Ribo-seq read counts for all genes (black dots) between wild type and

loc1Δ. Analysis was restricted to reads 28 bp in length, and each point is the average of

three biological replicas for each strain. Highlighted are read counts uniquely assignable

to _GAG_ in the zero frame (green), to either _GAG_ or p22 in the zero frame (light blue),

and reads from the _POL_ region in the +1 frame (orange).
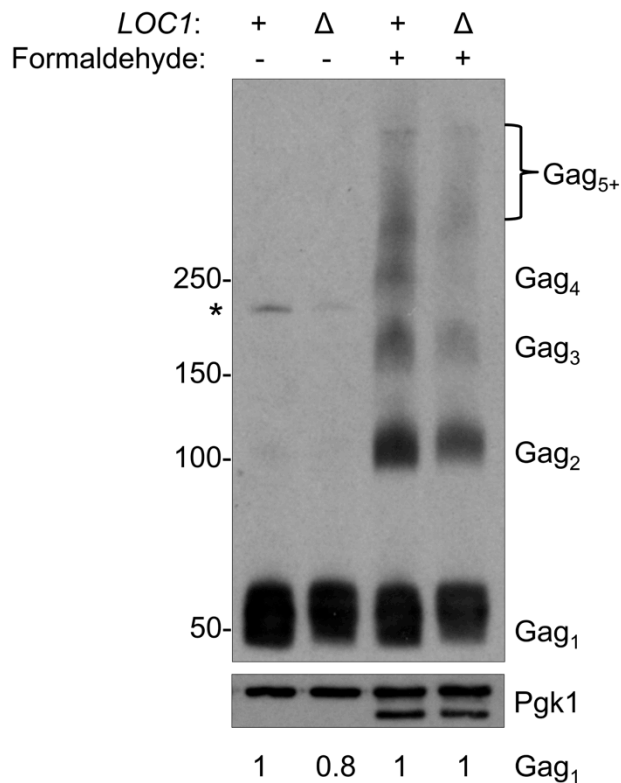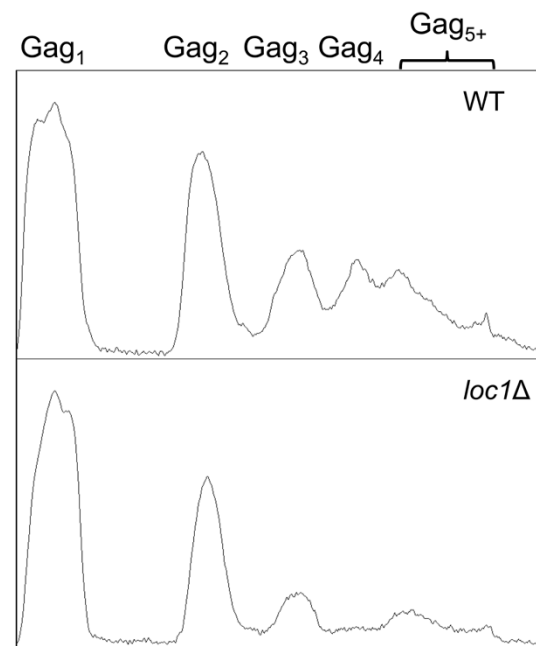
Figure 2.6 Gag multimerization in wild type and *loc1*Δ strains. (A) Total protein from pBDG606-induced WT and *loc1*Δ strains were cross-linked by formaldehyde treatment *in vivo*. Protein extracts from untreated (- formaldehyde) and treated (+ formaldehyde) cells were separated on a 3-8% Tris-Acetate gel and probed with Gag antiserum (TY-tag). Samples (7.5 µg/lane) separated on a 10% SDS-PAGE gel were probed with Pgk1 antiserum, which served as a loading control. Estimated protein sizes are shown in kilodaltons. Extracts from untreated cells contain monomeric Gag (Gag_1) and the Gag-Pol-p199 precursor (*), while cross-linked cells produce a ladder of Gag multimers. The gel conditions allow sufficient separation of Gag_{1-4}, while larger multimers (Gag_{5+}) are not resolved. An additional faster-migrating Pgk1 signal appeared in formaldehyde-

treated samples, which likely results from cross-linking. Numbers below the panel

indicate relative changes in Gag_1 level compared to wild type with and without

formaldehyde treatment, as determined by quantification of protein bands by

densitometry. Both Pgk1 bands in formaldehyde treated cells were used in the

quantification. (B) Densitometry traces of cross-linked Gag from WT and *loc1Δ*. The

pixel intensity for each Gag-multimer per lane is shown along the Y-axis.
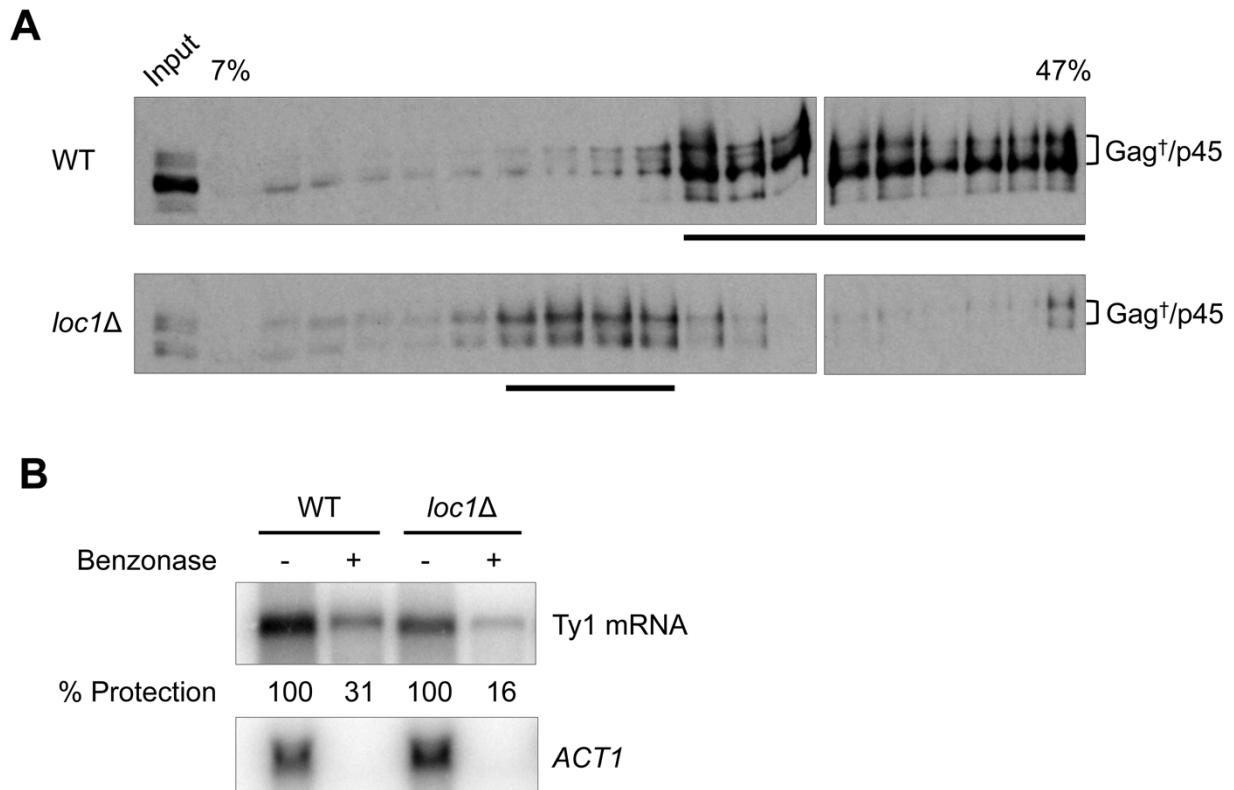
Figure 2.7 Gag sedimentation and nuclease protection of Ty1 mRNA in wild type and

loc1Δ strains. (A) Total cell extracts from BY4742 WT or a *loc1Δ* mutant were

centrifuged through a 7–47% continuous sucrose gradient. Ten micrograms of protein

extract (input) and equal volumes of each fraction (15 ul/lane) were subjected to western blot analysis using Gag antiserum (TY-tag), which recognizes Gag isoforms and Gag-p45. Fractions containing the highest concentrations of Gag, as determined by densitometry tracing, are underlined. Sucrose gradient analysis was repeated 3 times and a representative experiment is shown here. (B) Ty1 mRNA packaging as monitored by sensitivity to the nuclease benzonase. Equal aliquots of whole cell extract from BY4742 WT and a *loc1*Δ mutant were incubated with (+) or without (−) benzonase. To detect Ty1 mRNA, total RNA extracted from these samples was analyzed by northern blot analysis. Nuclease protection assays were repeated 3 times and a representative experiment is shown. WT protection ranged from 26.5-31.4% and *loc1*Δ protection ranged from 16.1-18.8%. *ACT1* was used as a control to confirm RNA degradation in the benzonase treated samples.

	DIC	DAPI	Ty1 mRNA	Gag	MERGE
WT					34.2%
loc1Δ					0%
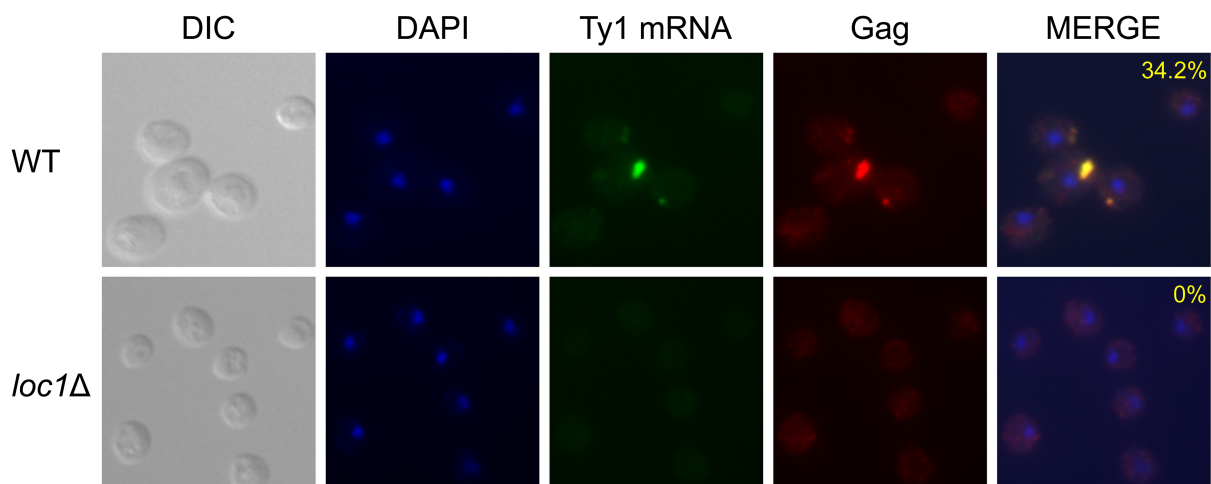
Figure 2.8 Retrosome formation in wild type and *loc1Δ* cells. BY4742 WT and *loc1Δ*

mutant strains were visualized by differential interference (DIC) and fluorescence

microscopy. DNA was stained with DAPI, Ty1 mRNA was detected using a cocktail of

GAG-DIG probes, and Gag was detected with VLP antiserum. The percentages of cells

containing retrosomes, defined as having at least one colocalized Ty1 mRNA/Gag foci,

is shown in the DAPI/Ty1 mRNA/Gag merge. The experiment was repeated twice and

>200 cells were analyzed for each strain.
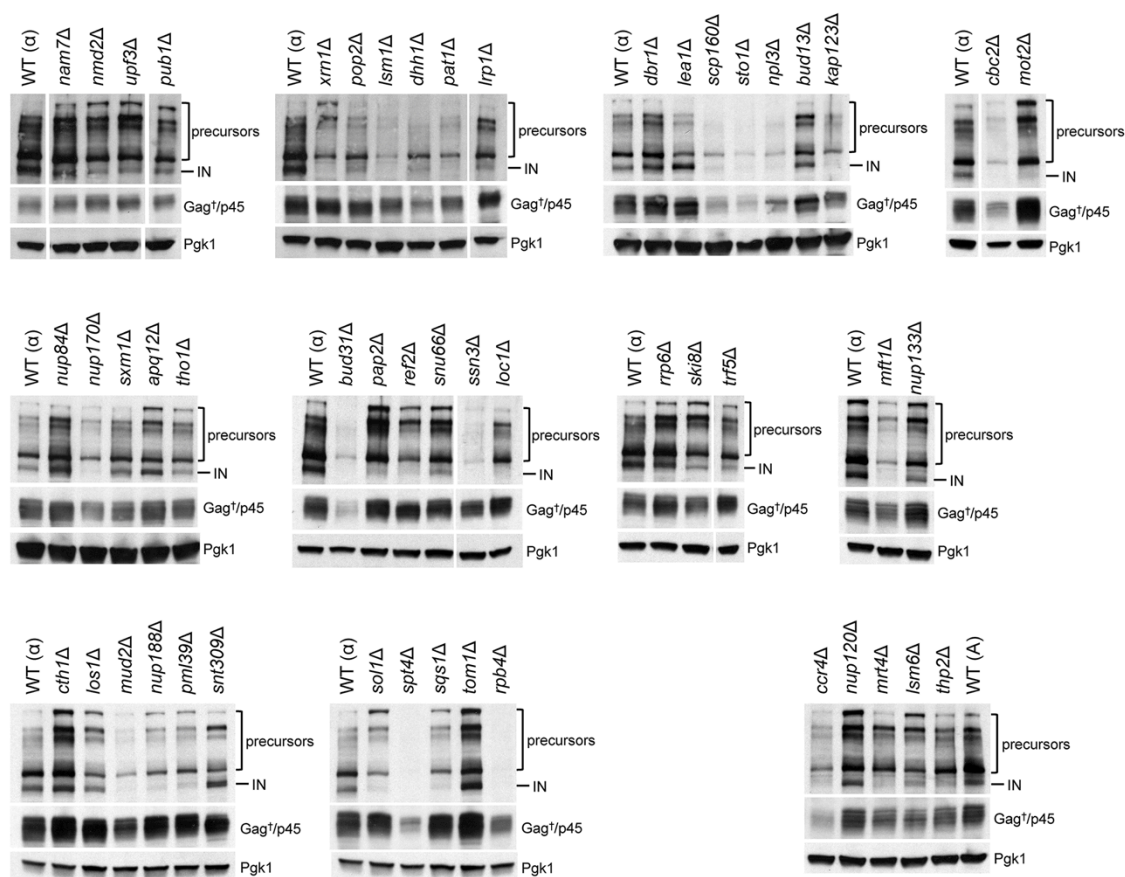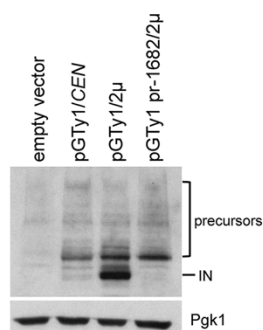
Figure S2.1 Ty1 IN and Gag levels in cofactor mutants, and in cells expressing

multicopy (*2μ*-based) versus low-copy (*CEN*-based) wild type and PR-defective pGTy1.

(A) Level of Ty1 IN and Gag in 51 Ty1 cofactor mutants involved in post-transcriptional

RNA biogenesis. Total protein extracted from wild type and mutant strains expressing a low-copy pGTy1*his3-AI*/Cen/*URA3* plasmid pBDG606 was subjected to western blot analysis (40 μg/lane) using IN and VLP antisera. Forty-six mutations were in the BY4742 (*MAT*α) strain background and five were in the closely related BY4741 (*MAT*a) strain. Pgk1 served as a loading control. (B) Differential expression of pGTy1-induced IN based on the copy number of the plasmid. BY4742 expressing either empty vector, pGTy1/Cen/*URA3* (pBDG606), pGTy1/2μ/*URA3* (pBDG1003, GS99) pGTy1/2μ/*URA3* carrying PR mutation *pr-1682* (pBDG1615, GS99 2-1) [95]. Total protein was subjected to western blot analysis using IN antiserum. Pgk1 served as a loading control.
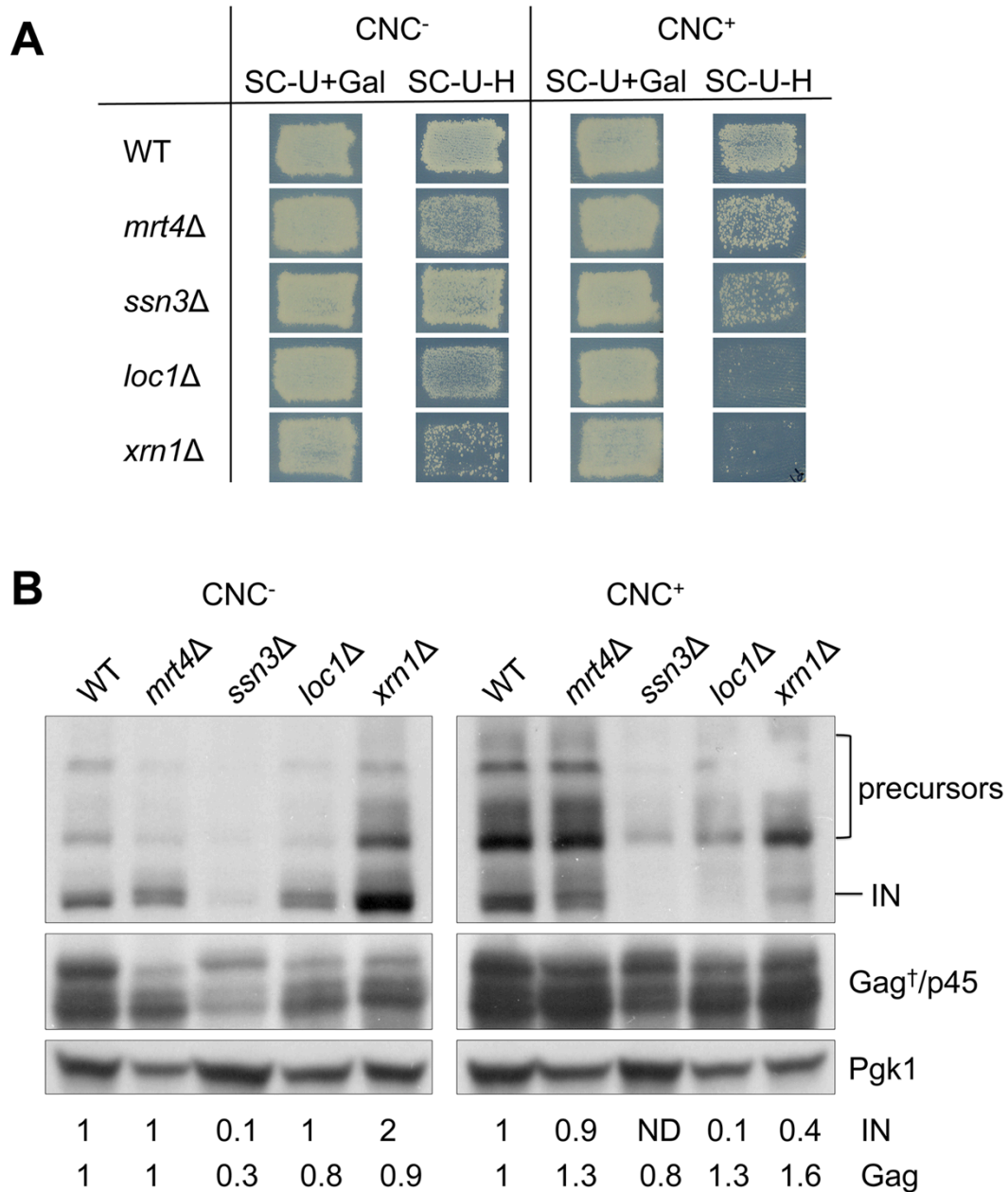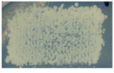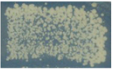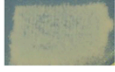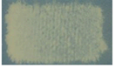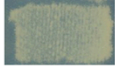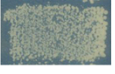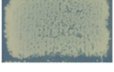
Figure S2.2 Qualitative Ty1*his3-AI* mobility and Ty1 protein level in CNC⁻ and CNC⁺ backgrounds containing cofactor mutations *ssn3Δ, mrt4Δ, loc1Δ,* and *xrn1Δ.* (A) Qualitative Ty1*his3-AI* mobility displayed by the cofactor mutants. Wild type and mutant strains containing pBDG606 were induced with galactose for 1 day at 22 °C. Ty1*HIS3* transposition events were detected as His⁺ papillae by replica plating the SC-
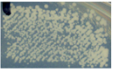
Ura+Galactose plates to SC-Ura-His, followed by incubation for 2 days at 30 °C. (B) IN and Gag level of the cofactor mutants. Total protein from strains expressing pBDG606 was subjected to western blot analysis (40 µg/lane) using IN and VLP antisera. Gag† isoforms and Gag-p45, Pol precursors, and mature IN are highlighted. Pgk1 served as a loading control. Numbers below each panel indicate relative changes in protein level compared to WT as determined by quantification of protein bands by densitometry. ND: not detected.
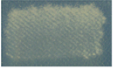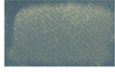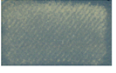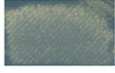
Figure S2.3 Qualitative Ty1*his3-AI* mobility in CNC⁻ and CNC⁺ strains containing mutations of genes involved in the localization of *ASH1* mRNA. Wild type and *loc1Δ*, *she2Δ*, *she3Δ*, and *scp160Δ* mutant strains containing pBDG606 were induced with galactose for one day at 22 °C. Ty1*HIS3* mobility events were detected as His⁺ papillae by replica plating the SC-Ura+Galactose plates to SC-Ura-His, followed by incubation for 2 days at 30 °C.
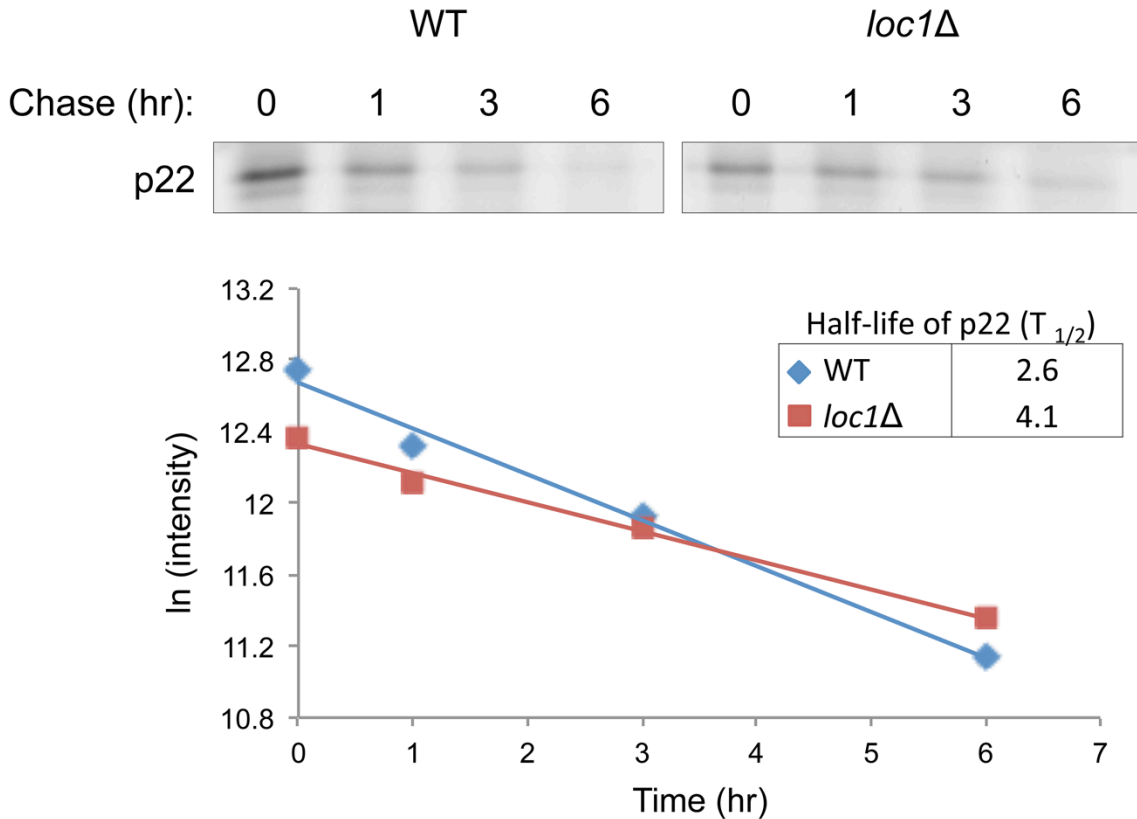
Figure S2.4 Stability of p22 in wild type and *loc1Δ* cells. p22 from cells were pulse-labeled with HPG for 30 min and chased with methionine for 0, 1, 3, or 6 hrs, as indicated. HPG-labeled p22 was immunoprecipitated with p18 antiserum and detected by conjugation to TAMRA. TAMRA-conjugated p22 was analyzed on a 15% SDS-PAGE gel and the fluorescent signal was visualized and quantified using a Typhoon scanner. The fluorescent signal in each strain relative to the 0 hr chase time point was plotted. This experiment is repeated twice and the top image is from a representative experiment. Signals from two experiments were averaged and the ln of the intensity was plotted versus time (hr). The half-life of p22 was determined as described previously [49]. The R^2 value for each trendline was 0.98775 for wild type and 0.9922 for the *loc1Δ* mutant.
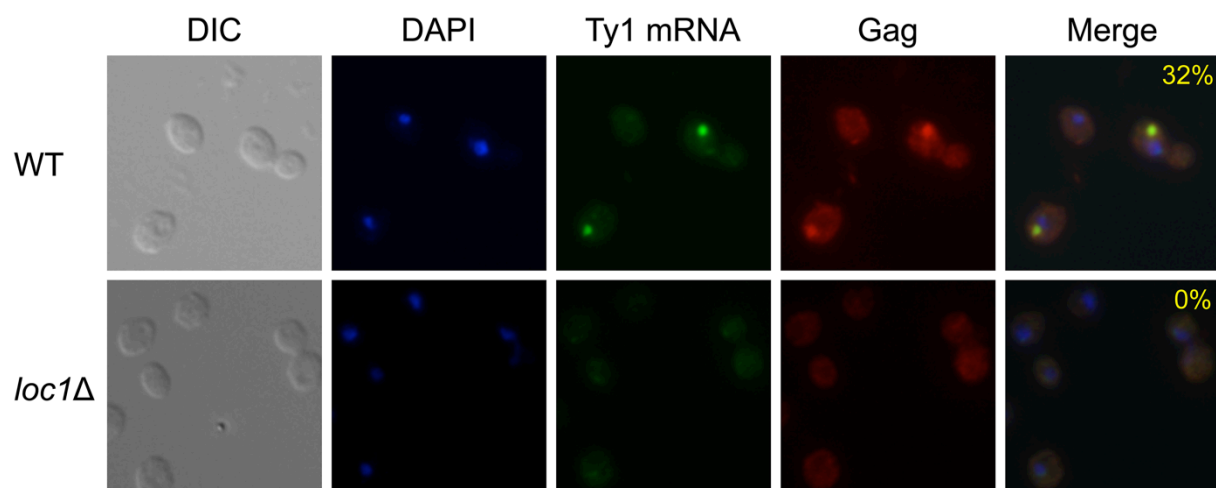
Figure S2.5 Retrosome formation in wild type and *loc1*Δ cells expressing pBDG606.

Cells grown at 20 °C were visualized by differential interference (DIC) and fluorescence microscopy. DNA was stained with DAPI, Ty1 mRNA was detected using a cocktail of GAG-DIG probes, and Gag was detected with VLP antiserum. The percentages of cells containing retrosomes, defined as having at least one colocalized Ty1 RNA/Gag foci, is shown in the DAPI/Ty1 mRNA/Gag merge. The experiment was repeated twice and >200 cells were analyzed for each strain.

Table 1. Ty1 cofactors with roles in post-transcriptional RNA biogenesis

Classification	Gene name
P-body components	*DHH1, LSM1, LSM6, PAT1, PUB1, XRN1*
Nonsense-mediated decay	*NAM7, NMD2, UPF3*
Exosome components	*LRP1, RRP6, SKI8*
snRNA/snoRNA processing	*BUD31, REF2, SNU66*
RNA splicing	*DBR1, LEA1, MUD2, SNT309, SQS1*
CCR4-NOT complex	*CCR4, MOT2, POP2*
TRAMP complex	*PAP2, TRF5*
Nuclear cap-binding complex	*CBC2, NPL3, STO1*
RNA export/transport	*APQ12, BUD13, KAP123, LOC1, LOS1, MFT1, NUP120, NUP133, NUP170, NUP188, NUP84, PML39, SCP160, SOL1, SXM1, THO1, THP2, TOM1*
Miscellaneous	*CTH1, MRT4, RPB4, SPT4, SSN3*

Table 2. Ty1*his3-AI* mobility in CNC⁻ and CNC⁺ *S. cerevisiae* strains

[a] Genotype	Ty1*his3-AI* mobility, x10⁻⁶ (SD)		Fold increase in CNC (CNC⁻/CNC⁺)
	CNC⁻	CNC⁺	
[b] Wild type	7481 (1805)	823 (318)	9
loc1Δ	2181 (534)	7 (2)	307
puf6Δ	1781 (748)	26 (4)	69
rps0aΔ	5002 (1585)	59 (48)	85
rps0bΔ	9592 (4285)	67 (22)	143
rpl7aΔ	5436 (1540)	89 (22)	61
rpl7bΔ	4212 (2308)	166 (55)	25
she2Δ	9370 (2038)	750 (231)	12

[a] *S. cerevisiae* CNC⁻ and isogenic CNC⁺ strains (WT and mutants) transformed with pBDG606 were induced with galactose. Cells were induced with galactose for 16 hrs at 22 °C. Then cells were diluted, spread onto SC-Ura (for total number of colonies) and SC-Ura-His (for His⁺ colonies) plates, and incubated for 3-4 days at 30 °C until colonies formed.

[b] Average of 5 trials.

Table S2.1 Yeast strains

Strain	Genotype	Plasmid	Reference or source
BY4741	*MATA, his3-Δ1, leu2-Δ0, lys2-Δ0, ura3-Δ0*		[43]
BY4742	*MATa, his3-Δ1, leu2-Δ0, lys2-Δ0, ura3-Δ0*		[43]
HWA215	BY4742 *loc1Δ::KanMX4*		This study
DG3453	*MATa, his3-Δ200, ura3, +1 Ty1*		This study
DG3648	*MATa, his3-Δ200, ura3, +20 Ty1*		This study
HWA210	DG3453 *loc1Δ::KanMX4*		This study
HWA211	DG3648 *loc1Δ::KanMX4*		This study
HWA410	DG3453 *rpl7aΔ::KanMX4*		This study
HWA411	DG3648 *rpl7aΔ::KanMX4*		This study
HWA408	DG3453 *rps0bΔ::KanMX4*		This study
HWA409	DG3648 *rps0bΔ::KanMX4*		This study
HWA15	BY4742	pBDG606 (pGTy1*his3-AI*/Cen-*URA3*)	This study
HWA169	HWA215	pBDG606	This study
DG3681	DG3453	pBDG606	This study

DG3682	DG3648	pBDG606	This study
HWA170	HWA210	pBDG606	This study
HWA197	HWA211	pBDG606	This study
HWA180	DG3453 *xrn1Δ::KanMX4*	pBDG606	This study
HWA195	DG3648 *xrn1Δ::KanMX4*	pBDG606	This study
HWA172	DG3453 *mrt4Δ::KanMX4*	pBDG606	This study
HWA196	DG3648 *mrt4Δ::KanMX4*	pBDG606	This study
HWA166	DG3453 *ssn3Δ::KanMX4*	pBDG606	This study
HWA194	DG3648 *ssn3Δ::KanMX4*	pBDG606	This study
HWA347	DG3453 *she2Δ::KanMX4*	pBDG606	This study
HWA348	DG3648 *she2Δ::KanMX4*	pBDG606	This study
HWA352	DG3453 *puf6Δ::KanMX4*	pBDG606	This study
HWA353	DG3648 *puf6Δ::KanMX4*	pBDG606	This study
HWA379	HWA410	pBDG606	This study
HWA380	HWA411	pBDG606	This study
HWA382	HWA408	pBDG606	This study
HWA383	HWA409	pBDG606	This study
HWA464	DG3453 *rpl7bΔ::KanMX4*	pBDG606	This study
HWA465	DG3648 *rpl7bΔ::KanMX4*	pBDG606	This study
HWA462	DG3453 *rps0aΔ::KanMX4*	pBDG606	This study
HWA463	DG3648 *rps0aΔ::KanMX4*	pBDG606	This study
HWA581	DG3453	pBDG1594	This study

HWA583	HWA210	pBDG1594	This study
HWA625	BY4742	pBDG1565	This study
HWA626	HWA215	pBDG1565	This study

Table S2.2 Functional annotation clustering of GO terms of Ty1 cofactors

[a] Cluster is a group of terms having similar biological meaning due to sharing similar gene members. Top 10 annotation clusters are shown here.

[b] Enrichment score is the overall enrichment score for the group based on the P-values of each term members. The higher, the more enriched.

[c] Category shows original database/resource where the terms orient. CC, cellular component. MF, molecular function. BP, biological processes.

[a] Annotation [b] Enrichment Score:
Cluster 1   7.082362706967514

[c] Category	Term	Count	%	PValue	Genes	List Total	Pop Hits	Pop Total	Fold Enrichment	Bonferroni	Benjamini	FDR
GOTERM_BP_FAT	GO:0000956~nuclear-transcribed mRNA catabolic process	25	5.23	5.32E-08	PAT1, KEM1, SKI8, LSM6, PUB1, DHH1, TRF5, NAM7, LSM1, NMD2, CTH1, CBC2, RRP6, PAP2, RPB4, STO1, UPF3, SSN3, LRP1, SSN2, CCR4, MRT4, HTZ1, POP2, MOT2	414	84	4870	3.50097769	6.76E-05	1.69E-05	8.66E-05
GOTERM_BP_FAT	GO:0006402~mRNA catabolic process	25	5.23	8.76E-08	PAT1, KEM1, SKI8, LSM6, PUB1, DHH1, TRF5, NAM7, LSM1, NMD2, CTH1, CBC2, RRP6, PAP2, RPB4, STO1, UPF3, SSN3, LRP1, SSN2, CCR4, MRT4, HTZ1, POP2, MOT2	414	86	4870	3.4195596	1.11E-04	1.59E-05	1.43E-04
GOTERM_BP_FAT	GO:0006401~RNA catabolic process	27	5.65	1.22E-07	PAT1, KEM1, POL32, LSM6, SKI8, PUB1, DHH1, TRF5, DBR1, NAM7, LSM1, NMD2, CTH1, CBC2, RRP6, PAP2, RPB4, STO1, UPF3, SSN3, LRP1, SSN2, CCR4, MRT4, HTZ1, POP2, MOT2	414	100	4870	3.17608696	1.54E-04	1.93E-05	1.98E-04

Annotation   Enrichment Score:
Cluster 2   7.079393899032657

Category	Term	Count	%	PValue	Genes	List Total	Pop Hits	Pop Total	Fold Enrichment	Bonferroni	Benjamini	FDR
GOTERM_CC_FAT	GO:0022626~cytosolic ribosome	37	7.74	2.43E-10	PAT1, RPS30A, RPS11A, RPL20B, RPL41B, RPL27A, RPL1B, RPL39, RPL34A, RPL15B, RPS27B, RPP2B, RPL43A, RPL14A, YGR054W, RPL7A, RPL19B, RPL19A, RPS25A, RPS19A, RPS19B, RPL21B, RPL21A, RPS9B, RPL2B, NAT1, RPS10A, RPS0B, RPL33B, RPL31A, RPL40A, RPL16B, RPL18A, RPP1A, RPL6A, RPL37A, ARD1, REI1	366	145	4595	3.20359902	9.13E-08	9.13E-08	3.36E-07
GOTERM_CC_FAT	GO:0033279~ribosomal subunit	42	8.79	1.44E-08	PAT1, RPS30A, RPS11A, RPL20B, RPL41B, RPL27A, RPL1B, RPL39, RPL34A, RPL15B, RPS27B, RPL43A, RPP2B, RPL14A, MRP17, YGR054W, RPL19B, MRPL39, RPL7A, RPL19A, RPS25A, RPS19A, RPS19B, RPL21B, RPL21A, RPS9B, MRPS28, RPL2B, MRPL8, RPS10A, MRPL7, RPS0B, RPL33B, RPL31A, RPL40A, RPL16B, RPL18A, RPP1A, RPL6A, MRPL49, RPL37A, RSM25, REI1	366	205	4595	2.57217113	5.43E-06	1.81E-06	2.00E-05
GOTERM_MF_FAT	GO:0003735~structural constituent of ribosome	39	8.16	1.14E-06	RPS30A, RPS11A, RPL20B, RPL41B, RPL27A, RPL1B, RPL39, RPL34A, RPL15B, RPS27B, RPL43A, RPP2B, RPL14A, MRP17, RPL19B, MRPL39, RPL7A, RPL19A, RPS25A, RPS19A, RPS19B, RPL21B, RPL21A, RPS9B, MRPS28, RPL2B, MRPL8, RPS10A, MRPL7, RPS0B, RPL33B, RPL31A, RPL40A, RPL16B, RPL18A, RPP1A, RPL6A, MRPL49, RPL37A, RSM25	359	199	4190	2.287342	5.64E-04	5.64E-04	0.001633
GOTERM_CC_FAT	GO:0005840~ribosome	48	10	1.21E-05	PAT1, RPS30A, RPL41B, MDY2, RPL1B, RPL15B, RPS27B, RPP2B, MRP17, MRPL39, RPL19B, RPL19A, RPL21B, RPL21A, RPS9B, MRPL8, MRPL7, RPS0B, RPL31A, RPL16B, RPP1A, MRPL49, ARD1, REI1, RPS11A, PML39, RPL20B, RPL27A, RPL39, RPL34A, RPL43A, RPL14A, RPL7A, GUF1, YGR054W, RPS25A, RPS19A, RPS19B, MRPS28, RPL2B, NAT1, RPS10A, RPL33B, RPL40A, RPL18A, ZUO1, RPL6A, RPL37A, RSM25	366	316	4595	1.90703465	0.0045363	5.05E-04	0.016713

Annotation Cluster 3   Enrichment Score: 5.320449483588934

Category	Term	Count	%	PValue	Genes	List Total	Pop Hits	Pop Total	Fold Enrichment	Bonferroni	Benjamini	FDR
GOTERM_BP_FAT	GO:0006403~RNA localization	32	6.69	7.77E-08	PML39, SCP160, NUP120, NUP188, TPM1, DHH1, SXM1, SHE4, LOS1, BUD13, SOL1, NAM7, THO1, RPS19A, RPS19B, APQ12, NUP170, NUP133, CBC2, RRP6, LOC1, RPS10A, RPS0B, NUP84, KAP123, STO1, MFT1, RPB4, LRP1, THP2, TOM1, NPL3	414	131	4870	2.8734742	9.87E-05	1.65E-05	1.27E-04
GOTERM_BP_FAT	GO:0050657~nucleic acid transport	26	5.44	4.60E-06	PML39, NUP120, NUP188, DHH1, SXM1, LOS1, BUD13, SOL1, THO1, RPS19A, RPS19B, APQ12, NUP133, NUP170, CBC2, LOC1, RPS10A, RPS0B, NUP84, STO1, MFT1, RPB4, KAP123, THP2, TOM1, NPL3	414	112	4870	2.73076259	0.0058209	5.31E-04	0.007484
GOTERM_BP_FAT	GO:0050658~RNA transport	26	5.44	4.60E-06	PML39, NUP120, NUP188, DHH1, SXM1, LOS1, BUD13, SOL1, THO1, RPS19A, RPS19B, APQ12, NUP133, NUP170, CBC2, LOC1, RPS10A, RPS0B, NUP84, STO1, MFT1, RPB4, KAP123, THP2, TOM1, NPL3	414	112	4870	2.73076259	0.0058209	5.31E-04	0.007484
GOTERM_BP_FAT	GO:0051236~establishment of RNA localization	26	5.44	4.60E-06	PML39, NUP120, NUP188, DHH1, SXM1, LOS1, BUD13, SOL1, THO1, RPS19A, RPS19B, APQ12, NUP133, NUP170, CBC2, LOC1, RPS10A, RPS0B, NUP84, STO1, MFT1, RPB4, KAP123, THP2, TOM1, NPL3	414	112	4870	2.73076259	0.0058209	5.31E-04	0.007484
GOTERM_BP_FAT	GO:0051028~mRNA transport	20	4.18	3.75E-05	APQ12, NUP133, NUP170, CBC2, LOC1, PML39, NUP84, NUP120, NUP188, RPB4, MFT1, STO1, KAP123, DHH1, SXM1, BUD13, THP2, TOM1, THO1, NPL3	414	82	4870	2.86909391	0.0465321	0.003398	0.061066
GOTERM_BP_FAT	GO:0015931~nucleobase, nucleoside, nucleotide and nucleic acid transport	27	5.65	4.22E-05	PML39, NUP120, NUP188, DHH1, SXM1, LOS1, BUD13, SOL1, THO1, RPS19A, RPS19B, APQ12, NUP133, NUP170, CBC2, LOC1, RPS10A, RPS0B, NUP84, FCY21, STO1, MFT1, RPB4, KAP123, THP2, TOM1, NPL3	414	134	4870	2.37021415	0.0521488	0.003564	0.068635

Annotation Cluster 4   Enrichment Score: 4.0214159391832105

Category	Term	Count	%	PValue	Genes	List Total	Pop Hits	Pop Total	Fold Enrichment	Bonferroni	Benjamini	FDR
GOTERM_BP_FAT	GO:0016575~histone deacetylation	12	2.51	5.35E-06	HDA3, UME1, SAP30, HST1, HDA1, PHO23, SNT1, HOS2, UME6, SIN3, DEP1, SET3	414	27	4870	5.22812668	0.0067686	5.66E-04	0.008706
GOTERM_BP_FAT	GO:0006476~protein amino acid deacetylation	12	2.51	1.77E-05	HDA3, UME1, SAP30, HST1, HDA1, PHO23, SNT1, HOS2, UME6, SIN3, DEP1, SET3	414	30	4870	4.70531401	0.0222508	0.001729	0.028842
GOTERM_CC_FAT	GO:0000118~histone deacetylase complex	12	2.51	2.68E-05	HDA3, UME1, SAP30, HST1, HDA1, PHO23, SNT1, HOS2, UME6, SIN3, DEP1, SET3	366	33	4595	4.56532538	0.0100079	0.001005	0.036969
GOTERM_CC_FAT	GO:0070210~Rpd3L-Expanded complex	9	1.88	5.09E-05	UME1, SAP30, PHO23, SNT1, HOS2, UME6, SIN3, DEP1, SET3	366	19	4595	5.94693701	0.018946	0.001737	0.070292
GOTERM_MF_FAT	GO:0004407~histone deacetylase activity	10	2.09	1.30E-04	HDA3, UME1, SAP30, HST1, HDA1, PHO23, HOS2, UME6, SIN3, DEP1	359	25	4190	4.66852368	0.0627328	0.021364	0.187383
GOTERM_MF_FAT	GO:0033558~protein deacetylase activity	10	2.09	1.30E-04	HDA3, UME1, SAP30, HST1, HDA1, PHO23, HOS2, UME6, SIN3, DEP1	359	25	4190	4.66852368	0.0627328	0.021364	0.187383
GOTERM_MF_FAT	GO:0019213~deacetylase activity	10	2.09	4.69E-04	HDA3, UME1, SAP30, HST1, HDA1, PHO23, HOS2, UME6, SIN3, DEP1	359	29	4190	4.02458938	0.2081243	0.056669	0.673278
GOTERM_MF_FAT	GO:0016811~hydrolase activity, acting on carbon-nitrogen (but not peptide) bonds, in linear amides	10	2.09	0.006554	HDA3, UME1, SAP30, HST1, HDA1, PHO23, HOS2, UME6, SIN3, DEP1	359	41	4190	2.84666078	0.9619289	0.278794	9.027943

Annotation Cluster 5   Enrichment Score: 3.8459811183353207

Category	Term	Count	%	PValue	Genes	List Total	Pop Hits	Pop Total	Fold Enrichment	Bonferroni	Benjamini	FDR
GOTERM_BP_FAT	GO:0031328~positive regulation of cellular biosynthetic process	35	7.32	4.24E-05	TUP1, SPT3, SPT4, INO2, INO4, SWI3, SNF5, SNF6, ISW1, HDA1, SWI6, MKT1, SPT20, RTF1, SPT23, GON7, GUF1, SNF1, RPN4, SNF2, CBF1, GCN5, RPS9B, HAC1, SUB1, RAS2, SNF12, SIP4, CTK1, ITC1, SPT8, SIN4, HOS2, SIN3, TEC1	414	198	4870	2.07936856	0.0524281	0.00336	0.069012
GOTERM_BP_FAT	GO:0009891~positive regulation of biosynthetic process	35	7.32	4.24E-05	TUP1, SPT3, SPT4, INO2, INO4, SWI3, SNF5, SNF6, ISW1, HDA1, SWI6, MKT1, SPT20, RTF1, SPT23, GON7, GUF1, SNF1, RPN4, SNF2, CBF1, GCN5, RPS9B, HAC1, SUB1, RAS2, SNF12, SIP4, CTK1, ITC1, SPT8, SIN4, HOS2, SIN3, TEC1	414	198	4870	2.07936856	0.0524281	0.00336	0.069012
GOTERM_BP_FAT	GO:0010557~positive regulation of macromolecule biosynthetic process	33	6.9	1.58E-04	TUP1, SPT3, SPT4, INO2, INO4, SWI3, SNF5, SNF6, ISW1, HDA1, SWI6, MKT1, SPT20, RTF1, SPT23, GON7, GUF1, RPN4, SNF2, CBF1, GCN5, RPS9B, HAC1, SUB1, RAS2, SNF12, ITC1, CTK1, SPT8, SIN4, HOS2, SIN3, TEC1	414	194	4870	2.00097116	0.18167	0.011724	0.256687
GOTERM_BP_FAT	GO:0010604~positive regulation of macromolecule metabolic process	35	7.32	1.73E-04	SPT3, TUP1, SPT4, INO2, INO4, SWI3, SNF5, SNF6, ISW1, HDA1, SWI6, MKT1, SPT20, RTF1, SPT23, GON7, GUF1, RPN4, SNF2, CBF1, GCN5, RPS9B, HAC1, SUB1, STE5, RAS2, SNF12, CTK1, ITC1, SPT8, SIN4, HOS2, SIN3, TEC1, NPL3	414	212	4870	1.94205177	0.1974165	0.012143	0.281528
GOTERM_BP_FAT	GO:0045941~positive regulation of transcription	30	6.28	2.18E-04	TUP1, SPT3, SPT4, INO2, INO4, SWI3, SNF5, SNF6, ISW1, HDA1, SWI6, SPT20, RTF1, SPT23, GON7, RPN4, SNF2, CBF1, GCN5, HAC1, SUB1, RAS2, SNF12, ITC1, CTK1, SPT8, SIN4, HOS2, SIN3, TEC1	414	172	4870	2.05173576	0.2419317	0.013754	0.354446

Category	Term	Count	%	PValue	Genes	List Total	Pop Hits	Pop Total	Fold Enrichment	Bonferroni	Benjamini	FDR
GOTERM_BP_FAT	GO:0010628~positive regulation of gene expression	30	6.28	2.42E-04	TUP1, SPT3, SPT4, INO2, INO4, SWI3, SNF5, SNF6, ISW1, HDA1, SWI6, SPT20, RTF1, SPT23, GON7, RPN4, SNF2, CBF1, GCN5, HAC1, SUB1, RAS2, SNF12, ITC1, CTK1, SPT8, SIN4, HOS2, SIN3, TEC1	414	173	4870	2.03987602	0.2643992	0.014516	0.392871
GOTERM_BP_FAT	GO:0045935~positive regulation of nucleobase, nucleoside, nucleotide and nucleic acid metabolic process	31	6.49	2.57E-04	TUP1, SPT3, SPT4, INO2, INO4, SWI3, SNF5, SNF6, ISW1, HDA1, SWI6, SPT20, RTF1, SPT23, GON7, RPN4, SNF2, CBF1, GCN5, HAC1, SUB1, RAS2, SNF12, ITC1, CTK1, SPT8, SIN4, HOS2, SIN3, TEC1, NPL3	414	182	4870	2.00363646	0.2781934	0.014709	0.41704
GOTERM_BP_FAT	GO:0051173~positive regulation of nitrogen compound metabolic process	31	6.49	2.57E-04	TUP1, SPT3, SPT4, INO2, INO4, SWI3, SNF5, SNF6, ISW1, HDA1, SWI6, SPT20, RTF1, SPT23, GON7, RPN4, SNF2, CBF1, GCN5, HAC1, SUB1, RAS2, SNF12, ITC1, CTK1, SPT8, SIN4, HOS2, SIN3, TEC1, NPL3	414	182	4870	2.00363646	0.2781934	0.014709	0.41704

Annotation Cluster 6   Enrichment Score: 3.5485229061900907

Category	Term	Count	%	PValue	Genes	List Total	Pop Hits	Pop Total	Fold Enrichment	Bonferroni	Benjamini	FDR
GOTERM_CC_FAT	GO:0031981~nuclear lumen	68	14.2	1.02E-04	HFI1, SSF1, SBP1, PHO23, TUP1, LSM6, PUF6, HDA3, LOS1, SAP30, HDA1, SWI6, SPT20, TRF5, RTF1, SRB8, LSM1, TGS1, GCN5, RPS9B, HST1, LOC1, RRP6, RRP8, HMO1, BUD22, RPA49, BUD21, UME1, SPT7, SPT8, TAF14, CCR4, MRT4, RSA3, TOM1, PGD1, HOS2, UME6, MED2, MOT2, NPL3, CAF40, SGF73, SPT3, SPT4, IWR1, DEP1, NCL1, SNF2, UTP30, REF2, PAP2, RPB4, MFT1, SET3, SSN3, NOP12, SSN2, CTK1, ITC1, THP2, DBP3, POP2, SIN4, SNT1, SIN3, DBP7	366	545	4595	1.5664511	0.0375284	0.002938	0.140518
GOTERM_CC_FAT	GO:0070013~intracellular organelle lumen	88	18.4	3.41E-04	HFI1, SSF1, MDJ1, SBP1, RPO41, PET130, PHO23, LSM6, TUP1, SAP30, LOS1, PUF6, HDA3, ATP1, HDA1, SWI6, SPT20, RTF1, TRF5, SRB8, MRP17, MRPL39, LSM1, TGS1, GCN5, RPS9B, HST1, RRP6, LOC1, MRPL8, MRPL7, KGD1, PIM1, RRP8, HMO1, MSW1, BUD22, RPA49, BUD21, UME1, SPT7, SPT8, CCR4, SCJ1, TAF14, MRT4, RSA3, MRPL49, TOM1, PGD1, HOS2, UME6, MED2, MOT2, NPL3, FMC1, CAF40, SGF73, SPT3, SPT4, IWR1, DEP1, NCL1, ISU1, GLO4, GUF1, SNF2, UTP30, REF2, MRPS28, PAP2, RPB4, MFT1, SET3, SSN3, NOP12, SSN2, CTK1, ITC1, THP2, POP2, DBP3, SIN4, SNT1, CCP1, RSM25, DBP7, SIN3	366	783	4595	1.41099456	0.1202322	0.008503	0.469801
GOTERM_CC_FAT	GO:0043233~organelle lumen	88	18.4	3.41E-04	HFI1, SSF1, MDJ1, SBP1, RPO41, PET130, PHO23, LSM6, TUP1, SAP30, LOS1, PUF6, HDA3, ATP1, HDA1, SWI6, SPT20, RTF1, TRF5, SRB8, MRP17, MRPL39, LSM1, TGS1, GCN5, RPS9B, HST1, RRP6, LOC1, MRPL8, MRPL7, KGD1, PIM1, RRP8, HMO1, MSW1, BUD22, RPA49, BUD21, UME1, SPT7, SPT8, CCR4, SCJ1, TAF14, MRT4, RSA3, MRPL49, TOM1, PGD1, HOS2, UME6, MED2, MOT2, NPL3, FMC1, CAF40, SGF73, SPT3, SPT4, IWR1, DEP1, NCL1, ISU1, GLO4, GUF1, SNF2, UTP30, REF2, MRPS28, PAP2, RPB4, MFT1, SET3, SSN3, NOP12, SSN2, CTK1, ITC1, THP2, POP2, DBP3, SIN4, SNT1, CCP1, RSM25, DBP7, SIN3	366	783	4595	1.41099456	0.1202322	0.008503	0.469801
GOTERM_CC_FAT	GO:0031974~membrane-enclosed lumen	91	19	5.42E-04	HFI1, SSF1, MDJ1, SBP1, RPO41, PET130, PHO23, LSM6, TUP1, SAP30, LOS1, PUF6, HDA3, ATP1, HDA1, SWI6, SPT20, RTF1, TRF5, SRB8, MRP17, MRPL39, LSM1, TGS1, GCN5, RPS9B, HST1, RRP6, LOC1, GAL83, MRPL8, MRPL7, KGD1, PIM1, RRP8, HMO1, MSW1, BUD22, RPA49, BUD21, UME1, SPT7, SPT8, CCR4, SCJ1, TAF14, MRT4, RSA3, MRPL49, TOM1, PGD1, HOS2, UME6, MED2, MOT2, NPL3, FMC1, CAF40, SGF73, SPT3, SPT4, IWR1, DEP1, NCL1, ISU1, GLO4, GUF1, SNF1, SNF2, MDM35, UTP30, REF2, MRPS28, PAP2, RPB4, MFT1, SET3, SSN3, NOP12, SSN2, CTK1, ITC1, THP2, POP2, DBP3, SIN4, SNT1, CCP1, RSM25, DBP7, SIN3	366	827	4595	1.38146636	0.1843835	0.01067	0.746445

Annotation Cluster 7   Enrichment Score: 3.3440855095847772

Category	Term	Count	%	PValue	Genes	List Total	Pop Hits	Pop Total	Fold Enrichment	Bonferroni	Benjamini	FDR
GOTERM_BP_FAT	GO:0006417~regulation of translation	31	6.49	2.10E-04	PAT1, RPS11A, SBP1, RPL20B, RPL1B, PUB1, DHH1, SXM1, PUF6, MKT1, SSZ1, YGR054W, NAM7, RPL7A, GUF1, NMD2, SNF1, GCN3, RPS9B, RPL2B, NAT1, HIS4, RPS0B, RPL31A, UPF3, KAP123, RPL18A, CTK1, ZUO1, RPL6A	414	180	4870	2.02589909	0.2341378	0.013942	0.341379
GOTERM_BP_FAT	GO:0010608~posttranscriptional regulation of gene expression	31	6.49	6.57E-04	PAT1, RPS11A, SBP1, RPL20B, RPL1B, PUB1, DHH1, SXM1, PUF6, MKT1, SSZ1, YGR054W, NAM7, RPL7A, GUF1, NMD2, SNF1, GCN3, EAP1, RPS9B, RPL2B, NAT1, HIS4, RPS0B, RPL31A, UPF3, KAP123, RPL18A, CTK1, ZUO1, RPL6A	414	192	4870	1.89928039	0.5660779	0.028379	1.064578

Category	Term	Count	%	PValue	Genes	List Total	Pop Hits	Pop Total	Fold Enrichment	Bonferroni	Benjamini	FDR
GOTERM_B P_FAT	GO:0032268~regulatio n of cellular protein metabolic process	32	6.69	6.73E-04	PAT1, RPS11A, SBP1, RPL20B, RPL1B, PUB1, DHH1, SXM1, PUF6, MKT1, SSZ1, YGR054W, GUF1, NAM7, RPL7A, NMD2, SNF1, GCN3, EAP1, RPS9B, RPL2B, NAT1, HIS4, STE5, RPS0B, RPL31A, UPF3, KAP123, RPL18A, CTK1, ZUO1, RPL6A	414	201	4870	1.87276179	0.5745362	0.026352	1.089541

Annotation Cluster 8 Enrichment Score: 2.9603929318321383

Category	Term	Count	%	PValue	Genes	List Total	Pop Hits	Pop Total	Fold Enrichment	Bonferroni	Benjamini	FDR
GOTERM_B P_FAT	GO:0010558~negative regulation of macromolecule biosynthetic process	35	7.32	3.57E-04	SPT2, SBP1, PHO23, SPT3, TUP1, SCP160, SPT4, MIG3, VPS72, SPT10, CHK1, HDA3, SAP30, ISW1, HDA1, SRB8, NAM7, SNF1, PST2, CBF1, EAP1, GCN5, HST1, IES3, SET3, SSN3, SSN2, ITC1, CCR4, SPT8, TAF14, HTZ1, SIN4, UME6, SIN3	414	220	4870	1.87143171	0.364477	0.018711	0.57943
GOTERM_B P_FAT	GO:0010629~negative regulation of gene expression	32	6.69	5.65E-04	SPT2, PHO23, TUP1, SCP160, SPT3, SPT4, MIG3, VPS72, SPT10, HDA3, SAP30, ISW1, HDA1, SRB8, NAM7, PST2, CBF1, GCN5, HST1, RRP6, IES3, PAP2, SET3, SSN3, SSN2, ITC1, SPT8, TAF14, HTZ1, SIN4, UME6, SIN3	414	199	4870	1.89158352	0.5119043	0.027209	0.915254
GOTERM_B P_FAT	GO:0010605~negative regulation of macromolecule metabolic process	39	8.16	6.46E-04	SPT2, NUM1, SBP1, PHO23, SPT3, TUP1, SCP160, SPT4, MIG3, VPS72, SPT10, CHK1, HDA3, SAP30, ISW1, HDA1, SRB8, NAM7, SNF1, PST2, CBF1, EAP1, GCN5, HST1, RRP6, NIP100, IES3, PAP2, SET3, SSN3, SSN2, ITC1, TAF14, SPT8, CCR4, HTZ1, SIN4, UME6, SIN3	414	263	4870	1.74436546	0.5596248	0.028865	1.045853
GOTERM_B P_FAT	GO:0031327~negative regulation of cellular biosynthetic process	35	7.32	8.86E-04	SPT2, SBP1, PHO23, SPT3, TUP1, SCP160, SPT4, MIG3, VPS72, SPT10, CHK1, HDA3, SAP30, ISW1, HDA1, SRB8, NAM7, SNF1, PST2, CBF1, EAP1, GCN5, HST1, IES3, SET3, SSN3, SSN2, ITC1, CCR4, SPT8, TAF14, HTZ1, SIN4, UME6, SIN3	414	231	4870	1.78231591	0.6753882	0.03255	1.431999
GOTERM_B P_FAT	GO:0009890~negative regulation of biosynthetic process	35	7.32	0.001034	SPT2, SBP1, PHO23, SPT3, TUP1, SCP160, SPT4, MIG3, VPS72, SPT10, CHK1, HDA3, SAP30, ISW1, HDA1, SRB8, NAM7, SNF1, PST2, CBF1, EAP1, GCN5, HST1, IES3, SET3, SSN3, SSN2, ITC1, CCR4, SPT8, TAF14, HTZ1, SIN4, UME6, SIN3	414	233	4870	1.76701706	0.7313269	0.035849	1.670698
GOTERM_B P_FAT	GO:0016481~negative regulation of transcription	30	6.28	0.001935	SPT2, PHO23, TUP1, SCP160, SPT3, SPT4, MIG3, VPS72, SPT10, HDA3, SAP30, ISW1, HDA1, SRB8, NAM7, PST2, CBF1, GCN5, HST1, IES3, SET3, SSN3, SSN2, ITC1, SPT8, TAF14, HTZ1, SIN4, UME6, SIN3	414	196	4870	1.80050281	0.9145995	0.055612	3.104883
GOTERM_B P_FAT	GO:0051172~negative regulation of nitrogen compound metabolic process	32	6.69	0.002999	SPT2, PHO23, SPT3, TUP1, SCP160, SPT4, MIG3, VPS72, SPT10, CHK1, HDA3, SAP30, ISW1, HDA1, SRB8, NAM7, PST2, CBF1, GCN5, HST1, IES3, SET3, SSN3, SSN2, ITC1, CCR4, SPT8, TAF14, HTZ1, SIN4, UME6, SIN3	414	220	4870	1.71102328	0.9779575	0.076399	4.772695
GOTERM_B P_FAT	GO:0045934~negative regulation of nucleobase, nucleoside, nucleotide and nucleic acid metabolic process	32	6.69	0.002999	SPT2, PHO23, SPT3, TUP1, SCP160, SPT4, MIG3, VPS72, SPT10, CHK1, HDA3, SAP30, ISW1, HDA1, SRB8, NAM7, PST2, CBF1, GCN5, HST1, IES3, SET3, SSN3, SSN2, ITC1, CCR4, SPT8, TAF14, HTZ1, SIN4, UME6, SIN3	414	220	4870	1.71102328	0.9779575	0.076399	4.772695

Annotation Cluster 9 Enrichment Score: 2.673324067021937

Category	Term	Count	%	PValue	Genes	List Total	Pop Hits	Pop Total	Fold Enrichment	Bonferroni	Benjamini	FDR
GOTERM_B P_FAT	GO:0006405~RNA export from nucleus	19	3.97	6.60E-04	APQ12, NUP133, NUP170, RPS10A, RPS0B, NUP84, NUP120, NUP188, RPB4, MFT1, SXM1, LOS1, BUD13, THP2, SOL1, RPS19A, THO1, NPL3, RPS19B	414	93	4870	2.40325178	0.5677329	0.02757	1.069424
GOTERM_B P_FAT	GO:0051169~nuclear transport	24	5.02	0.002956	APQ12, NUP133, NUP170, HIS3, RPS10A, RPS0B, NUP84, NUP120, NUP188, RPB4, MFT1, KAP123, SXM1, LOS1, BUD13, THP2, TOM1, SOL1, NPL6, REI1, RPS19A, THO1, NPL3, RPS19B	414	148	4870	1.90755973	0.976708	0.076877	4.705364
GOTERM_B P_FAT	GO:0006913~nucleocy toplasmic transport	24	5.02	0.002956	APQ12, NUP133, NUP170, HIS3, RPS10A, RPS0B, NUP84, NUP120, NUP188, RPB4, MFT1, KAP123, SXM1, LOS1, BUD13, THP2, TOM1, SOL1, NPL6, REI1, RPS19A, THO1, NPL3, RPS19B	414	148	4870	1.90755973	0.976708	0.076877	4.705364
GOTERM_B P_FAT	GO:0051168~nuclear export	21	4.39	0.003513	APQ12, NUP133, NUP170, HIS3, RPS10A, RPS0B, NUP84, NUP120, NUP188, RPB4, MFT1, SXM1, LOS1, BUD13, THP2, SOL1, REI1, RPS19A, THO1, NPL3, RPS19B	414	124	4870	1.99216924	0.9885406	0.085501	5.567932

Annotation Cluster 10 Enrichment Score: 2.630927527064271

Category	Term	Count	%	PValue	Genes	List Total	Pop Hits	Pop Total	Fold Enrichment	Bonferroni	Benjamini	FDR
GOTERM_B P_FAT	GO:0008361~regulatio n of cell size	20	4.18	0.001411	DIA1, VPH1, MUC1, SSF1, DFG10, GAL83, SPT3, STE5, STE4, RAS2, TPM1, LGE1, RPA49, KTI12, VMA9, DBR1, TOM1, SNF1, NPL3, TEC1	414	107	4870	2.19874486	0.8336671	0.044952	2.273273
GOTERM_B P_FAT	GO:0016049~cell growth	16	3.35	0.002231	DIA1, VPH1, MUC1, DFG10, GAL83, SPT3, STE5, STE4, RAS2, TPM1, KTI12, VMA9, DBR1, SNF1, NPL3, TEC1	414	79	4870	2.38243747	0.9413428	0.059791	3.570363
GOTERM_B P_FAT	GO:0032535~regulatio n of cellular component size	20	4.18	0.004066	DIA1, VPH1, MUC1, SSF1, DFG10, GAL83, SPT3, STE5, STE4, RAS2, TPM1, LGE1, RPA49, KTI12, VMA9, DBR1, TOM1, SNF1, NPL3, TEC1	414	117	4870	2.01081795	0.9943381	0.09471	6.417599

Table S2.3 Ty1 protein levels and Ty1his3-AI mobility of mutant strains listed in Table

2.1

[a] Wild type control intensity was set to 1 and fold change between wild type and mutant

was calculated.

[b] Ty1his3-AI mobility: * (No decrease/similar to wild type control) ** (mild) *** (moderate)

**** (severe) ***** (very severe) decrease

Deletion strain tested	[a] IN level fold change	[a] Gag level fold change	IN/Gag	[b] Decrease in Ty1*his3-AI* mobility	Reference
apq12Δ	1.99	1.02	1.8	***	[34, 35]
bud13Δ	1.21	1.05	1.2	**	[34]
bud31Δ	0.05	0.31	0.1	****	[34]
cbc2Δ	0.03	0.58	0.0	***	[33]
ccr4Δ	0.31	0.33	0.9	*	[35]
cth1Δ	1.15	1.22	1.1	*	[35]
dbr1Δ	1.33	1.15	1.3	***	[33]
dhh1Δ	0.28	0.32	0.9	****	[34, 35]
kap123Δ	0.17	0.83	0.3	*****	[34]
lea1Δ	1.50	1.13	1.6	***	[33]
loc1Δ	0.14	1.07	0.1	*****	[34, 35]
los1Δ	0.78	0.98	0.7	*	[35]
lrp1Δ	0.45	0.99	0.5	**	[34]
lsm1Δ	0.05	0.58	0.1	****	[33-35]
lsm6Δ	0.68	0.68	1.0	****	[35]
mft1Δ	0.08	0.71	0.1	**	[33]
mot2Δ	0.37	1.34	0.2	*****	[34]
mrt4Δ	0.35	0.79	0.5	****	[35]
mud2Δ	0.08	0.73	0.1	***	[35]
nam7Δ	0.68	1.16	0.7	**	[34, 35]
nmd2Δ	0.47	1.28	0.4	**	[34]
npl3Δ	0.11	0.48	0.3	****	[34]
nup120Δ	1.26	1.00	1.2	***	[35]
nup133Δ	0.60	1.16	0.5	***	[33, 35]
nup170Δ	0.18	0.63	0.3	****	[34, 35]
nup188Δ	0.16	0.94	0.1	***	[35]
nup84Δ	2.02	1.01	1.7	***	[33, 34]
pap2Δ	0.30	0.78	0.3	****	[34]

pat1Δ	0.12	0.48	0.2	**	[33, 34]
pml39Δ	0.25	0.94	0.2	*	[35]
pop2Δ	0.40	0.69	0.6	****	[33, 34]
pub1Δ	0.42	1.05	0.5	*	[34]
ref2Δ	0.19	0.74	0.2	****	[34, 35]
rpb4Δ	0.04	0.52	0.1	****	[35]
rrp6Δ	0.92	1.22	0.8	***	[34]
scp160Δ	0.13	0.36	0.4	****	[33, 34]
ski8Δ	0.42	0.87	0.53	**	[34, 35]
snt309Δ	1.43	0.85	1.2	***	[35]
snu66Δ	0.40	0.81	0.5	***	[34]
sol1Δ	0.18	0.87	0.2	*	[35]
spt4Δ	0.00	0.34	0.0	****	[33, 34]
sqs1Δ	0.22	0.89	0.2	*	[35]
ssn3Δ	0.01	0.61	0.0	****	[34]
sto1Δ	0.08	0.27	0.3	***	[33]
sxm1Δ	1.34	0.82	1.6	*	[34]
tho1Δ	0.96	0.89	1.0	*	[34]
thp2Δ	0.52	0.76	0.6	**	[33, 35]
tom1Δ	1.74	1.00	1.4	*	[35]
trf5Δ	0.44	1.05	0.4	**	[35]
upf3Δ	0.51	1.24	0.5	**	[34, 35]
xrn1Δ	0.06	0.84	0.1	*****	[34]

Ty1*his3-AI* mobility criteria

Table S2.4

(1) List of trasncripts showing increased expression in *loc1*Δ (two-fold or more) compared to WT and padj of <0.05

Gene name	id	base Mean	Base MeanA (loc1)	Base MeanB (WT)	fold Change	log2 Fold Change	pval	padj
CSS1	YIL169C	149.66	258.37	40.95	6.31	2.66	2.17E-35	1.96E-32
COB	Q0105	103.16	175.35	30.97	5.66	2.50	6.86E-06	0.000209311
SNO1	YMR095C	29.57	48.05	11.09	4.33	2.12	1.60E-08	1.13E-06
Q0255	Q0255	47.92	77.50	18.35	4.22	2.08	0.00131282	0.012844842
YGL117W	YGL117W	783.40	1246.39	320.40	3.89	1.96	5.35E-12	7.33E-10
SIP4	YJL089W	28.28	44.47	12.09	3.68	1.88	2.75E-06	9.53E-05
YOR329W-A	YOR329W-A	8.57	13.45	3.70	3.64	1.86	0.006360754	0.039879712
ISU2	YOR226C	1153.61	1803.20	504.02	3.58	1.84	0.000236806	0.003438052
PRM2	YIL037C	57.11	88.72	25.51	3.48	1.80	5.26E-05	0.001076978
NDT80	YHR124W	65.00	100.86	29.13	3.46	1.79	0.000498572	0.006006693
CPA2	YJR109C	2039.89	3159.11	920.66	3.43	1.78	7.83E-34	4.94E-31
MUC1	YIR019C	14.78	22.77	6.79	3.35	1.75	0.001466907	0.013795494
ATP6	Q0085	154.15	236.97	71.33	3.32	1.73	0.006400147	0.040007269
HPF1	YOL155C	438.07	672.79	203.35	3.31	1.73	7.52E-28	3.65E-25
PCL5	YHR071W	573.81	878.57	269.04	3.27	1.71	1.04E-05	0.000282342
COX1	Q0045	197.69	299.90	95.48	3.14	1.65	0.00262183	0.020885148
CRC1	YOR100C	80.46	121.23	39.69	3.05	1.61	8.19E-08	4.78E-06
COX2	Q0250	423.72	638.22	209.22	3.05	1.61	0.000358247	0.004673527
IZH4	YOL101C	628.90	947.12	310.67	3.05	1.61	2.06E-26	9.26E-24
BI3	Q0115	23.83	35.73	11.93	2.99	1.58	0.000140337	0.002314327
AI3	Q0060	62.79	93.81	31.76	2.95	1.56	6.69E-05	0.00129625
AI2	Q0055	444.50	662.63	226.37	2.93	1.55	0.006064996	0.038601558
FUR4	YBR021W	212.97	310.13	115.81	2.68	1.42	3.96E-12	5.58E-10
LEU4	YNL104C	2500.44	3636.05	1364.84	2.66	1.41	1.71E-25	6.74E-23
HSP12	YFL014W	54.75	79.54	29.97	2.65	1.41	6.27E-07	2.62E-05
YJR149W	YJR149W	290.01	421.28	158.73	2.65	1.41	1.08E-15	2.44E-13

PHO92	YDR374C	18.37	26.63	10.10	2.64	1.40	0.002379281	0.019520635
YLR460C	YLR460C	88.45	127.96	48.94	2.61	1.39	1.68E-07	8.21E-06
SRT1	YMR101C	40.81	58.89	22.73	2.59	1.37	1.33E-05	0.0003426
RIM4	YHL024W	21.72	31.31	12.12	2.58	1.37	0.001218743	0.012131598
MUP3	YHL036W	268.09	382.02	154.16	2.48	1.31	0.005131584	0.034324959
MPC2	YHR162W	7553.65	10681.49	4425.81	2.41	1.27	3.95E-15	8.03E-13
AI4	Q0065	153.22	216.42	90.01	2.40	1.27	0.000441046	0.005481319
GCV3	YAL044C	6527.11	9217.05	3837.17	2.40	1.26	2.47E-15	5.19E-13
MDN1	YLR106C	1821.19	2571.07	1071.31	2.40	1.26	2.26E-18	6.48E-16
HXT2	YMR011W	666.59	941.03	392.16	2.40	1.26	0.000126269	0.002133036
REC8	YPR007C	92.99	130.93	55.05	2.38	1.25	0.000176586	0.002740563
FMP43	YGR243W	122.26	171.99	72.54	2.37	1.25	1.06E-05	0.000287137
SLH1	YGR271W	857.93	1204.95	510.92	2.36	1.24	3.01E-17	7.29E-15
MAL33	YBR297W	245.61	344.76	146.46	2.35	1.24	5.55E-13	9.20E-11
DAN4	YJR151C	46.55	65.21	27.88	2.34	1.23	0.003753212	0.027308303
GTO3	YMR251W	36.27	50.62	21.91	2.31	1.21	0.002078184	0.017600318
YLR112W	YLR112W	133.87	186.53	81.21	2.30	1.20	2.34E-09	1.94E-07
FLO5	YHR211W	68.17	94.88	41.46	2.29	1.19	8.88E-06	0.000250986
GLC3	YEL011W	178.01	247.03	108.99	2.27	1.18	9.45E-10	8.57E-08
ARG1	YOL058W	394.80	544.55	245.06	2.22	1.15	3.14E-13	5.66E-11
ARG4	YHR018C	2905.01	4005.05	1804.98	2.22	1.15	1.31E-17	3.31E-15
PDH1	YPR002W	264.80	363.96	165.64	2.20	1.14	1.36E-09	1.21E-07
NCA3	YJL116C	28.23	38.71	17.75	2.18	1.13	0.002387596	0.01953798
GAS2	YLR343W	23.92	32.80	15.05	2.18	1.12	0.006458835	0.04029418
YIL014C-A	YIL014C-A	52.40	71.73	33.07	2.17	1.12	0.000218778	0.003248914
AI5_ALPHA	Q0070	101.22	138.45	63.99	2.16	1.11	0.002162621	0.01809651
ARO10	YDR380W	2268.48	3102.02	1434.95	2.16	1.11	0.000111155	0.001924139
ARG7	YMR062C	1330.93	1819.83	842.03	2.16	1.11	2.14E-16	4.99E-14
DLS1	YJL065C	91.11	124.48	57.73	2.16	1.11	0.004053602	0.028828159
YHR033W	YHR033W	91.40	124.69	58.11	2.15	1.10	8.83E-06	0.000250986
TOM1	YDR457W	1709.88	2316.85	1102.91	2.10	1.07	7.91E-13	1.28E-10
DIA3	YDL024C	37.59	50.86	24.32	2.09	1.06	0.002018591	0.017257995

MSN4	YKL062W	193.73	262.00	125.46	2.09	1.06	1.89E-08	1.29E-06
YDL199C	YDL199C	114.45	154.73	74.16	2.09	1.06	1.41E-06	5.29E-05
FIG2	YCR089W	152.02	205.32	98.72	2.08	1.06	0.000518039	0.006193856
LYS1	YIR034C	415.32	560.76	269.88	2.08	1.06	4.71E-11	5.40E-09
RIB5	YBR256C	2507.96	3386.18	1629.75	2.08	1.06	2.88E-08	1.89E-06
ARO3	YDR035W	1885.27	2542.22	1228.32	2.07	1.05	2.35E-14	4.49E-12
ARO1	YDR127W	1700.87	2275.91	1125.83	2.02	1.02	5.16E-13	8.80E-11
ATO3	YDR384C	497.54	664.97	330.11	2.01	1.01	0.000726402	0.007973973
YAR009C	YAR009C	47.23	63.01	31.46	2.00	1.00	0.000967396	0.009911484

(2) List of trasncripts showing decreased expression in *loc1*Δ (two-fold or more) compared to WT and padj of <0.05

Gene name	id	base Mean	base Mean A (loc1)	base Mean B (WT)	fold Change	log2 Fold Change	pval	padj
LOC1	YFR001W	713.81	2.26	1425.36	0.00	-9.30	1.07E-218	6.74E-215
YBR144C	YBR144C	4.02	0.52	7.52	0.07	-3.85	0.003821506	0.027600209
HO	YDL227C	851.37	129.33	1573.41	0.08	-3.60	5.34E-110	1.68E-106
VPS61	YDR136C	4.88	0.79	8.98	0.09	-3.51	0.000845376	0.008952461
TIR1	YER011W	76.63	12.66	140.60	0.09	-3.47	6.06E-33	3.18E-30
YLR030W	YLR030W	100.28	18.85	181.71	0.10	-3.27	6.15E-36	6.46E-33
YLR366W	YLR366W	83.23	19.95	146.51	0.14	-2.88	3.30E-18	9.03E-16
FCY21	YER060W	214.48	55.18	373.78	0.15	-2.76	1.16E-44	1.83E-41
YPR195C	YPR195C	10.75	3.04	18.47	0.16	-2.60	4.79E-05	0.001010448
OPT2	YPR194C	91.46	26.45	156.47	0.17	-2.56	7.17E-24	2.38E-21
YOL118C	YOL118C	10.63	3.11	18.16	0.17	-2.55	0.00024567	0.003534168
PGU1	YJR153W	25.48	7.62	43.35	0.18	-2.51	3.09E-09	2.50E-07
RPS22B	YLR367W	2395.22	719.11	4071.32	0.18	-2.50	8.52E-35	6.13E-32
YGR069W	YGR069W	8.43	2.78	14.08	0.20	-2.34	0.00163346	0.014830713
ANB1	YJR047C	203.88	70.27	337.48	0.21	-2.26	2.67E-33	1.53E-30
DAN1	YJR150C	24.78	9.02	40.55	0.22	-2.17	0.002002255	0.017188297
RPL9B	YNL067W	4279.57	1611.03	6948.11	0.23	-2.11	6.84E-45	1.44E-41

YHL012W	YHL012W	47.67	17.99	77.36	0.23	-2.10	6.34E-11	7.01E-09
RPL8A	YHL033C	3843.25	1457.94	6228.57	0.23	-2.09	4.84E-43	6.10E-40
YGP1	YNL160W	115.86	46.76	184.96	0.25	-1.98	3.12E-08	2.02E-06
YLR031W	YLR031W	119.27	50.63	187.92	0.27	-1.89	4.81E-09	3.65E-07
EDS1	YBR033W	22.64	9.65	35.62	0.27	-1.88	2.25E-05	0.000535753
RPL7B	YPL198W	426.93	183.54	670.33	0.27	-1.87	6.55E-19	1.97E-16
IMD2	YHR216W	699.32	301.08	1097.57	0.27	-1.87	2.14E-22	6.74E-20
WSC4	YHL028W	163.66	70.90	256.42	0.28	-1.85	5.53E-06	0.000173394
MSH5	YDL154W	276.74	124.64	428.84	0.29	-1.78	1.13E-24	3.95E-22
YKL071W	YKL071W	28.18	13.13	43.24	0.30	-1.72	4.35E-06	0.000143609
MET2	YNL277W	92.05	43.10	140.99	0.31	-1.71	7.53E-05	0.001426561
YLR346C	YLR346C	222.03	104.55	339.51	0.31	-1.70	0.000676131	0.007586305
HOP1	YIL072W	41.38	19.64	63.13	0.31	-1.68	1.31E-07	6.82E-06
BDH2	YAL061W	151.22	71.77	230.67	0.31	-1.68	8.13E-08	4.78E-06
TDH2	YJR009C	4415.96	2132.73	6699.19	0.32	-1.65	8.75E-35	6.13E-32
YRO2	YBR054W	68.94	33.32	104.55	0.32	-1.65	4.14E-06	0.000138833
MDH2	YOL126C	168.25	81.99	254.51	0.32	-1.63	1.48E-12	2.22E-10
FCY22	YER060W-A	28.52	14.16	42.87	0.33	-1.60	0.003244093	0.024509631
YOR293C-A	YOR293C-A	12.95	6.52	19.39	0.34	-1.57	0.004534417	0.031431638
RPI1	YIL119C	220.54	111.38	329.69	0.34	-1.57	8.79E-13	1.39E-10
SDP1	YIL113W	63.77	32.64	94.90	0.34	-1.54	1.37E-08	9.91E-07
SSE2	YBR169C	172.96	88.78	257.13	0.35	-1.53	1.35E-15	2.93E-13
RRT5	YFR032C	21.58	11.26	31.90	0.35	-1.50	0.000399395	0.005094305
ARN2	YHL047C	617.48	324.49	910.47	0.36	-1.49	0.001101532	0.011117995
ICS2	YBR157C	389.10	205.15	573.04	0.36	-1.48	1.02E-13	1.89E-11
ENO1	YGR254W	1377.78	742.45	2013.11	0.37	-1.44	9.73E-25	3.61E-22
YGR107W	YGR107W	20.97	11.30	30.64	0.37	-1.44	0.001690395	0.015237734
URA1	YKL216W	3976.00	2142.79	5809.22	0.37	-1.44	2.77E-26	1.16E-23
SPO74	YGL170C	15.75	8.51	22.99	0.37	-1.43	0.006974231	0.042499639
FIT1	YDR534C	47.07	25.52	68.61	0.37	-1.43	0.000296233	0.004067593
FIT2	YOR382W	371.22	204.94	537.51	0.38	-1.39	0.00439113	0.030736278
FIG1	YBR040W	20.45	11.30	29.60	0.38	-1.39	0.001455678	0.013725733
SPS4	YOR313C	202.60	112.33	292.86	0.38	-1.38	1.73E-08	1.20E-06

YCL042W	YCL042W	35.46	19.77	51.14	0.39	-1.37	0.006155847	0.038897992
PHR1	YOR386W	163.29	93.07	233.52	0.40	-1.33	6.88E-12	9.23E-10
FIT3	YOR383C	349.21	200.58	497.84	0.40	-1.31	0.003748134	0.027302882
SET4	YJL105W	29.55	17.00	42.11	0.40	-1.31	0.000210451	0.003149761
RSB1	YOR049C	90.37	52.38	128.37	0.41	-1.29	0.000345062	0.004548611
AAC3	YBR085W	449.59	263.64	635.53	0.41	-1.27	3.99E-12	5.58E-10
EMI2	YDR516C	352.19	206.97	497.41	0.42	-1.26	0.004246211	0.029961227
YMR317W	YMR317W	56.47	33.24	79.70	0.42	-1.26	4.53E-06	0.000148613
DCS2	YOR173W	44.43	26.19	62.68	0.42	-1.26	7.16E-05	0.00136261
DUR1,2	YBR208C	470.25	277.33	663.16	0.42	-1.26	4.53E-11	5.31E-09
YHR022C	YHR022C	89.99	53.10	126.88	0.42	-1.26	0.000288229	0.004000292
VRP1	YLR337C	121.17	71.94	170.39	0.42	-1.24	4.23E-09	3.33E-07
RPS14B	YJL191W	1118.35	668.23	1568.47	0.43	-1.23	1.23E-10	1.29E-08
YKL151C	YKL151C	377.80	225.93	529.67	0.43	-1.23	0.00044516	0.005499907
TDH1	YJL052W	227.97	136.37	319.57	0.43	-1.23	2.35E-12	3.45E-10
YHR140W	YHR140W	76.41	46.37	106.44	0.44	-1.20	1.79E-06	6.59E-05
ASG7	YJL170C	64.31	39.10	89.53	0.44	-1.20	7.06E-06	0.000212985
COA4	YLR218C	105.46	64.42	146.51	0.44	-1.19	0.000436311	0.005433193
YGR053C	YGR053C	89.87	54.91	124.84	0.44	-1.18	0.001547787	0.014321003
FET4	YMR319C	2492.38	1533.82	3450.95	0.44	-1.17	8.08E-18	2.12E-15
GAD1	YMR250W	310.21	191.06	429.36	0.44	-1.17	1.29E-07	6.75E-06
XYL2	YLR070C	31.17	19.27	43.06	0.45	-1.16	0.001354203	0.013116546
PHO5	YBR093C	582.53	364.92	800.14	0.46	-1.13	1.02E-07	5.70E-06
ECM4	YKR076W	47.88	30.00	65.76	0.46	-1.13	0.000129325	0.002178812
BNA2	YJR078W	29.63	18.64	40.61	0.46	-1.12	0.006236283	0.039216386
DDR48	YMR173W	74.87	47.29	102.45	0.46	-1.12	3.89E-06	0.000132655
HSP78	YDR258C	732.26	463.34	1001.18	0.46	-1.11	2.15E-10	2.18E-08
SDS24	YBR214W	123.95	78.95	168.96	0.47	-1.10	1.19E-07	6.33E-06
YBL029W	YBL029W	151.85	96.97	206.73	0.47	-1.09	2.53E-08	1.69E-06
PET117	YER058W	124.66	79.92	169.40	0.47	-1.08	2.73E-07	1.23E-05
ZRT1	YGL255W	302.41	193.96	410.86	0.47	-1.08	7.91E-05	0.001482755
PXA1	YPL147W	286.32	183.80	388.83	0.47	-1.08	1.08E-10	1.16E-08
PEX30	YLR324W	192.82	124.15	261.50	0.47	-1.07	1.10E-08	8.05E-07

RPL18B	YNL301C	2027.13	1310.31	2743.95	0.48	-1.07	4.47E-15	8.79E-13
GLK1	YCL040W	550.73	356.69	744.77	0.48	-1.06	0.000918141	0.009546541
YJL163C	YJL163C	116.45	75.52	157.39	0.48	-1.06	9.00E-07	3.59E-05
PDR15	YDR406W	295.90	192.18	399.62	0.48	-1.06	1.47E-05	0.00037112
YBR012C	YBR012C	514.63	334.97	694.29	0.48	-1.05	0.000214186	0.003198068
ATG40	YOR152C	118.89	77.45	160.33	0.48	-1.05	4.57E-07	1.96E-05
URA8	YJR103W	732.61	477.83	987.40	0.48	-1.05	4.21E-13	7.38E-11
STF1	YDL130W-A	84.19	55.04	113.35	0.49	-1.04	0.000338862	0.004514107
GIP1	YBR045C	108.93	71.40	146.47	0.49	-1.04	8.39E-06	0.000241401
HSP82	YPL240C	2111.95	1388.90	2835.01	0.49	-1.03	9.45E-08	5.37E-06
AIM2	YAL049C	477.96	314.70	641.22	0.49	-1.03	6.41E-07	2.64E-05
YTP1	YNL237W	31.97	21.11	42.83	0.49	-1.02	0.006378513	0.039911632
TYE7	YOR344C	296.88	196.18	397.58	0.49	-1.02	1.81E-05	0.000447617
TSL1	YML100W	254.01	167.85	340.17	0.49	-1.02	0.001839416	0.016164802
YNL092W	YNL092W	38.90	25.86	51.93	0.50	-1.01	0.002821016	0.022163615
FDC1	YDR539W	528.48	351.75	705.21	0.50	-1.00	8.33E-12	1.09E-09
PDR5	YOR153W	12722.76	8468.13	16977.39	0.50	-1.00	0.00084141	0.008925462
HSP104	YLL026W	1807.52	1203.19	2411.85	0.50	-1.00	9.85E-13	1.51E-10
AIM10	YER087W	376.14	250.38	501.89	0.50	-1.00	2.40E-07	1.13E-05

(3) Functional annotation clustering of GO terms of genes showing decrease in *loc1Δ*

[a] Cluster is a group of terms having similar biological meaning due to sharing similar gene members.

[b] Enrichment score is the overall enrichment score for the group based on the P-values of each term members. The higher, the more enriched.

[c] Category shows original database/resource where the terms orient. CC, cellular component. MF, molecular function. BP, biological processes

Annotation   Enrichment Score:
Cluster 1   5.545979739525703

Category	Term	Count	%	PValue	Genes	List Total	Pop Hits	Pop Total	Fold Enrichment	Bonferroni	Benjamini	FDR
GOTERM_BP_FAT	GO:0009266~response to temperature stimulus	17	17.89	1.15E-07	PDR15, YKL151C, SDS24, YRO2, ECM4, GLK1, TSL1, YCL042W, WSC4, HSP78, DDR48, SSE2, BDH2, YGP1, GAD1, YTP1, HSP104	76	217	4870	5.02	6.64E-05	6.64E-05	1.69E-04
GOTERM_BP_FAT	GO:0009408~response to heat	16	16.84	2.97E-07	PDR15, YKL151C, SDS24, YRO2, ECM4, GLK1, TSL1, YCL042W, WSC4, HSP78, DDR48, SSE2, YGP1, GAD1, YTP1, HSP104	76	203	4870	5.05	1.72E-04	8.59E-05	4.36E-04
GOTERM_BP_FAT	GO:0034605~cellular response to heat	15	15.79	4.47E-07	PDR15, YKL151C, YRO2, SDS24, ECM4, GLK1, TSL1, YCL042W, HSP78, DDR48, SSE2, YGP1, GAD1, YTP1, HSP104	76	181	4870	5.31	2.58E-04	8.62E-05	6.56E-04
GOTERM_BP_FAT	GO:0009628~response to abiotic stimulus	19	20.00	2.84E-06	PDR15, YKL151C, SDS24, VRP1, YRO2, ECM4, GLK1, HSP82, TSL1, YCL042W, WSC4, HSP78, DDR48, SSE2, BDH2, GAD1, YGP1, YTP1, HSP104	76	344	4870	3.54	0.0016421	4.11E-04	0.004172
GOTERM_BP_FAT	GO:0033554~cellular response to stress	18	18.95	0.004319	PDR15, YKL151C, SDS24, YRO2, MSH5, ECM4, GLK1, PHR1, TSL1, YCL042W, HSP78, DDR48, SSE2, PHO5, YGP1, GAD1, YTP1, HSP104	76	562	4870	2.05	0.9184019	0.1177656	6.162961

Annotation   Enrichment Score:
Cluster 2   5.489467661227756

Category	Term	Count	%	PValue	Genes	List Total	Pop Hits	Pop Total	Fold Enrichment	Bonferroni	Benjamini	FDR
GOTERM_CC_FAT	GO:0009277~fungal-type cell wall	10	10.53	2.09E-06	TDH1, FIG1, DAN1, FIT3, TIR1, FIT1, TDH2, FIT2, PHO5, YGP1	57	98	4595	8.23	2.44E-04	2.44E-04	0.002383
GOTERM_CC_FAT	GO:0005618~cell wall	10	10.53	4.04E-06	TDH1, FIG1, DAN1, FIT3, TIR1, FIT1, TDH2, FIT2, PHO5, YGP1	57	106	4595	7.61	4.72E-04	2.36E-04	0.004614
GOTERM_CC_FAT	GO:0030312~external encapsulating structure	10	10.53	4.04E-06	TDH1, FIG1, DAN1, FIT3, TIR1, FIT1, TDH2, FIT2, PHO5, YGP1	57	106	4595	7.61	4.72E-04	2.36E-04	0.004614

Annotation   Enrichment Score:
Cluster 3   2.643698820042149

Category	Term	Count	%	PValue	Genes	List Total	Pop Hits	Pop Total	Fold Enrichment	Bonferroni	Benjamini	FDR
GOTERM_BP_FAT	GO:0006096~glycolysis	6	6.32	7.99E-05	TDH1, TDH2, EMI2, GLK1, MDH2, ENO1	76	30	4870	12.82	0.0452155	0.0092112	0.117381
SP_PIR_KEYWORDS	glycolysis	5	5.26	4.95E-04	TDH1, TDH2, EMI2, GLK1, ENO1	95	26	6448	13.05	0.060044	0.060044	0.571223
GOTERM_BP_FAT	GO:0006007~glucose catabolic process	6	6.32	9.35E-04	TDH1, TDH2, EMI2, GLK1, MDH2, ENO1	76	50	4870	7.69	0.41816	0.0744487	1.365279
GOTERM_BP_FAT	GO:0044275~cellular carbohydrate catabolic process	7	7.37	0.001015	TDH1, TDH2, PGU1, EMI2, GLK1, MDH2, ENO1	76	76	4870	5.90	0.4445566	0.0708625	1.481455
GOTERM_BP_FAT	GO:0019320~hexose catabolic process	6	6.32	0.001569	TDH1, TDH2, EMI2, GLK1, MDH2, ENO1	76	56	4870	6.87	0.5971109	0.079322	2.281204
GOTERM_BP_FAT	GO:0016052~carbohydrate catabolic process	7	7.37	0.001713	TDH1, TDH2, PGU1, EMI2, GLK1, MDH2, ENO1	76	84	4870	5.34	0.6293293	0.0793759	2.487726
GOTERM_BP_FAT	GO:0046365~monosaccharide catabolic process	6	6.32	0.002139	TDH1, TDH2, EMI2, GLK1, MDH2, ENO1	76	60	4870	6.41	0.7104904	0.0745479	3.097505
GOTERM_BP_FAT	GO:0046164~alcohol catabolic process	6	6.32	0.003047	TDH1, TDH2, EMI2, GLK1, MDH2, ENO1	76	65	4870	5.91	0.8291579	0.0888072	4.386252
GOTERM_BP_FAT	GO:0005996~monosaccharide metabolic process	7	7.37	0.017668	TDH1, TDH2, EMI2, GLK1, MDH2, XYL2, ENO1	76	136	4870	3.30	0.9999671	0.2994546	23.04781
GOTERM_BP_FAT	GO:0006006~glucose metabolic process	6	6.32	0.018339	TDH1, TDH2, EMI2, GLK1, MDH2, ENO1	76	100	4870	3.84	0.9999778	0.2922742	23.81682
GOTERM_BP_FAT	GO:0019318~hexose metabolic process	6	6.32	0.038979	TDH1, TDH2, EMI2, GLK1, MDH2, ENO1	76	122	4870	3.15	1	0.5129543	44.25337

130

Annotation Cluster 4 Enrichment Score: 2.3914501432371957

Category	Term	Count	%	PValue	Genes	List Total	Pop Hits	Pop Total	Fold Enrichment	Bonferroni	Benjamini	FDR
GOTERM_BP_FAT	GO:0015891~siderophore transport	4	4.21	2.76E-04	FIT3, FIT1, FIT2, ARN2	76	9	4870	28.48	0.1475941	0.0262643	0.40454
SP_PIR_KEYWORDS	iron transport	5	5.26	4.95E-04	FIT3, FIT1, FIT2, ARN2, FET4	95	26	6448	13.05	0.060044	0.060044	0.571223
GOTERM_BP_FAT	GO:0006826~iron ion transport	5	5.26	0.001514	FIT3, FIT1, FIT2, ARN2, FET4	76	33	4870	9.71	0.5840822	0.0839889	2.202227
GOTERM_BP_FAT	GO:0000041~transition metal ion transport	6	6.32	0.001838	FIT3, FIT1, FIT2, ZRT1, ARN2, FET4	76	58	4870	6.63	0.6552486	0.0786523	2.666992
GOTERM_BP_FAT	GO:0015674~di-, tri-valent inorganic cation transport	5	5.26	0.00605	FIT3, FIT1, FIT2, ARN2, FET4	76	48	4870	6.67	0.9702137	0.1416728	8.532936
GOTERM_BP_FAT	GO:0030001~metal ion transport	6	6.32	0.010011	FIT3, FIT1, FIT2, ZRT1, ARN2, FET4	76	86	4870	4.47	0.9970481	0.2007253	13.74556
SP_PIR_KEYWORDS	ion transport	6	6.32	0.016239	FIT3, FIT1, FIT2, ZRT1, ARN2, FET4	95	102	6448	3.99	0.8708254	0.2033971	17.24916
SP_PIR_KEYWORDS	iron	6	6.32	0.01959	FIT3, FIT1, FIT2, BNA2, ARN2, FET4	95	107	6448	3.81	0.9156689	0.2190939	20.45003
GOTERM_MF_FAT	GO:0005506~iron ion binding	6	6.32	0.040962	FIT3, FIT1, FIT2, BNA2, ARN2, FET4	65	125	4190	3.09	0.99881	0.99881	39.68351

Annotation Cluster 5 Enrichment Score: 2.281823448316349

Category	Term	Count	%	PValue	Genes	List Total	Pop Hits	Pop Total	Fold Enrichment	Bonferroni	Benjamini	FDR
GOTERM_BP_FAT	GO:0046364~monosaccharide biosynthetic process	5	5.26	0.00105	TDH1, TDH2, MDH2, XYL2, ENO1	76	30	4870	10.68	0.4556964	0.06535	1.532107
GOTERM_BP_FAT	GO:0006094~gluconeogenesis	4	4.21	0.002847	TDH1, TDH2, MDH2, ENO1	76	19	4870	13.49	0.8081357	0.0876398	4.104182
GOTERM_BP_FAT	GO:0019319~hexose biosynthetic process	4	4.21	0.004378	TDH1, TDH2, MDH2, ENO1	76	22	4870	11.65	0.9211559	0.1139358	6.244706
GOTERM_BP_FAT	GO:0034637~cellular carbohydrate biosynthetic process	6	6.32	0.006672	TDH1, TSL1, TDH2, MDH2, XYL2, ENO1	76	78	4870	4.93	0.9792709	0.1491411	9.370742
GOTERM_BP_FAT	GO:0046165~alcohol biosynthetic process	5	5.26	0.008038	TDH1, TDH2, MDH2, XYL2, ENO1	76	52	4870	6.16	0.990653	0.1704798	11.18466
GOTERM_BP_FAT	GO:0006090~pyruvate metabolic process	4	4.21	0.010569	TDH1, TDH2, MDH2, ENO1	76	30	4870	8.54	0.9978706	0.203756	14.45773
GOTERM_BP_FAT	GO:0016051~carbohydrate biosynthetic process	6	6.32	0.014352	TDH1, TSL1, TDH2, MDH2, XYL2, ENO1	76	94	4870	4.09	0.9997684	0.2583958	19.14156

Annotation Cluster 6 Enrichment Score: 1.8294610954096568

Category	Term	Count	%	PValue	Genes	List Total	Pop Hits	Pop Total	Fold Enrichment	Bonferroni	Benjamini	FDR
SP_PIR_KEYWORDS	gpi-anchor	5	5.26	0.010461	DAN1, FIT3, TIR1, FIT1, FIT2	95	59	6448	5.75	0.7313897	0.171207	11.45043
GOTERM_CC_FAT	GO:0031225~anchored to membrane	5	5.26	0.010501	DAN1, FIT3, TIR1, FIT1, FIT2	57	71	4595	5.68	0.7092101	0.2188834	11.36382
SP_PIR_KEYWORDS	cell wall	5	5.26	0.014569	DAN1, FIT3, TIR1, FIT1, FIT2	95	65	6448	5.22	0.8403035	0.204918	15.60931
UP_SEQ_FEATURE	propeptide:Removed in mature form	5	5.26	0.030055	DAN1, FIT3, TIR1, FIT1, FIT2	95	81	6448	4.19	0.9995285	0.9221682	32.7455

Annotation Cluster 7 Enrichment Score: 1.673856650339587

Category	Term	Count	%	PValue	Genes	List Total	Pop Hits	Pop Total	Fold Enrichment	Bonferroni	Benjamini	FDR
GOTERM_CC_FAT	GO:0022626~cytosolic ribosome	7	7.37	0.007882	RPL9B, RPL8A, RPL18B, ANB1, RPS14B, RPL7B, RPS22B	57	145	4595	3.89	0.6037829	0.2066162	8.644948
GOTERM_CC_FAT	GO:0044445~cytosolic part	7	7.37	0.02691	RPL9B, RPL8A, RPL18B, RPS14B, RPL7B, RPS22B, ENO1	57	190	4595	2.97	0.9588967	0.3661559	26.78067
KEGG_PATHWAY	sce03010:Ribosome	6	6.32	0.044864	RPL9B, RPL8A, RPL18B, RPS14B, RPL7B, RPS22B	25	121	1439	2.85	0.6520633	0.2319755	30.49921

(4) Functional annotation clustering of GO terms of genes showing increase in *loc1Δ*

[a] Cluster is a group of terms having similar biological meaning due to sharing similar gene members. [b] Enrichment score is the overall enrichment score for the group based on the P-values of each term members. The higher, the more enriched.

<superscript>c</superscript> Category shows original database/resource where the terms orient. CC, cellular component. MF, molecular function. BP, biological processes

Annotation Cluster 1 Enrichment Score: 4.595957892361109

Category	Term	Count	%	PValue	Genes	List Total	Pop Hits	Pop Total	Fold Enrichment	Bonferroni	Benjamini	FDR
INTERPRO	IPR000883:Cytochrome c oxidase, subunit I	5	7.94	6.29E-07	AI5_ALPHA, COX1, AI4, AI3, AI2	57	7	4676	58.60	8.81E-05	8.81E-05	7.42E-04
GOTERM_MF_FAT	GO:0046906~tetrapyrrole binding	6	9.52	2.55E-05	AI5_ALPHA, COX2, COX1, AI4, AI3, AI2	49	32	4190	16.03	0.00391417	0.001959	0.030545
GOTERM_MF_FAT	GO:0020037~heme binding	6	9.52	2.55E-05	AI5_ALPHA, COX2, COX1, AI4, AI3, AI2	49	32	4190	16.03	0.00391417	0.001959	0.030545
GOTERM_BP_FAT	GO:0009060~aerobic respiration	7	11.11	0.001013	AI5_ALPHA, COX2, COX1, COB, AI4, AI3, AI2	57	102	4870	5.86	0.35058835	0.069421	1.416255

Annotation Cluster 2 Enrichment Score: 3.998909694708192

Category	Term	Count	%	PValue	Genes	List Total	Pop Hits	Pop Total	Fold Enrichment	Bonferroni	Benjamini	FDR
GOTERM_MF_FAT	GO:0009055~electron carrier activity	9	14.29	2.52E-06	BI3, LYS1, AI5_ALPHA, COX2, COX1, COB, AI4, AI3, AI2	49	80	4190	9.62	3.88E-04	3.88E-04	0.003025
GOTERM_BP_FAT	GO:0045333~cellular respiration	8	12.70	3.94E-04	BI3, AI5_ALPHA, COX2, COX1, COB, AI4, AI3, AI2	57	120	4870	5.70	0.15471888	0.054488	0.553848
GOTERM_BP_FAT	GO:0009060~aerobic respiration	7	11.11	0.001013	AI5_ALPHA, COX2, COX1, COB, AI4, AI3, AI2	57	102	4870	5.86	0.35058835	0.069421	1.416255

Annotation Cluster 3 Enrichment Score: 3.11067309032934

Category	Term	Count	%	PValue	Genes	List Total	Pop Hits	Pop Total	Fold Enrichment	Bonferroni	Benjamini	FDR
SP_PIR_KEYWORDS	amino-acid biosynthesis	8	12.70	3.15E-05	LYS1, ARG1, ARG4, CPA2, YHR033W, ARO1, LEU4, ARO3	62	97	6448	8.58	0.00468384	0.004684	0.037576
GOTERM_BP_FAT	GO:0008652~cellular amino acid biosynthetic process	8	12.70	7.69E-04	LYS1, ARG1, ARG4, CPA2, YHR033W, ARO1, LEU4, ARO3	57	134	4870	5.10	0.27933217	0.063415	1.076539
GOTERM_BP_FAT	GO:0009309~amine biosynthetic process	8	12.70	0.001084	LYS1, ARG1, ARG4, CPA2, YHR033W, ARO1, LEU4, ARO3	57	142	4870	4.81	0.37006784	0.063888	1.515408
GOTERM_BP_FAT	GO:0016053~organic acid biosynthetic process	8	12.70	0.003263	LYS1, ARG1, ARG4, CPA2, YHR033W, ARO1, LEU4, ARO3	57	172	4870	3.97	0.75154121	0.129988	4.496767
GOTERM_BP_FAT	GO:0046394~carboxylic acid biosynthetic process	8	12.70	0.003263	LYS1, ARG1, ARG4, CPA2, YHR033W, ARO1, LEU4, ARO3	57	172	4870	3.97	0.75154121	0.129988	4.496767

Annotation Cluster 4 Enrichment Score: 2.595525223581457

Category	Term	Count	%	PValue	Genes	List Total	Pop Hits	Pop Total	Fold Enrichment	Bonferroni	Benjamini	FDR
SP_PIR_KEYWORDS	gpi-anchor	6	9.52	2.17E-04	MUC1, DAN4, FIG2, FLO5, GAS2, HPF1	62	59	6448	10.58	0.03182972	0.016044	0.258614
GOTERM_CC_FAT	GO:0031225~anchored to membrane	6	9.52	5.20E-04	MUC1, DAN4, FIG2, FLO5, GAS2, HPF1	45	71	4595	8.63	0.03572523	0.035725	0.536401
SP_PIR_KEYWORDS	lipoprotein	7	11.11	0.001236	MUC1, DAN4, FIG2, COX1, FLO5, GAS2, HPF1	62	128	6448	5.69	0.16833033	0.059591	1.464697
GOTERM_CC_FAT	GO:0030312~external encapsulating structure	6	9.52	0.003156	MUC1, DIA3, DAN4, FIG2, FLO5, HPF1	45	106	4595	5.78	0.19849597	0.104732	3.218367
GOTERM_CC_FAT	GO:0005618~cell wall	6	9.52	0.003156	MUC1, DIA3, DAN4, FIG2, FLO5, HPF1	45	106	4595	5.78	0.19849597	0.104732	3.218367
SP_PIR_KEYWORDS	cell wall	5	7.94	0.003191	MUC1, DAN4, FIG2, FLO5, HPF1	62	65	6448	8.00	0.37890053	0.051542	3.740853
SP_PIR_KEYWORDS	Secreted	5	7.94	0.006711	MUC1, DAN4, FIG2, FLO5, HPF1	62	80	6448	6.50	0.63332986	0.080208	7.71755
GOTERM_CC_FAT	GO:0005576~extracellular region	5	7.94	0.010679	MUC1, DAN4, FIG2, FLO5, HPF1	45	91	4595	5.61	0.52836512	0.139559	10.51623
GOTERM_CC_FAT	GO:0009277~fungal-type cell wall	5	7.94	0.013752	DIA3, DAN4, FIG2, FLO5, HPF1	45	98	4595	5.21	0.62066693	0.149182	13.35164

Annotation Cluster 5 Enrichment Score: 2.0299487358484716

Category	Term	Count	%	PValue	Genes	List Total	Pop Hits	Pop Total	Fold Enrichment	Bonferroni	Benjamini	FDR
SP_PIR_KEYWORDS	arginine biosynthesis	3	4.76	0.003037	ARG1, ARG4, CPA2	62	9	6448	34.67	0.36440033	0.055074	3.562857
GOTERM_BP_FAT	GO:0009084~glutamine family amino acid biosynthetic process	4	6.35	0.003849	ARG1, ARG4, CPA2, YHR033W	57	28	4870	12.21	0.80657469	0.138733	5.283626
GOTERM_BP_FAT	GO:0006526~arginine biosynthetic process	3	4.76	0.005509	ARG1, ARG4, CPA2	57	10	4870	25.63	0.9049489	0.154727	7.481242
GOTERM_BP_FAT	GO:0006525~arginine metabolic process	3	4.76	0.014056	ARG1, ARG4, CPA2	57	16	4870	16.02	0.99759538	0.298635	18.06586
KEGG_PATHWAY	sce00250:Alanine, aspartate and glutamate metabolism	3	4.76	0.078258	ARG1, ARG4, CPA2	27	26	1439	6.15	0.76934408	0.519733	45.15641

Annotation Enrichment Score:
Cluster 6   1.931823491999084

Category	Term	Count	%	PValue	Genes	List Total	Pop Hits	Pop Total	Fold Enrichment	Bonferroni	Benjamini	FDR
SP_PIR_KEYWORDS	respiratory chain	4	6.35	0.002005	AI5_ALPHA, COX2, COX1, COB	62	27	6448	15.41	0.25845198	0.048614	2.365275
GOTERM_CC_FAT	GO:0070469~respiratory chain	4	6.35	0.003348	AI5_ALPHA, COX2, COX1, COB	45	32	4595	12.76	0.20923828	0.075269	3.411246
GOTERM_BP_FAT	GO:0022900~electron transport chain	5	7.94	0.007403	BI3, AI5_ALPHA, COX2, COX1, COB	57	68	4870	6.28	0.95779972	0.179493	9.93048
GOTERM_MF_FAT	GO:0004129~cytochrome-c oxidase activity	3	4.76	0.012272	AI5_ALPHA, COX2, COX1	49	15	4190	17.10	0.8506561	0.146542	13.76722
GOTERM_MF_FAT	GO:0016676~oxidoreductase activity, acting on heme group of donors, oxygen as acceptor	3	4.76	0.012272	AI5_ALPHA, COX2, COX1	49	15	4190	17.10	0.8506561	0.146542	13.76722
GOTERM_MF_FAT	GO:0016675~oxidoreductase activity, acting on heme group of donors	3	4.76	0.012272	AI5_ALPHA, COX2, COX1	49	15	4190	17.10	0.8506561	0.146542	13.76722
GOTERM_MF_FAT	GO:0015002~heme-copper terminal oxidase activity	3	4.76	0.012272	AI5_ALPHA, COX2, COX1	49	15	4190	17.10	0.8506561	0.146542	13.76722
SP_PIR_KEYWORDS	electron transport	4	6.35	0.013494	AI5_ALPHA, COX2, COX1, COB	62	53	6448	7.85	0.8679151	0.144198	14.96025
SP_PIR_KEYWORDS	copper	3	4.76	0.028422	AI5_ALPHA, COX2, COX1	62	28	6448	11.14	0.98637968	0.223312	29.10154
SP_PIR_KEYWORDS	heme	3	4.76	0.030344	AI5_ALPHA, COX1, COB	62	29	6448	10.76	0.98986015	0.225139	30.75673
GOTERM_MF_FAT	GO:0005507~copper ion binding	3	4.76	0.042876	AI5_ALPHA, COX2, COX1	49	29	4190	8.85	0.99882757	0.362316	40.88564

Annotation Enrichment Score:
Cluster 7   1.886005143232497

Category	Term	Count	%	PValue	Genes	List Total	Pop Hits	Pop Total	Fold Enrichment	Bonferroni	Benjamini	FDR
GOTERM_MF_FAT	GO:0043169~cation binding	18	28.57	0.010338	AI4, GLC3, AI3, AI2, MAL33, SIP4, AI5_ALPHA, ISU2, MSN4, COX2, COX1, COB, CPA2, ARO10, YAR009C, PDH1, IZH4, YLR460C	49	850	4190	1.81	0.79817222	0.162903	11.72033
GOTERM_MF_FAT	GO:0043167~ion binding	18	28.57	0.010461	AI4, GLC3, AI3, AI2, MAL33, SIP4, AI5_ALPHA, ISU2, MSN4, COX2, COX1, COB, CPA2, ARO10, YAR009C, PDH1, IZH4, YLR460C	49	851	4190	1.81	0.80199788	0.149514	11.85183
GOTERM_MF_FAT	GO:0046914~transition metal ion binding	15	23.81	0.014126	AI4, AI3, AI2, MAL33, SIP4, ISU2, AI5_ALPHA, MSN4, COX2, COX1, COB, CPA2, PDH1, YAR009C, YLR460C	49	665	4190	1.93	0.88818935	0.155098	15.68973
GOTERM_MF_FAT	GO:0046872~metal ion binding	17	26.98	0.018704	AI4, AI3, AI2, MAL33, SIP4, AI5_ALPHA, ISU2, MSN4, COX2, COX1, COB, CPA2, ARO10, PDH1, YAR009C, IZH4, YLR460C	49	830	4190	1.75	0.945402	0.187547	20.26832

Annotation Enrichment Score:
Cluster 8   1.8583533314467826

Category	Term	Count	%	PValue	Genes	List Total	Pop Hits	Pop Total	Fold Enrichment	Bonferroni	Benjamini	FDR
SP_PIR_KEYWORDS	Intron homing	3	4.76	0.001289	AI5_ALPHA, AI4, AI3	62	6	6448	52.00	0.17481446	0.046901	1.526418
SP_PIR_KEYWORDS	Endonuclease	4	6.35	0.03249	AI5_ALPHA, AI4, AI3, YAR009C	62	74	6448	5.62	0.99271076	0.228191	32.5624
SP_PIR_KEYWORDS	nuclease	4	6.35	0.063537	AI5_ALPHA, AI4, AI3, YAR009C	62	97	6448	4.29	0.99994349	0.346405	54.29699

Annotation Enrichment Score:
Cluster 9   1.7027310890850527

Category	Term	Count	%	PValue	Genes	List Total	Pop Hits	Pop Total	Fold Enrichment	Bonferroni	Benjamini	FDR
GOTERM_BP_FAT	GO:0006314~intron homing	5	7.94	3.13E-06	BI3, AI5_ALPHA, AI4, Q0255, AI3	57	10	4870	42.72	0.00133235	0.001332	0.004405
SP_PIR_KEYWORDS	Intron homing	3	4.76	0.001289	AI5_ALPHA, AI4, AI3	62	6	6448	52.00	0.17481446	0.046901	1.526418
SP_PIR_KEYWORDS	mrna splicing	4	6.35	0.074971	BI3, AI5_ALPHA, AI4, AI3	62	104	6448	4.00	0.99999094	0.371529	60.52649
GOTERM_BP_FAT	GO:0008380~RNA splicing	5	7.94	0.079387	BI3, AI5_ALPHA, AI4, AI3, AI2	57	142	4870	3.01	1	0.783907	68.7858
SP_PIR_KEYWORDS	mrna processing	4	6.35	0.197235	BI3, AI5_ALPHA, AI4, AI3	62	162	6448	2.57	1	0.664167	92.72307
GOTERM_BP_FAT	GO:0006397~mRNA processing	4	6.35	0.418474	BI3, AI5_ALPHA, AI4, AI3	57	204	4870	1.68	1	0.998053	99.95146
GOTERM_BP_FAT	GO:0016071~mRNA metabolic process	4	6.35	0.608024	BI3, AI5_ALPHA, AI4, AI3	57	270	4870	1.27	1	0.999953	99.99981

Annotation Enrichment Score:
Cluster 10   1.518362402712603

Category	Term	Count	%	PValue	Genes	List Total	Pop Hits	Pop Total	Fold Enrichment	Bonferroni	Benjamini	FDR
SP_PIR_KEYWORDS	oxidative phosphorylation	4	6.35	0.00223	COX2, COX1, COB, ATP6	62	28	6448	14.86	0.28300963	0.046416	2.628141
SP_PIR_KEYWORDS	mitochondrial inner membrane	4	6.35	0.004252	COX2, COX1, COB, ATP6	62	35	6448	11.89	0.46997554	0.06151	4.955024
GOTERM_BP_FAT	GO:0022904~respiratory electron transport chain	4	6.35	0.005642	BI3, COX2, COX1, COB	57	32	4870	10.68	0.9101945	0.14843	7.65462
GOTERM_MF_FAT	GO:0015078~hydrogen ion transmembrane transporter activity	5	7.94	0.006496	AI5_ALPHA, COX2, COX1, COB, ATP6	49	66	4190	6.48	0.63348203	0.133581	7.520666
GOTERM_MF_FAT	GO:0015077~monovalent inorganic cation transmembrane transporter activity	5	7.94	0.0084	AI5_ALPHA, COX2, COX1, COB, ATP6	49	71	4190	6.02	0.72722144	0.149888	9.624278
GOTERM_MF_FAT	GO:0022890~inorganic cation transmembrane transporter activity	6	9.52	0.010599	AI5_ALPHA, COX2, COX1, ATO3, COB, ATP6	49	118	4190	4.35	0.80619074	0.138578	11.99868
SP_PIR_KEYWORDS	electron transfer	3	4.76	0.014992	COX2, COX1, COB	62	20	6448	15.60	0.8946845	0.148513	16.48818
GOTERM_CC_FAT	GO:0005746~mitochondrial respiratory chain	3	4.76	0.02703	COX2, COX1, COB	45	27	4595	11.35	0.85312415	0.174542	24.69257
GOTERM_BP_FAT	GO:0006119~oxidative phosphorylation	4	6.35	0.029747	COX2, COX1, COB, ATP6	57	59	4870	5.79	0.99999741	0.491901	34.62743
SP_PIR_KEYWORDS	membrane-associated complex	3	4.76	0.036409	COX2, COX1, ATP6	62	32	6448	9.75	0.99601856	0.231372	35.74944
GOTERM_BP_FAT	GO:0042775~mitochondrial ATP synthesis coupled electron transport	3	4.76	0.04056	COX2, COX1, COB	57	28	4870	9.15	0.99999998	0.586016	44.16734
GOTERM_BP_FAT	GO:0042773~ATP synthesis coupled electron transport	3	4.76	0.04056	COX2, COX1, COB	57	28	4870	9.15	0.99999998	0.586016	44.16734
GOTERM_CC_FAT	GO:0044455~mitochondrial membrane part	4	6.35	0.122828	COX2, COX1, COB, ATP6	45	128	4595	3.19	0.99989626	0.480693	74.23814
GOTERM_BP_FAT	GO:0016310~phosphorylation	4	6.35	0.440177	COX2, COX1, COB, ATP6	57	211	4870	1.62	1	0.998502	99.97158
GOTERM_BP_FAT	GO:0006796~phosphate metabolic process	4	6.35	0.642847	COX2, COX1, COB, ATP6	57	284	4870	1.20	1	0.999971	99.99995
GOTERM_BP_FAT	GO:0006793~phosphorus metabolic process	4	6.35	0.654764	COX2, COX1, COB, ATP6	57	289	4870	1.18	1	0.999966	99.99997

Annotation Enrichment Score:
Cluster 11   1.2681781618854904

Category	Term	Count	%	PValue	Genes	List Total	Pop Hits	Pop Total	Fold Enrichment	Bonferroni	Benjamini	FDR
SP_PIR_KEYWORDS	mitochondrion inner membrane	7	11.11	0.006594	BI3, AI5_ALPHA, COX2, CRC1, COX1, COB, ATP6	62	179	6448	4.07	0.62684304	0.085716	7.587909
GOTERM_CC_FAT	GO:0005743~mitochondrial inner membrane	7	11.11	0.01625	BI3, AI5_ALPHA, COX2, CRC1, COX1, COB, ATP6	45	217	4595	3.29	0.68236864	0.15112	15.59622
GOTERM_CC_FAT	GO:0019866~organelle inner membrane	7	11.11	0.01989	BI3, AI5_ALPHA, COX2, CRC1, COX1, COB, ATP6	45	227	4595	3.15	0.75495408	0.14466	18.77245
GOTERM_CC_FAT	GO:0031966~mitochondrial membrane	8	12.70	0.027933	BI3, AI5_ALPHA, COX2, CRC1, COX1, COB, PDH1, ATP6	45	314	4595	2.60	0.86236425	0.164969	25.41255
GOTERM_CC_FAT	GO:0005740~mitochondrial envelope	8	12.70	0.042913	BI3, AI5_ALPHA, COX2, CRC1, COX1, COB, PDH1, ATP6	45	344	4595	2.37	0.95359178	0.225747	36.48697
GOTERM_CC_FAT	GO:0044429~mitochondrial part	9	14.29	0.116362	BI3, AI5_ALPHA, ISU2, COX2, CRC1, COX1, COB, PDH1, ATP6	45	519	4595	1.77	0.99982653	0.4863	72.20363
GOTERM_CC_FAT	GO:0031967~organelle envelope	8	12.70	0.158562	BI3, AI5_ALPHA, COX2, CRC1, COX1, COB, PDH1, ATP6	45	471	4595	1.73	0.99999436	0.553211	83.24899
GOTERM_CC_FAT	GO:0031975~envelope	8	12.70	0.158562	BI3, AI5_ALPHA, COX2, CRC1, COX1, COB, PDH1, ATP6	45	471	4595	1.73	0.99999436	0.553211	83.24899
GOTERM_CC_FAT	GO:0031090~organelle membrane	8	12.70	0.516228	BI3, AI5_ALPHA, COX2, CRC1, COX1, COB, PDH1, ATP6	45	701	4595	1.17	1	0.940625	99.94551

Annotation Enrichment Score:
Cluster 12   0.5730322756514752

Category	Term	Count	%	PValue	Genes	List Total	Pop Hits	Pop Total	Fold Enrichment	Bonferroni	Benjamini	FDR
GOTERM_BP_FAT	GO:0001403~invasive growth in response to glucose limitation	3	4.76	0.112184	MUC1, SIP4, DIA3	57	50	4870	5.13	1	0.86835	81.26734
GOTERM_BP_FAT	GO:0070783~growth of unicellular organism as a thread of attached cells	3	4.76	0.255759	MUC1, SIP4, DIA3	57	85	4870	3.02	1	0.984922	98.43591
GOTERM_BP_FAT	GO:0030447~filamentous growth	3	4.76	0.37477	MUC1, SIP4, DIA3	57	113	4870	2.27	1	0.997217	99.86539
GOTERM_BP_FAT	GO:0040007~growth	3	4.76	0.474616	MUC1, SIP4, DIA3	57	138	4870	1.86	1	0.999115	99.98837

Annotation Enrichment Score:
Cluster 13   0.3469726327106529

Category	Term	Count	%	PValue	Genes	List Total	Pop Hits	Pop Total	Fold Enrichment	Bonferroni	Benjamini	FDR
GOTERM_MF_FAT	GO:0003700~transcription factor activity	4	6.35	0.200262	SIP4, NDT80, MSN4, MAL33	49	135	4190	2.53	1	0.867926	93.14895
SP_PIR_KEYWORDS	dna-binding	5	7.94	0.471308	SIP4, NDT80, MSN4, YAR009C, MAL33	62	372	6448	1.40	1	0.912402	99.95006
GOTERM_MF_FAT	GO:0043565~sequence-specific DNA binding	4	6.35	0.541808	SIP4, NDT80, MSN4, MAL33	49	246	4190	1.39	1	0.998211	99.99141
GOTERM_MF_FAT	GO:0030528~transcription regulator activity	4	6.35	0.800498	SIP4, NDT80, MSN4, MAL33	49	364	4190	0.94	1	0.999996	100

134

Annotation Cluster 14 Enrichment Score: 0.2913308021032751

Category	Term	Count	%	PValue	Genes	List Total	Pop Hits	Pop Total	Fold Enrichment	Bonferroni	Benjamini	FDR
SP_PIR_KEYWORDS	cell wall biogenesis/degradation	3	4.76	0.212603	FIG2, GAS2, HPF1	62	91	6448	3.43	1	0.682999	94.22147
GOTERM_BP_FAT	GO:0007047~cell wall organization	3	4.76	0.792902	FIG2, GAS2, HPF1	57	251	4870	1.02	1	0.999998	100
GOTERM_BP_FAT	GO:0045229~external encapsulating structure organization	3	4.76	0.792902	FIG2, GAS2, HPF1	57	251	4870	1.02	1	0.999998	100

Annotation Cluster 15 Enrichment Score: 0.20137625325768072

Category	Term	Count	%	PValue	Genes	List Total	Pop Hits	Pop Total	Fold Enrichment	Bonferroni	Benjamini	FDR
SP_PIR_KEYWORDS	transcription regulation	6	9.52	0.46689	SIP4, NDT80, MSN4, TOM1, DLS1, MAL33	62	473	6448	1.32	1	0.915114	99.94485
SP_PIR_KEYWORDS	Transcription	6	9.52	0.542248	SIP4, NDT80, MSN4, TOM1, DLS1, MAL33	62	514	6448	1.21	1	0.941564	99.99105
GOTERM_BP_FAT	GO:0045449~regulation of transcription	7	11.11	0.785807	SIP4, NDT80, MSN4, PCL5, TOM1, DLS1, MAL33	57	657	4870	0.91	1	0.999998	100
GOTERM_BP_FAT	GO:0006350~transcription	6	9.52	0.786623	SIP4, NDT80, MSN4, TOM1, DLS1, MAL33	57	559	4870	0.92	1	0.999998	100

Annotation Cluster 16 Enrichment Score: 0.13643072010402085

Category	Term	Count	%	PValue	Genes	List Total	Pop Hits	Pop Total	Fold Enrichment	Bonferroni	Benjamini	FDR
SP_PIR_KEYWORDS	meiosis	3	4.76	0.282869	REC8, NDT80, RIM4	62	111	6448	2.81	1	0.787367	98.10503
GOTERM_BP_FAT	GO:0051327~M phase of meiotic cell cycle	3	4.76	0.63153	REC8, NDT80, RIM4	57	184	4870	1.39	1	0.999969	99.99992
GOTERM_BP_FAT	GO:0007126~meiosis	3	4.76	0.63153	REC8, NDT80, RIM4	57	184	4870	1.39	1	0.999969	99.99992
GOTERM_BP_FAT	GO:0051321~meiotic cell cycle	3	4.76	0.649158	REC8, NDT80, RIM4	57	190	4870	1.35	1	0.999969	99.99996
GOTERM_BP_FAT	GO:0000279~M phase	4	6.35	0.751433	REC8, NDT80, RIM4, TOM1	57	335	4870	1.02	1	0.999997	100
GOTERM_BP_FAT	GO:0022403~cell cycle phase	4	6.35	0.875446	REC8, NDT80, RIM4, TOM1	57	422	4870	0.81	1	1	100
GOTERM_BP_FAT	GO:0007049~cell cycle	5	7.94	0.942157	REC8, NDT80, RIM4, PCL5, TOM1	57	627	4870	0.68	1	1	100
GOTERM_BP_FAT	GO:0022402~cell cycle process	4	6.35	0.952321	REC8, NDT80, RIM4, TOM1	57	530	4870	0.64	1	1	100
GOTERM_CC_FAT	GO:0043232~intracellular non-membrane-bounded organelle	3	4.76	0.999934	REC8, NDT80, TOM1	45	1120	4595	0.27	1	1	100
GOTERM_CC_FAT	GO:0043228~non-membrane-bounded organelle	3	4.76	0.999934	REC8, NDT80, TOM1	45	1120	4595	0.27	1	1	100

Annotation Cluster 17 Enrichment Score: 0.046100558208698406

Category	Term	Count	%	PValue	Genes	List Total	Pop Hits	Pop Total	Fold Enrichment	Bonferroni	Benjamini	FDR
SP_PIR_KEYWORDS	atp-binding	6	9.52	0.729748	ARG1, CPA2, ARO1, MDN1, SLH1, YAR009C	62	634	6448	0.98	1	0.98926	99.99998
SP_PIR_KEYWORDS	nucleotide-binding	6	9.52	0.836749	ARG1, CPA2, ARO1, MDN1, SLH1, YAR009C	62	731	6448	0.85	1	0.996804	100
GOTERM_MF_FAT	GO:0005524~ATP binding	6	9.52	0.887097	ARG1, CPA2, ARO1, MDN1, SLH1, YAR009C	49	651	4190	0.79	1	0.999999	100
GOTERM_MF_FAT	GO:0032559~adenyl ribonucleotide binding	6	9.52	0.888753	ARG1, CPA2, ARO1, MDN1, SLH1, YAR009C	49	653	4190	0.79	1	0.999999	100
GOTERM_MF_FAT	GO:0030554~adenyl nucleotide binding	6	9.52	0.917872	ARG1, CPA2, ARO1, MDN1, SLH1, YAR009C	49	693	4190	0.74	1	1	100
GOTERM_MF_FAT	GO:0001883~purine nucleoside binding	6	9.52	0.919768	ARG1, CPA2, ARO1, MDN1, SLH1, YAR009C	49	696	4190	0.74	1	0.999999	100
GOTERM_MF_FAT	GO:0001882~nucleoside binding	6	9.52	0.922847	ARG1, CPA2, ARO1, MDN1, SLH1, YAR009C	49	701	4190	0.73	1	0.999999	100
GOTERM_MF_FAT	GO:0032555~purine ribonucleotide binding	6	9.52	0.946739	ARG1, CPA2, ARO1, MDN1, SLH1, YAR009C	49	747	4190	0.69	1	1	100
GOTERM_MF_FAT	GO:0032553~ribonucleotide binding	6	9.52	0.946739	ARG1, CPA2, ARO1, MDN1, SLH1, YAR009C	49	747	4190	0.69	1	1	100
GOTERM_MF_FAT	GO:0000166~nucleotide binding	7	11.11	0.961947	ARG1, RIM4, CPA2, ARO1, MDN1, SLH1, YAR009C	49	897	4190	0.67	1	1	100
GOTERM_MF_FAT	GO:0017076~purine nucleotide binding	6	9.52	0.961991	ARG1, CPA2, ARO1, MDN1, SLH1, YAR009C	49	787	4190	0.65	1	1	100

Annotation Cluster 18 Enrichment Score: 0.021271961002659392

Category	Term	Count	%	PValue	Genes	List Total	Pop Hits	Pop Total	Fold Enrichment	Bonferroni	Benjamini	FDR
GOTERM_BP_FAT	GO:0044257~cellular protein catabolic process	3	4.76	0.920081	TOM1, GLC3, HSP12	57	351	4870	0.73	1	1	100
GOTERM_BP_FAT	GO:0030163~protein catabolic process	3	4.76	0.929837	TOM1, GLC3, HSP12	57	364	4870	0.70	1	1	100
GOTERM_BP_FAT	GO:0044265~cellular macromolecule catabolic process	3	4.76	0.978014	TOM1, GLC3, HSP12	57	475	4870	0.54	1	1	100
GOTERM_BP_FAT	GO:0009057~macromolecule catabolic process	3	4.76	0.982504	TOM1, GLC3, HSP12	57	496	4870	0.52	1	1	100

Annotation   Enrichment Score:
Cluster 19   0.013799147943721835

Category	Term	Count	%	PValue	Genes	List Total	Pop Hits	Pop Total	Fold Enrichment	Bonferroni	Benjamini	FDR
GOTERM_BP_FAT	GO:0034470~ncRNA processing	3	4.76	0.906365	ISU2, TOM1, MDN1	57	335	4870	0.77	1	1	100
GOTERM_BP_FAT	GO:0034660~ncRNA metabolic process	3	4.76	0.947763	ISU2, TOM1, MDN1	57	393	4870	0.65	1	1	100
GOTERM_CC_FAT	GO:0043233~organelle lumen	3	4.76	0.997383	ISU2, TOM1, MDN1	45	783	4595	0.39	1	1	100
GOTERM_CC_FAT	GO:0070013~intracellular organelle lumen	3	4.76	0.997383	ISU2, TOM1, MDN1	45	783	4595	0.39	1	1	100
GOTERM_CC_FAT	GO:0031974~membrane-enclosed lumen	3	4.76	0.998338	ISU2, TOM1, MDN1	45	827	4595	0.37	1	1	100

CHAPTER 3

CONCLUSIONS

Saccharomyces cerevisiae is doing a remarkable job at restraining transposable elements in its genome to only 3% of its haploid genome [1]. Instead of using genome defense mechanisms commonly found in other eukaryotes, Ty1 modulates its copy number by a self-encoded restriction factor p22 in *S. cerevisiae* and *S. paradoxus* [2, 3]. This study found ribosome biogenesis, a conserved cellular process of assembling ribosomal RNAs and proteins into a ribosome, modulates CNC. In addition, these findings contributed to our understanding of CNC, and how retroelements utilize host factors to facilitate their movement.

The candidate gene approach utilized in this study was successful at revealing an important cellular process that modulates CNC. Yet, genome-wide screens to identify Ty1 host factors have all used yeast knockout collections that do not include 1105 essential genes [4-7]. About 40% of essential yeast genes are conserved in human [8], emphasizing their fundamental importance. Therefore, surveying essential genes can reveal additional CNC-specific Ty1 host modulators that are directly required for an organism to survive such as genes involved in cell metabolism and cell division/reproduction. There are yeast library collections developed for this purpose. The Decreased Abundance by mRNA Perturbation (DAmP) library is a collection of yeast strains containing compromised alleles of essential genes that exhibit modest

growth defects. This is based on disruption of the 3′ UTR with an antibiotic resistance cassette, which destabilizes the corresponding transcript and can reduce mRNA levels 4-10 fold [9]. Another approach that can be used is Tet-promoter-fusions to essential yeast genes available from the Hughes Collection (yTHC). The endogenous promoters of essential genes have been replaced with a Tet-titratable promoter. This allows the expression of essential genes to be regulated by doxycycline [8].

Ty1 CNC is enhanced when ribosome biogenesis is defective in *S. cerevisiae* (Figure 3.1), because more p22 is present in several ribosome biogenesis mutants. In ribosome biogenesis mutants, there is a modest increase in the level of Ty1i RNA relative to Ty1 mRNA, and in the stability of p22 protein. As discussed in Chapter 2, analysis of extant Ribo-seq dataset with cells grown at optimal growth condition for yeast (30 °C) did not reveal noticeable difference between wild type and *loc1Δ*. Repeating Ribo-seq with *S. cerevisiae* CNC strains (CNC⁻ with 1 Ty1 element and CNC⁺ with 20 Ty1 elements, wild type and *loc1Δ*) grown at optimal growth condition for Ty1 retrotransposition (20 °C) would better address the question if the translational control is truly involved in increasing p22 level. The increase in p22 disrupts Gag-Gag interactions, and replication steps targeted by p22 are disrupted more aggressively. Besides changes in p22, enhanced CNC may be caused by changes in the cellular environment due to defects in ribosome biogenesis. This change may be responsible for activating some downstream stress response pathways that contribute to the p22 translation or stability. Also, it may alter cellular environment that sensitizes the process of Ty1 retrotransposition to the inhibitory effect of p22.

This study suggests that increase in Ty1i RNA is one reason for the increased level

of p22 increase in ribosome biogenesis mutants. However, the increase of Ty1i RNA is modest when compared to the increase in the level of p22 (compare quantification of Figure 2.4A and 4B). These results suggest there could be other reason(s) contributing to the p22 increase. In Chapter 2, we have discussed several possibilities for the increase in Ty1i RNA. However, we could not pinpoint a specific transcription factor or pathway. Recently, the Mediator complex is found to be involved in discriminating between transcription initiation of Ty1 and Ty1i mRNA. The relative level of Mediator at the Ty1i promoter versus the Ty1 promoter determines the ratio of Ty1i/Ty1 RNA [10]. How Mediator aids transcription of Ty1i RNA and Ty1 mRNA when ribosome biogenesis is defective requires further investigation. One possibility is that the Mediator aids transcription initiation of cellular pathways such as TOR pathway that is critical for expressing genes involved in ribosome biogenesis [11].

The increase in Ty1i RNA relative to Ty1 mRNA would contribute to the level of Gag and p22 protein. p22 disrupts Gag-Gag interaction by interacting with Gag, and the CNC-resistant mutations within a distinct region of *GAG* excludes p22 from VLPs [3, 12]. Increasing the level of p22 relative to Gag likely alters Gag-Gag interactions, and affects Gag assembly during formation of retrosomes and VLPs. Altered Ty1iRNA/Ty1 mRNA level would affect Gag and p22 level. As shown in Figure 2.4B, Gag level decreased while p22 level increased in CNC$^+$ ribosome biogenesis mutants compared to wild type. Indeed, the impact of the defect retrosome formation is quite remarkable in that retrosomes are completely abolished even without presence of puncta in a *loc1Δ* mutant (Figure 2.8). Ty1 retrosomes are nucleated when Gag binds to Ty1 mRNA/SRP/ribosome complex residing at the ER surface [13], implying a connection

between ribosome biogenesis and retrosome formation. It is still an open question if there are any cellular factors modulating the protein-protein interaction of Gag and if their expression or activity is altered when ribosome biogenesis is defective.

Since we identified ribosome biogenesis proteins with different roles (assembly factor, small and large ribosomal subunit proteins), it is unlikely there is a specific stage of ribosome biogenesis that modulates CNC. However, further examination for how ribosome biogenesis is altered in these mutant cells might be informative. Previous studies including the results in Chapter 2 were focused on analyzing Ty1 defects in mutant background. However, defects in other steps of ribosome biogenesis have been observed in mammalian cells as well as in yeast. For example, ribosome biogenesis is altered during virus infection such as expression changes of ribosome biogenesis genes in response to HIV-1, and changes in the rRNA maturation in response to herpes simplex virus type 1 [14]. Posttranscriptional modification of rRNA is suggested as the "bottleneck" in ribosome biogenesis, as demonstrated in a study of yeast killer virus [15].

We proposed earlier that retrosomes are biomolecular condensates or membrane-less organelles. The fascinating characteristic of biomolecular condensates is their formation is reversible and dynamic depending on cellular environment or stress conditions. In cells expressing endogenous Ty1 elements, Ty1 mRNA as well as Gag are abundantly expressed and retrosomes are formed almost similar level to cells overexpressing Ty1*his3-AI* (compare percentage of wild type cells containing retrosomes in Figure 2.8 and S2.5). However, VLPs are only detected in cells overexpressing Ty1 [16]. Besides the role in Ty1 as sites for VLP assembly, retrosomes

could serve as a biomolecular condensates formed by unknown cellular conditions, although different from other known biomolecular condensates in yeast such as P bodies or stress granules [13, 16]. In yeast, another biomolecular condensate formed by prion protein Sup35 was recently characterized [17]. The N-terminus sensor domain of Sup35 detects low cytosolic pH in the cell due to starvation, and reversible condensates are formed by N-terminus prion domain of the protein. These condensates are thought to be used to generating protein-specific environmental responses and promote cellular fitness [17]. Moreover, many DNA and RNA virus proteins contain disordered histone-like sequence and act as histone mimics, termed viral histone mimicry. Interaction of those proteins can promote phase separation and form biomolecular condensates. Those condensates may compete with host chromatin and affect gene expression of the host, or elicit nucleolar stress if formed in the nucleolus [18].

REFERENCES

1.   Kim, J.M., et al., *Transposable elements and genome organization: a comprehensive survey of retrotransposons revealed by the complete Saccharomyces cerevisiae genome sequence.* Genome Res, 1998. **8**(5): p. 464-78.

2.   Garfinkel, D.J., et al., *Post-transcriptional cosuppression of Ty1 retrotransposition.* Genetics, 2003. **165**(1): p. 83-99.

3.   Saha, A., et al., *A trans-dominant form of Gag restricts Ty1 retrotransposition and mediates copy number control.* J Virol, 2015. **89**(7): p. 3922-38.

4.   Cliften, P., et al., *Finding functional features in Saccharomyces genomes by phylogenetic footprinting.* Science, 2003. **301**(5629): p. 71-6.

5.   Kellis, M., et al., *Sequencing and comparison of yeast species to identify genes and regulatory elements.* Nature, 2003. **423**(6937): p. 241-54.

6.   Brachmann, C.B., et al., *Designer deletion strains derived from Saccharomyces cerevisiae S288c: a useful set of strains and plasmids for PCR-mediated gene disruption and other applications.* Yeast, 1998. **14**(2): p. 115-32.

7.   Giaever, G., et al., *Functional profiling of the Saccharomyces cerevisiae genome.* Nature, 2002. **418**(6896): p. 387-91.

8.   Mnaimneh, S., et al., *Exploration of essential gene functions via titratable promoter alleles.* Cell, 2004. **118**(1): p. 31-44.

9.   Breslow, D.K., et al., *A comprehensive strategy enabling high-resolution functional analysis of the yeast genome.* Nat Methods, 2008. **5**(8): p. 711-8.

10.   Salinero, A.C., et al., *The Mediator co-activator complex regulates Ty1 retromobility by controlling the balance between Ty1i and Ty1 promoters.* PLoS Genet, 2018. **14**(2): p. e1007232.

11.   Mohammadi, S., S. Subramaniam, and A. Grama, *Inferring the effective TOR-dependent network: a computational study in yeast.* BMC Syst Biol, 2013. **7**: p. 84.

12.   Tucker, J.M., et al., *The Ty1 retrotransposon restriction factor p22 targets Gag.* PLoS Genet, 2015. **11**(10): p. e1005571.

13. Doh, J.H., S. Lutz, and M.J. Curcio, *Co-translational localization of an LTR-retrotransposon RNA to the endoplasmic reticulum nucleates virus-like particle assembly sites.* PLoS Genet, 2014. **10**(3): p. e1004219.

14. Belin, S., et al., *Uncoupling ribosome biogenesis regulation from RNA polymerase I activity during herpes simplex virus type 1 infection.* RNA, 2010. **16**(1): p. 131-40.

15. Meskauskas, A., et al., *Delayed rRNA processing results in significant ribosome biogenesis and functional defects.* Mol Cell Biol, 2003. **23**(5): p. 1602-13.

16. Checkley, M.A., et al., *P-body components are required for Ty1 retrotransposition during assembly of retrotransposition-competent virus-like particles.* Mol Cell Biol, 2010. **30**(2): p. 382-98.

17. Franzmann, T.M., et al., *Phase separation of a yeast prion protein promotes cellular fitness.* Science, 2018. **359**(6371).

18. Tarakhovsky, A. and R.K. Prinjha, *Drawing on disorder: How viruses use histone mimicry to their advantage.* J Exp Med, 2018. **215**(7): p. 1777-1787.
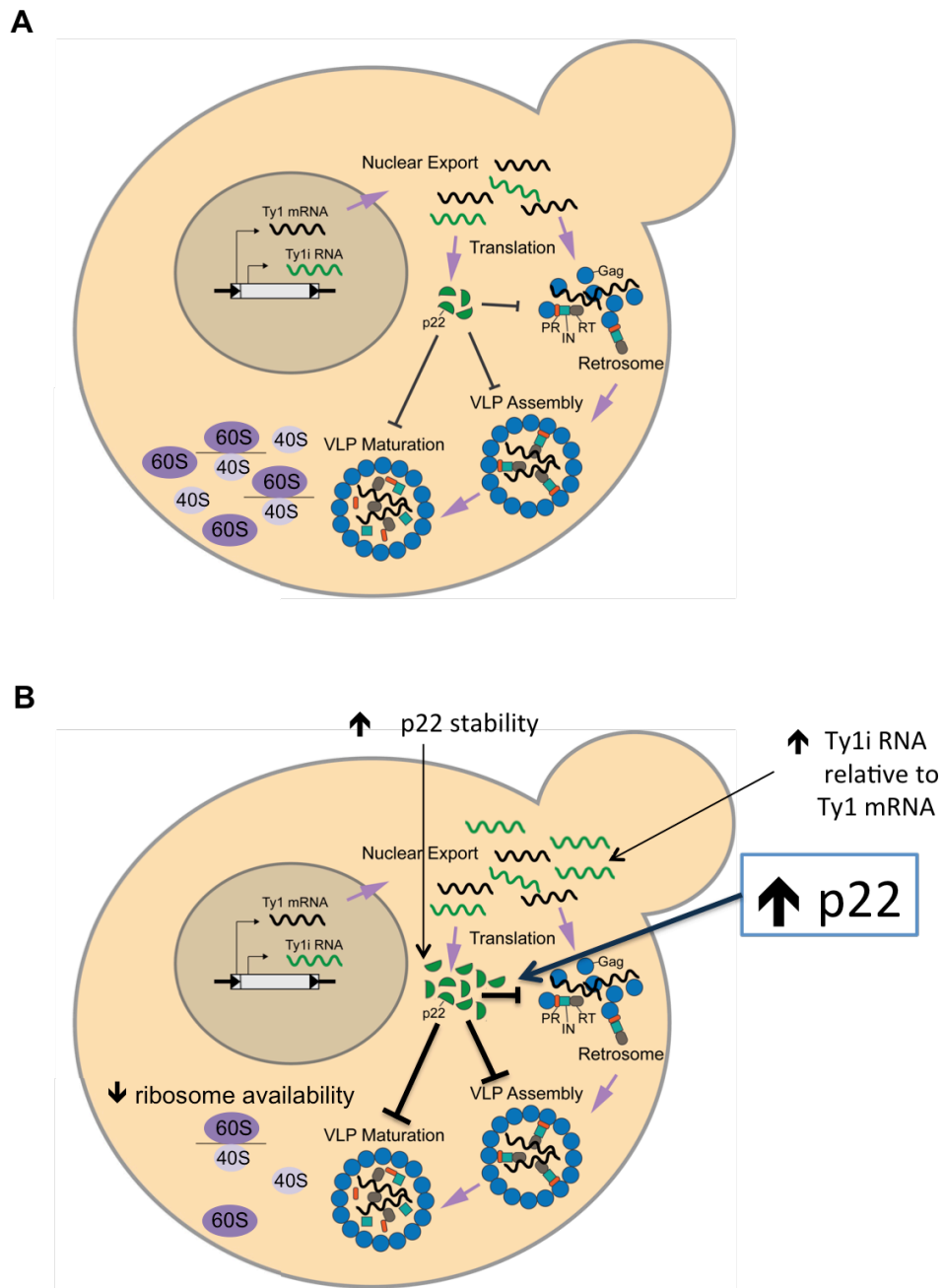
Figure 3.1 Enhanced CNC in cells defective for ribosome biogenesis

CNC phenotypes in CNC$^+$ wild type (A) and ribosome biogenesis mutants (B) are

shown here. (A) Ty1 mRNA and Ty1i RNA are transcribed and exported to the

cytoplasm. Ty1 mRNA is translated to Gag and Gag-Pol, and Ty1i RNA is translated to p22. p22 disrupts retrosome formation, VLP assembly, and VLP maturation. (B) Ty1i RNA level is increased relative to Ty1 mRNA. Also, p22 stability is enhanced. More p22 aggravates disruption of retrosome formation, VLP assembly, and VLP maturation. Bolder line indicates stronger inhibition. Changes in the cellular environment resulting from defects in ribosome biogenesis such as reduced ribosome availability may also enhance CNC either by eliciting stress response that may affect translation or stability of p22 or sensitizing the inhibitory effect of p22.

APPENDIX

Suresh S, **Ahn HW**, Joshi K, Dakshinamurthy A, Kananganat A, Garfinkel DJ,

Farabaugh PJ. 2015. Ribosomal protein and biogenesis factors affect multiple steps

during movement of the Saccharomyces cerevisiae Ty1 retrotransposon. Mobile DNA.

2015 Dec 8;6:22.

Abstract

Background: A large number of Saccharomyces cerevisiae cellular factors modulate

the movement of the retrovirus-like transposon Ty1. Surprisingly, a significant number of

chromosomal genes required for Ty1 transposition encode components of the

translational machinery, including ribosomal proteins, ribosomal biogenesis factors,

protein trafficking proteins and protein or RNA modification enzymes.

Results: To assess the mechanistic connection between Ty1 mobility and the

translation machinery, we have determined the effect of these mutations on ribosome

biogenesis and Ty1 transcriptional and post-transcriptional regulation. Lack of genes

encoding ribosomal proteins or ribosome assembly factors causes reduced

accumulation of the ribosomal subunit with which they are associated. In addition, these

mutations cause decreased Ty1 + 1 programmed translational frameshifting, and

reduced Gag protein accumulation despite at least normal levels of Ty1 mRNA. Several

ribosome subunit mutations increase the level of both an internally initiated Ty1 transcript and its encoded truncated Gag-p22 protein, which inhibits transposition.

Conclusions: Together, our results suggest that this large class of cellular genes modulate Ty1 transposition through multiple pathways. The effects are largely post-transcriptional acting at a variety of levels that may include translation initiation, protein stability and subcellular protein localization.