DISCOVERY OF A NOVEL RESTRICTION FACTOR ENCODED BY THE RETROTRANSPOSON TY1 IN *SACCHAROMYCES CEREVISIAE*

By

AGNIVA SAHA

(Under the Direction of David J. Garfinkel)

ABSTRACT

Ty1 is the most active long terminal repeat (LTR) retrotransposon in *Saccharomyces cerevisiae* and resembles retroviruses in genome organization and replication mechanisms. Ty1 encodes the capsid protein of virus-like particles (VLPs) called Gag and enzymes required for its protein processing, reverse transcription and integration. *S. cerevisiae* and its sister species *S. paradoxus* lack canonical transposition inhibition mechanisms like RNAi, yet maintain tight control over the Ty1 replication via the mechanism of copy number control (CNC). Work presented here focuses on the discovery of a new Ty1 protein named p22, which we show is both necessary and sufficient for CNC. This Ty1 restriction factor is encoded by a previously uncharacterized subgenomic Ty1 sense transcript termed Ty1i RNA. It initiates ~ 800 bp downstream of Ty1 mRNA in *GAG* coding sequence and is translated in the same

reading frame *as GAG*. Therefore, p22 shares protein sequence with the C-terminal half of Gag. Ty1i RNA and p22 are present in several wild type *S. cerevisiae* and *S. paradoxus* strains. Interestingly their levels increase in the cytoplasmic exoribonuclease x*rn1Δ* and Ty1 transcription factor *spt3Δ* mutants. Co-sedimentation analyses suggest that p22 associates with VLPs and co-immunoprecipitation of p22 and Gag suggests that this association is a result of p22 binding to Gag. This p22-VLP association leads to aberrant protein processing and abnormal VLP morphology as demonstrated by electron microscopy. However, the earliest step in the Ty1 life cycle that is affected by p22 is during formation of retrosomes, the cytoplasmic sites where VLP assembly takes place. Fluorescence in situ hybridization and immunoflouorescence experiments show that p22 disrupts retrosomes, perhaps leading to above mentioned defective VLPs. In summary, the work presented here focuses on the discovery of the novel Ty1 restriction factor p22 and how it alters Ty1 VLP structure and function, thereby bringing about Ty1 CNC.

Index words: Retrotransposon, retrovirus, LTR, virus-like particles, restriction factor, copy number control, genome defense

DISCOVERY OF A NOVEL RESTRICTION FACTOR ENCODED BY THE RETROTRANSPOSON TY1 IN *SACCHAROMYCES CEREVISIAE*

by

AGNIVA SAHA

B.S., St. Joseph's College, Bangalore, India – 2005

M.S., CMR Institute of Management Studies, Bangalore, India – 2007

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

2017

DISCOVERY OF A NOVEL RESTRICTION FACTOR ENCODED BY THE RETROTRANSPOSON TY1 IN *SACCHAROMYCES CEREVISIAE*

by

AGNIVA SAHA

Major Professor: David J. Garfinkel

Committee: Stephen Hajduk

Walter K. Schmidt

Christopher West

Electronic Version Approved:

Suzanne Barbour

Dean of the Graduate School

The University of Georgia

August 2017

DEDICATION

I dedicate this dissertation to my parents Amitava Saha and Namita Saha. I could not have done it without your unconditional love and support. I also dedicate this to my sister Poulomee Das for being my pillar of strength in a foreign country and helping me in every way possible. I would also like to dedicate my thesis to my high school teachers Sharmila Dutta Bose, and my college professors Saraswati Raman, Nayeemullah Khan, Betty Daniel and Romea Pramodh. Thanks for believing in me and encouraging me to pursue research.

ACKNOWLEDGEMENTS

First, I would thank my PI and mentor David J. Garfinkel for giving me the opportunity to complete my PhD is his lab. You are a great mentor and I have learned a lot from you not only as a scientist but also as a good human being. I would like to thank my thesis committee members Stephen Hajduk, Walter Schmidt and Christopher West for being there for me whenever I needed them. I want to thank all my lab mates and colleagues I have worked with especially Jessica Tucker, Hyowon Ahn, Jonathan Hildreth and Ellen Krall. Jessica and Hyowon, you are the best lab mates one can ever dream of and I want to thank you for being the nicest colleagues and sweetest friends. I would also like to thank Linda Eizenstat and Michael Carlock for being awesome friends and helping everyone in the lab making our work so much easier. I would also like to thank my fellow graduate students of BCMB and all my teachers here at UGA. You have made my experience here at UGA so much richer and you have taught me a lot over the years.

TABLE OF CONTENTS

CHAPTER 1

INTRODUCTION AND LITERATURE REVIEW

Transposable elements (TEs) or transposons are specific fragments of DNA that are able to physically move from one genomic location to another. Hence they have been historically referred to as 'jumping genes' after their discovery by Barbara McClintock in the 1940s (1). Her seminal work on 'controlling elements' in *Zea mays* was awarded the Nobel Prize in Physiology or Medicine in 1983 and paved the path for transposon research in other organisms including humans. Transposon content of genomes varies greatly between organisms. For example, 75% of the maize genome is comprised of transposons (2) whereas that number is only 3% for the budding yeast *Saccharomyces cerevisiae* (3, 4). The Human Genome Project revealed that approximately 45% of human genomes are comprised of transposon derived DNA (5-7). Several examples showing positive roles of transposons have been reported, which include placentation and successful pregnancy in mammals (8) and even co-option by interferon induced genes that now function as part of the innate immunity system in humans (9). Numerous reports have also linked transposons to various human diseases including cancer, hemophilia, neurofibromatosis and macular degeneration (10, 11). The focus of my research is the Ty1 retrotransposon of budding yeast and intrinsic mechanisms that control Ty1 replication. Studying Ty1 is medically important because Ty1 replicates using very similar mechanisms as the human immunodeficiency virus (HIV). Therefore, understanding how yeast restricts Ty1 transposition can advance our

knowledge on retroelement propagation and control in other eukaryotes including humans. *Saccharomyces cerevisiae* is a great model to study novel transposon control mechanisms since it lacks conserved genome defense pathways employed by eukaryotes to combat TEs such as RNA interference (12) and DNA methylation (13, 14).

Mammalian retrotransposons

Human genomes are host to two broad classes of TEs. Class I TEs known as retrotransposons, replicate via a 'copy-and-paste' mechanism that involves reverse transcription of their genomic RNAs (15, 16). Class II TEs are DNA transposons which replicate via a 'cut-and-paste' mechanism and unlike retrotransposons, do not cause as dramatic an increase in the number of genomic copies as Class I TEs. Genomes of some mammals, like bats, have large numbers of DNA transposons which have been active quite recently during evolution (17). However, most DNA transposons are transpositionally inactive in humans and other mammals (18) and will not be further described here. Retrotransposons can be classified into two types: those that have Long Terminal Repeats (LTRs) and those that do not (non-LTR). Both subtypes are medically important. LTR retrotransposons like Human Endogenous Retroviruses (HERVs) have been implicated in neurodegenerative diseases like multiple sclerosis (MS) (19) whereas non-LTR retrotransposons like LINE-1 or L1 elements have been linked to different cancers (10).

Long INterspersed Element-1 or L1 elements

LINE-1 or L1 elements in humans are a class of non-LTR TE. L1 transposons are autonomous i.e. they encode all the proteins necessary to be transposition competent.

They are the most active transposons in humans with over 500,000 insertions present in our genome (10, 11). However, most of these insertions have lost the capacity to transpose further due to mutations accumulated over evolutionary time including 5' truncations, point mutations and large rearrangements (7, 20-22). Less than 100 of these L1 elements remain transpositionally active in humans (23, 24). Active L1 elements are 6 kb long, encode two open reading frames, ORF1p and ORF2p, and also contain 5' and 3' UTR regions that serve important roles in gene expression and priming reverse transcription. ORF1 and ORF2 are separated by a short 63 bp inter-ORF spacer and terminate in a long poly A tail (25). OFR1p is ~ 40 kD and has RNA binding (26, 27) and nucleic acid chaperone activities (28). ORF2p is ~ 150 kD and contains endonuclease (29) as well as reverse transcriptase activities (30). These two proteins are essential for L1 transposition (31). A recently discovered L1 ORF called ORF0, is transcribed from the antisense strand of primate L1 elements (32). It encodes a 70 kD protein which increases L1 transposition by 41% in cell culture when overexpressed (33). However, the role of ORF0p in replication of L1 is not fully understood.

L1 is transcribed by RNA pol II from its native promoter (34) located in the 5' UTR. L1 RNAs are 5' capped as suggested by presence of non-template guanosine residues present in the 5' end of full length genomic copies (35). L1 RNA transcription is terminated by a polyA signal sequence present in the 3' UTR (31, 32, 36). Sometimes, transcription read-through takes place past the L1 polyA signal and terminates at alternate polyA signals located downstream of the L1 element in the genome. This transcriptional read-through results in hybrid L1 transcripts comprised of non-L1 sequence at the 3' end. The new sequence can be transposed to other genomic

locations by a process termed as 3' transduction (36-38), and can lead to genomic expansions and shuffling of protein-coding exons (36, 39).

Following transcription, L1 RNA is exported into the cytoplasm where it is translated. ORF2p is translated by a unique mechanism where ORF1p translation termination is immediately followed by reinitiation of ORF2p translation (40). Translation of L1 proteins is followed by assembly of a L1 ribonucleoprotein (RNP), which is comprised of several ORF1p trimers and at least two molecules of ORF2p (41, 42). Several non-L1 cellular RNAs and proteins are also part of these RNPs although the roles of these cellular factors in L1 RNA assembly are not clear (42-45).

Genomic insertions of L1 transposons take place by a mechanism known as target-primed reverse transcription and has been described in great detail using the insect *Bombyx mori* R2 elements. (46, 47). R2 differs from human L1 elements in that it encodes only one ORF. However, this one R2 protein contains RNA binding, endonuclease as well as reverse transcriptase functions (46, 48). The target site for L1 integration may result, at least in part, by the endonuclease specificity of ORF2p, which has been shown to recognize and cleave the consensus sequence 5'- TTTT/AA-3' where '/' indicates the cleavage site (29, 31, 49-51). ORF2p catalyzed cleavage of the bottom strand of a genomic L1 element creates a free 3' OH center that is used to prime ORF2p mediated reverse transcription of the L1 RNA to generate a single stranded L1 cDNA (42, 47). This is followed by the cleavage of the DNA top strand and is used by ORF2p to synthesize the second cDNA strand of L1 using DNA-dependent DNA synthesis activity of ORF2p (52). Please see figure 1-1 for stages of L1 replication. It is still not completely understood what other determinants influence L1 target site

specificity. Open chromatin states and perhaps protein-protein interactions between L1 and cellular proteins may affect integration site specificity for L1 integration (53).

Endogenous Retroviruses (ERVs) are genomic fossils of ancient retroviral infections

Endogenous retroviruses (ERVs) are a class of LTR retrotransposons present in the genomes of almost all vertebrates (54). ERVs likely originated from ancient exogenous retroviruses that infected host germline cells and integrated into their genomes as proviruses. Over evolutionary time, these viral sequences were endogenized, passed onto future generations and became permanent residents of vertebrate genomes (55, 56). The first reported ERV integration in humans occurred about 100 million years ago (57, 58) with the most recent integration reported to have taken place about 100,000 years ago (59). Remarkably, HERVs constitute about 8% of the human genome. HERVs can be broadly classified into three categories based on similarities with present exogenous retroviruses; Class I contains *Gammaretrovirus*-like HERVs whereas Classes II and III contain *Betaretrovirus*-like and *Spumaretrovirus*-like elements, respectively. HERVs can be further classified in 31 families (HERV-W, -K, -H, etc.), based on tRNA primer binding site sequence similarities (60). All ERVs share a common genome structure that is very similar to modern retroviruses; two open reading frames encoding viral proteins Gag and Pol are bracketed by LTRs which contain promoter and enhancer sequences for transcription (see figure 1-2). Gag is the structural protein of ERV virus-like particles (VLPs) which are viral counterparts of protein capsids that packages dimeric genomic RNAs to form an intact virion or VLP. Like infectious retroviruses, the Pol polyprotein contains segments with reverse transcriptase, integrase and protease activities which are required for reverse

transcription of ERV RNA into cDNA, integration of the cDNA into the host genome and proteolytic cleavage and processing of ERV proteins respectively. Unlike retroviruses, ERV VLPs are non-infectious; they do not bud from the host cell and infect neighboring cells to propagate themselves. This major difference between exogenous retroviruses and ERVs is due to the lack of the envelope (ENV) gene that allow retroviruses to form infectious virions. In fact, modern retroviruses are speculated to have evolved from Endogenous Viral Elements (EVEs) or LTR retrotransposons by acquisition of envelope gene (61). The fact that LTR retrotransposons like HERVs have been linked with diseases such as cancer (62), MS (63, 64), schizophrenia (65, 66) as well as autoimmune disorders (67), and their similarities with pathogenic retroviruses like HIV, make their research medically relevant.

Impact of retroelements on host genomes

Non-LTR (L1, Alu, etc) and LTR (ERV) retrotransposons impact their host genomes in a variety of negative and positive ways. The autonomous non-LTR retrotransposon L1 and the non-autonomous Short INterspersed Element (SINE) Alu are the most active retrotransposons in humans and are linked with most transposon related human diseases. SINEs like Alu depend on the autonomous L1 element encoded machinery for transposition since Alu elements lack their own transposition proteins. At least 124 diseases-causing mutations have been reported in humans that involve L1, Alu or SINEs like SINE-VNTR-Alu (SVA) elements (20, 52, 68-75). A large number of these diseases are caused by insertional inactivation of genes that occur during *de novo* transposition of these TEs. One of the earliest reports of a L1 mediated disease was of colon cancer caused by the insertion of a full length element in the

adenomatous polyposis coli (APC) tumor suppressor gene (76). A truncated insertion of a L1 element in exon 6 of the phosphatase and tensin gene was also reported to cause endometrial carcinoma (77). More recently, full length *de novo* L1 and SVA insertions in the retinoblastoma tumor suppressor RB1 and caspase 8 genes resulting in retinoblastoma and cutaneous basal cell carcinoma, respectively, have been reported (78, 79). Other studies have discovered the role of L1 elements may not be limited to cancer initiation. Several groups have reported that L1 elements may function in cancer progression as evidenced by elevated L1 mobility and protein expression in various cancers (77, 80-83). L1 elements have also been implicated in autoimmune diseases like Aicairdi-Goutieres syndrome (84-86), lupus erythematosus (87) as well as in other diseases such as age-related macular degeneration (88).

LTR retroelements like HERVs, during evolutionary time, have lost their ability to cause infections unlike their exogenous retroviral counterparts (89). Although most HERVs are silent due to epigenetic repression, a small number are expressed (90-92) and associated with human diseases. The association of HERV-W with MS has been reported (63, 64). Syncytin, the protein product of the *env* gene of HERV-W is overexpressed in brain astrocytes, which produces inflammatory cytokines and leads to elevated protein oxidation in the neurons (93). These processes are thought to be the important in causing MS, which suggests a pathogenic role of HERV-W. Expression of HERV-K elements in melanomas, germ cell tumors and ovarian cancers is much higher than in healthy tissues implying a role of HERV-K in cancers (94-100).

Millions of years of retrotransposon and human genome coevolution have led to transposon domestication, a process where transposon encoded proteins have

acquired host cellular functions (101, 102). One of the most fascinating examples of transposon domestication or ‘exaptation’ is the co-option of ERV sequences for activation of interferon induced genes during innate immunity in mammals (9). MER41, a primate specific ERV, has enhancer functions that help activate transcription of interferon-γ (INFNG) controlled genes upon viral infections including activation of the AIM2 inflammasome. Importantly, Cas9 mediated elimination of the ERV leads to impairment of this IFNG response. Another example of domestication of retroviral sequences by mammalian genomes is the co-option of ancient retroviral *env* genes which encode glycoproteins called syncytin. Syncytin proteins help form the syncytiotrophoblast, which is crucial for normal placental development in mammals. Knocking out *syncytin-A* in mice results is disrupted placenta leading to embryo lethality *in utero* (103). In humans HERV-W and HERV-FRD express *syncytin-1* and *syncytin-2* glycoproteins respectively (104, 105). These genes are also exclusively expressed in the human placenta and are thought to carry out the same functions as in mice. *Syncytin-2* is also thought to have immunosuppressive functions that may help produce immunological tolerance to prevent fetus rejection in mammals (104). HERV LTR sequences in synthesizing noncoding RNAs may play a role in pluripotency (92) or act as binding sites of pluripotency specific transcription factors like LBP9. Disruption of transcription factor binding to HERV-H LTR abolishes stem cell renewal (106). Protection against exogenous pathogenic viruses is another well documented example of retroelements positively impacting their host. Sheep carrying endogenous versions of the lung cancer causing Jaagsiekte Sheep Retrovirus (JSRV) have been shown to be immune to exogenous JSRV (107). The Gag protein encoded by the endogenous JSRV

enJS56A1 interferes with exogenous JSRV virion assembly in the late stages of the viral life cycle (108). A similar phenomenon has been demonstrated in mice immunity against Murine Leukemia Virus (MLV) mediated by Gag protein from an endogenous retrovirus homologous to the human HERV-L retroelement (109).

Retrotransposon landscape in budding yeast

The *S. cerevisiae* reference strain harbors 5 families of LTR retrotransposons denoted Ty1-Ty5. These retrotransposons constitute 3.4% of the genome (3, 4) and to date are the only transposons present in budding yeast. Ty1 is the most abundant among the five retrotransposon families in the reference strain with 32 full length copies followed by 13 copies of Ty2 and 1-3 copies of Ty3,4 and 5 (3). Ty1, 2, 4 and 5 can be classified under the *Pseudoviridae* family of retroelements, a family which also includes copia elements in *Drosophila*, Tnt1 and Tto1 elements in the tobacco plant among others. Ty3 is classified under the *Metaviridae* family, which include Gyspy retroelements in *Drosophila*, the Athila and Tat4 viral elements in *Arabidopsis*, among others. The *Metaviridae* Ty3/Gypsy superfamily most closely resembles retroviruses based on genome organization as well as protein sequence homology (4, 110, 111). Ty1 and Ty2 are closely related with almost identical LTRs except for one base deletion and some protein sequence divergence primarily in Gag (3). Ty3 and Ty4 elements probably invaded the yeast genome more recently than the other Ty elements since their sequences are less heterogeneous (3). Ty5 is inactive in *S. cerevisiae* and contains extremely heterogeneous LTRs and deletions. However, active Ty5 elements are present in the closely related yeast, *S. paradoxus* (112).

Gene structure and replication of Ty1

Ty1 is a 5918-bp retrotransposon bracketed on each end with 334 bp LTRs. The 'reference' Ty1 called Ty1-H3 is the best characterized Ty1 element (16) and all nucleotide coordinates used here correspond to Ty1-H3. Each LTR has subdomains called unique 3' sequence or U3 (240 nucleotides), a repetitious sequence called R (56 nucleotides) and unique 5' sequence called U5 (38 nucleotides). These subdomains in the LTRs are defined by their nucleotide positions on the Ty1 mRNA (also called genomic RNA), which is ~ 5.7 kb long and is transcribed by RNA pol II from the R region in the 5' LTR to the end of the R region in the 3' LTR. The terminally redundant R sequences on the Ty1 mRNA are crucial for accurate reverse transcription into Ty1 cDNA. Ty1 contains two ORFs called *GAG* and *POL* that encode proteins Gag and a polyprotein Pol, which are analogous to retroviral Gag and Pol proteins respectively. Coding sequences of *GAG* and *POL* have a 38 bp overlap and a +1 ribosomal frame shift between overlapping leucine codons results in the production of the Gag- Pol precursor. Gag is the main structural protein of Ty1 virus-like particles (VLPs). Gag also possesses nucleic acid chaperone activity (NAC), which helps form and package dimeric Ty1 mRNA and mediates strand-transfer reactions during reverse transcription (113, 114). Retroviruses also package their genomic RNAs as dimers, perhaps to generate genetic diversity via recombination during reverse transcription (115). The Gag-Pol polyprotein is made of three individual Ty1 proteins in addition to Gag: protease (PR), integrase (IN) and reverse transcriptase (RT). PR is essential for processing of Gag and Gag-Pol proteins into their mature forms, IN catalyzes integration of cDNA, and reverse transcriptase (RT) catalyzes the conversion of Ty1

RNA into linear double-stranded cDNA. Like retroviruses, an association between RT and IN is essential for reverse transcription in vivo (116). Ty1 usually integrates upstream of RNA pol III transcribed genes through interactions between IN and pol III subunits (117-121). Please see figure 1-3 for the genetic organization of Ty1.

Expression of Ty1 and VLP assembly

Surprisingly, 0.1% - 0.8% of total cellular RNA and ~ 10% of mRNA is comprised of Ty1 transcripts (122, 123). Ty1 RNA is reported to have a long half-life of ~ 5 hours (124, 125) that may be due to an association with Gag (126). Although Ty1 mRNA is transcribed by RNA Pol II, only 15% is polyadenylated (127). Not all the Ty1s are transcribed efficiently as suggested by the 50-fold difference in expression of 31 out of the 32 Ty1s in the yeast genome monitored using fusions to LacZ (128). Transcription initiates and terminates in the R regions of the 5' and 3' LTRs, respectively, and results in a terminally redundant genomic transcript (129, 130). Like retroviruses, the termini of Ty1 RNA contain R-U5 and U3-R motifs. The 5' LTR houses a weak promoter activity. Ty1 sequences containing enhancer activity and binding sites for several transcription factors that modulate transcription are located in about 700 nucleotides of *GAG* sequence downstream of the transcription initiation site (131-134). In particular, Ste12 and Tec1 activate Ty1 transcription under normal growth conditions in haploid cells (135, 136), whereas Ty1 expression is reduced ~10 fold due to binding of the MATa1/α2 repressor in diploids (137). Tye7 enhances Ty1 transcription under adenine starvation possibly by downregulating expression of Ty1 antisense RNAs (138). Chromatin remodelers like Swi/Snf, ISWI (Imitation Switch homolog of *Drosophila*) and SAGA (Spt-Ada-Gcn5 acetyltransferase complex) also influence Ty1 transcription (128, 135, 136,

138-147). In cells lacking *SPT3*, which is a key component of SAGA complex, Ty1 genomic RNA level decreases dramatically (148). Termination of transcription occurs by endonucleolytic cleavage and polyadenylation of the Ty1 mRNA, and two sequences known as TS_1 and TS_2 located in the 3' LTR help in 3' end formation (149).

Once transcribed, Ty1 mRNA is exported by a Mex67 dependent mechanism (150) into the cytoplasm. Mex67 is an essential poly(A) RNA binding protein involved in RNA export and homologous with human Tap (151-154). There is also evidence that Ty1 Gag may help in the nuclear export of the mRNA into the cytoplasm (126). Once in the cytoplasm, Ty1 mRNA is translated into 49 kD Gag (Gag-p49) and the 199 kD Gag-Pol polyprotein (Gag-Pol-p199). The latter is formed by a +1 ribosomal frameshifting mechanism in the 38 nt overlap region between end of *GAG* and beginning of *POL* ORFs on the translating mRNA. This overlap region contains a hepta-nucleotide sequence 5' CUU-AGG-C 3' which is crucial for frameshifting. CUU and the +1 codon UUA both encode leucine via $tRNA^{Leu}$ (UAG) (155) and sometimes the translation slips from CUU into the +1 frame UUA due to ribosome pausing at the AGG codon, which encodes the very scarce $tRNA^{Arg}$ (CCU). This rare $tRNA^{Arg}$ (CCU) is encoded by a single copy of the gene *HSX1* (156, 157). This ribosomal frameshifting takes place with an efficiency of 3-13%, which leads to cellular ratio of Gag:Gag-Pol of about 20:1 (157, 158). Translational frameshifting is a common strategy used by viruses such as in HIV-1 (159), the L-A killer virus in yeast (160) and Mouse Mammary Tumor Virus (161) as well as retrotransposons like Ty3 (162) in order to produce a molar excess of Gag relative to Pol proteins. This ratio is required to form functional virus particles (157, 163-166). Other ways retroelements produce excess capsid relative to Pol proteins, include

inefficient readthrough translation across stop codons at the Gag-Pol junction in murine leukemia virus (MLV) (167) and selective posttranslational degradation of IN (Pol) in the Tf1 retrotransposon (168). Nonessential ribosome biogenesis factors Bud21 and Bud22 may be required for efficient Ty1 mRNA translation (169-171). The factors are involved in 18s rRNA maturation and formation of ribosomal small subunits.

Gag and Gag-Pol proteins localize in distinct cytoplasmic foci called retrosomes. These are thought to be nucleation or assembly sites for Ty1 VLPs and are analogous to assembly sites of retroviral particles. These foci were first described as T-bodies by Malagon et al. in 2008 and (150). However, the term ‘retrosomes’ was used earlier to describe Ty3 cytoplasmic granules (172). Ty1 retrosomes are detectable by a combination of FISH (Ty1 mRNA) and IF (Gag and Gag-Pol) techniques when cells are grown at 20°C, the permissive temperature for Ty1 transposition (150, 169, 173). At endogenous levels of Ty1 expression, VLPs cannot be detected by transmission electron microscopy. However, VLPs accumulate to high levels when cells express a multicopy pGTy1 plasmid containing Ty1 fused to the strong regulated *GAL1* promoter (173, 174). Another interesting aspect of Ty1 retrosome formation is the requirement for processing body (P-body) components. P-bodies are cytoplasmic granules that contain translationally repressed mRNAs, 5’-3’ exoribonucleases such as Xrn1, mRNA decapping activating enzymes Lsm1 and Pat1, as well as additional cellular proteins (175-180). Initially, P-body proteins and Ty3 retrosomes colocalize in the cytoplasm and Ty3 VLPs assembly occurs in P-bodies (181). However, the situation with Ty1 is more complex. Although P-body components are required for Ty1 transposition and normal retrosome appearance, Ty1 Gag and RNA fail to colocalize with P-body proteins, and

conditions such as glucose deprivation that disrupt retrosomes promote P-body formation (150, 173, 182). In particular, deleting *XRN1* markedly affects Ty1 retrosome formation, transposition, VLP assembly, and RNA packaging (182).

Interestingly, translation of Ty1 mRNA is related to formation of retrosomes. A recent study has unraveled the role of the signal recognition particle (SRP) on contranslational insertion of Gag into the endoplasmic reticulum (ER), retrosome nucleation and VLP assembly (169). Doh et al. (2014) suggest that Ty1 mRNA complexed with the translation apparatus associates with the signal recognition particles (SRP) on the ER. Following cotranslational insertion, Gag exits the ER by retrotranslocation and then binds Ty1 mRNA to nucleate retrosome formation. Mutations in SRP components Srp54 and Srp72 that slow down cotranslational insertion of Gag into ER or tunicamycin-mediated inhibition of translation elongation modulate the abundance of Ty1 retrosomes. In addition, their work explains earlier observations that unlike L1, Ty1 proteins can act *in trans* on different Ty1 transcripts during the process of retrotransposition (16, 183).

A major unanswered question addressed in my work is understanding what keeps VLPs from assembling. Binding of Gag molecules to Ty1 RNA in retrosomes nucleates multimerization of Gag, which may lead to dissociation of the RNA from the translation machinery (169). However, the Ty1 protein/RNA complexes in retrosomes are structurally distinct from VLPs, and overexpression of a Ty1 is required to detect assembled VLPs within retrosomes (173). VLPs are protein shells comprised of immature Gag, Gag-Pol and dimeric Ty1 RNA, and are analogous to retroviral RNP cores or virions without envelope. Retroviral Gag contain distinct capsid (CA),

nucleocapsid (NC) and short spacer (SP) domains, which carry out functions such as virus assembly, folding and dimerization of viral RNA as distinct proteins following cleavage by PR. Ty1 Gag carries out all the same functions as a single protein, and lacks identifiable domains based on sequence homology with its retroviral counterparts (184). Gag expressed in *E. coli* can assemble into VLPs which shows that the only Ty1 protein necessary for VLP formation is Gag (185, 186). The average molecular weight of Ty1 VLPs are ~ 14 MD and they possess icosahedral symmetry with T numbers of 7 and 9 (187-189). The VLPs are composed of an electron dense Gag shell with extensions or spike like structures. The shell is porous allowing the entry of small molecules like nucleotides, as well as globular proteins like RNaseA (~18 kD). Larger proteins like the nuclease benzonase (~30 kD), cannot access the packaged Ty1 mRNA (187, 188). Immunological probing of Ty1 VLPs have revealed the orientation of proteins inside these particles. The N-terminus of Gag and Gag-Pol is exposed on the VLP surface whereas the C-terminus of these proteins face the inside of VLPs (190). Residues that are important for Gag to form VLPs are located throughout most of the protein (191). Mutations in these regions of Gag result in abnormally large VLPs which are up to 8 times larger than the normal sized 30-80 nm VLP (192). Recent studies have predicted residues 1-172 and 355-401 to be highly disordered while regions 173-354 is predicted to be α-helical (114, 192). The C-terminal 355-401 region of Ty1 Gag has been recently shown to have nucleic acid chaperone activity, and this is consistent with previously published results carried out with a Gag peptide (TYA1-D) (113).

Once Ty1 VLPs have formed, PR cleaves immature Gag-p49 to mature Gag-p45 using a C-terminal processing site. PR also cleaves Gagl-Pol-p199 into mature Gag,

PR, RT and IN proteins (see figure 1-4). Ty1 PR is an aspartyl protease and is thought to be active only in VLPs (187). Processing of Ty1 proteins by PR is essential for normal retrotransposition of Ty1 (193-195). Active site mutations in Ty1 PR lead to defects in reviser transcription, VLP RNA packaging and less efficient RNA dimerization (195, 196).

Following Ty1 VLP maturation by PR, RT and IN reverse transcribes the packaged RNA dimer. An interaction between the N-terminus of RT and the C-terminus of IN is necessary for reverse transcription *in vivo* (116). An initiator $tRNA_i^{Met}$ is selectively packaged inside VLPs to primer cDNA synthesis (197, 198). The 3' end of the $tRNA_i^{Met}$ hybridizes with the 10-nt primer binding site (PBS) located in the 5' end of Ty1 mRNA. Two other binding sites for $tRNA_i^{Met}$ called Box0 and Box1 have also been mapped (199, 200). A long-range interaction between a short sequence in the 5' end of Ty1 mRNA called CYC5 and its complimentary sequence on the 3' end of the RNA called CYC3 is also crucial for efficient reverse transcription of Ty1 cDNA (113, 201). Another intramolecular interaction in Ty1 mRNA was identified as base pairing between part of the 5' R region (5' LTR) and a short 6 nt sequence downstream of the PBS (202). This interaction contributes to a RNA pseudoknot that is important for transposition (124). Following reverse transcription, Ty1 cDNA forms a pre-integration complex (PIC) with IN. The PIC is imported into the nucleus via a C-terminus bipartite nuclear localization signal in IN by the classical importin-α pathway (203). IN catalyzes integration of Ty1 upstream of genes transcribed by RNA pol III such as yeast tRNA genes, although insertions in genes transcribed by RNA pol II occur at a lower frequency. The specificity of integration was recently shown to be due to a subunit of

RNA pol III called AC40, which interacts with Ty1 IN to direct cDNA integration (121). The RNA Pol III subcomplex Rpc53/57 has also been shown to interact with Ty1 IN suggesting its role in guiding Ty1 integration upstream of Pol III transcribed genes (118).

Ty1 Copy Number Control (CNC)

Ty1 is the most active retrotransposon in the *S. cerevisiae* reference strain (204-208). The 32 copies of Ty1 constitute about 3% of the 12 Mbp yeast genome, a much smaller fraction than in human and other eukaryotes, whose genomes are almost half transposon-derived. Although Ty1 RNA is very abundant, transposition is extremely low with about one in a million cells undergoing a single transposition event per generation (209-211). A unique CNC mechanism helps maintain this low frequency of Ty1 transposition (212). Ty1 CNC was first demonstrated in a natural Ty1-less of *S. paradoxus* strain that probably lost its original complement of Ty1 elements by LTR-LTR recombination. In this strain, transposition of an introduced Ty1 element marked with a retrotransposon indicator (RIG) gene *his3-AI* (209) decreases with increasing numbers of additional Ty1 elements. This phenomenon of CNC was also demonstrated in *S. cerevisiae*. Interestingly, CNC can work *in trans* on genomic as well as plasmid-borne Ty1 copies. A multicopy, *GAL1*-promoted Ty1 pGTy1 plasmid confers CNC on a chromosomal Ty1*his3-AI* element when cells are grown in glucose and pGTy1 expression in repressed. However, CNC is overcome when Ty1 overexpression is induced. These observations suggested that the factor responsible for CNC is encoded by Ty1, independent of normal Ty1 expression, and titratable (212). The minimal region of pGTy1 that is crucial for conferring CNC contains 5' R-U5 and *GAG* sequences (212).

Subsequent work by Matsuda and Garfinkel (2009) suggested that three Ty1 antisense (AS) RNAs, named AS RNA I, II and III were responsible for Ty1 CNC (213). The antisense transcripts map to the CNC region and their expression level was related to the level of CNC. Mutations in the R-U5 region that destabilized these AS RNAs also abolished Ty1 CNC. A model of CNC was proposed where the Ty1 AS RNAs from the CNC region associated with VLPs to bring about defects in VLP protein processing and lower IN and RT levels. The combined defects resulted in inhibition of reverse transcription, and hence transposition. A major caveat in the AS-RNA based CNC model was that ectopic expression of any of the AS transcripts failed to confer CNC or lower Ty1 RNA level (213, 214). Thus, the AS RNAs may require additional factors to confer CNC. My research presented in chapter 2 clearly shows that Ty1 AS RNAs play little if any role in Ty1 CNC. Instead, a subgenomic sense Ty1 RNA called Ty1i, is transcribed from the CNC region and encodes a protein (p22), which exerts a trans-dominant negative effect on Ty1 transposition (215). The protein sequence of p22 or its processed form p18, are part of the C-terminal half of p49/p45-Gag since Ty1i RNA transcription initiates in *GAG* about 800 nucleotides downstream of the Ty1 mRNA initiation site. Initiation of p22 translation occurs at two internal AUG codons present on Ty1i RNA that are in the same reading frame as *GAG* (114, 215, 216). p22/p18 demonstrates all the characteristics of a CNC factor. In cells with high Ty1 copy number (CNC^+), p22/p18 associates with VLPs to cause defects in VLP protein processing, lower levels of IN and RT, and less Ty1 cDNA synthesis (215).

Anti-retroviral restriction factors

Mammalian cells express a diverse array of cellular proteins that provide innate immunity against retroviruses, and in some cases other viruses as well. These proteins are called restriction factors and several well-studied examples will be discussed in this section. The apolipoprotein B messenger RNA editing enzyme catalytic polypeptide-like 3 (APOBEC3) proteins such as APOBEC3G (A3G), F and H, are anti-HIV restriction factors. The APOBEC3 family consists of 11 proteins with cytidine deaminase activity that are expressed in multiple human tissues (217, 218). Out of these eleven proteins, APOBEC3G is the best characterized restriction factor. A3G expression is induced by type 1 interferons, a hallmark of several innate immunity factors (217, 218). A3G is packaged into HIV-1 virions in the cytoplasm via a combination of viral RNA binding and an interaction between its N-terminus and the HIV-1 nucleocapsid (219). Once packaged, the C-terminus of A3G deaminates cytosine residues to uracil in the nascent negative strand of viral cDNA during reverse transcription. This results in guanosine to adenosine substitutions in the plus-strand of the cDNA, leading to loss of genomic integrity (220-223). A3G can also interfere with HIV-1 replication in a way that is independent of its cytidine deaminase function. A3G impedes HIV-1 reverse transcriptase translocation along the viral RNA, resulting in lower levels of cDNA during HIV-1 infection. However, the molecular mechanism behind this is unclear (224, 225). Importantly, HIV-1 has evolved a way to counter restriction by A3G via a HIV-1 encoded accessory protein called viral infectivity factor or Vif. Vif binding to A3G, recruits a ubiquitin ligase complex that leads to proteasome mediated degradation and inhibits A3G packaging into HIV-1 virions (223, 226-228). When ectopically expressed in yeast,

human A3G restricts Ty1 transposition by causing mutations in Ty1 cDNA (182, 229-231). Another member of the APOBEC3 family called APOBEC3A inhibits L1 retrotransposition in human cells (232). A novel, interferon induced HIV restriction factor called SAMHD1 is a nuclear protein with a sterile alpha motif (SAM) and a HD domain (HD), (233). This protein is expressed in myeloid cells including dendritic cells, monocytes and macrophages (234), and can be induced by type-I as well as type-II interferons (234-236). SAMHD1 was shown to be a deoxynucleoside triphospho-hydrolase that exclusively cleaves dNTPs but not ribonucleotides, DNA or RNA (237, 238). It is thought to reduce HIV cDNA synthesis by lowering intracellular dNTP pools. These assertions are supported by a correlation between SAMHD1's antiviral function and concentration of intracellular dNTPs in myeloid cells (239-241). Similar Vif, the HIV accessory protein Vpx inhibits SAMHD1 activity by ubiquitination followed by proteasomal degradation of SAMHD1 (242-245).

The anti-retroviral restriction factors most relevant to my research are ones that bind to viral capsid as part of their restriction mechanism. The first such anti-retroviral restriction factor identified was Fv1. The gene encoding Fv1 was found to control susceptibility of mice to leukemia caused by the Friend murine leukemia virus (246). When the gene at the genetically defined Fv1 locus was cloned by complementation analyses (109), it was found to be very closely related to the Gag gene of the endogenous MERV-L retrovirus found in both mice and humans. It was proposed that Fv1 was domesticated after an ancient MERV-L-like retrovirus infected the mouse germ cells about 7 million years ago (109). The mechanism by which Fv1 restricts MLV infection is not clearly understood, in part because its binding determinant on MLV was

difficult to identify. In 2011, it was shown that Fv1 could bind recombinant CA proteins assembled into Gag lattice, which is only present into mature MLV particles (247). Fv1 inhibits virus integration into host DNA by preventing pre-integration complex association with mitotic chromosomes (248). Another important cellular restriction factor called TRIM5α (Tripartite Motif 5α) was found to provide immunity against HIV through interactions with Gag (249). A more potent version of TRIM5α was discovered as gene fusion with Cyclophillin A (249, 250). TRIM5-CypA provides complete resistance to HIV in rhesus macaques and owl monkeys. Since its discovery TRIM5α proteins have been found in humans (251, 252), cattle (253, 254) as well as other vertebrates (255, 256). The C-terminal domain of TRIM5α called B30.2/PRYSPRY or simply SPRY, is critical in recognition and binding to retroviral capsids (257-260). In TRIM-Cyp, that function is carried out by the C-terminal Cyclophillin A (Cyp) domain, which has replaced the SPRY domain of TRIM5α in a number of monkeys (261-263). The mechanism of how TRIM5α restricts HIV-1 is not well understood. Evidence suggests that TRIM5α may interfere with retroviral infection by binding to viral capsid via the SPRY domain, resulting in accelerated uncoating and inhibition of reverse transcription (264, 265). Interestingly, although human TRIM5α (huTRIM5α) is not very efficient in virus restriction (249), single amino acid changes in the SPRY domain can make it as potent as an anti-retroviral restriction factor like monkey TRIM5α (258, 260). Another restriction factor called MX2 also blocks nuclear entry and cDNA integration of HIV-1 by binding to viral capsid (266-268). MX2 is a dynamin-like GTPase (269), although its GTPase function is not important for viral restriction (269-271). It brings about HIV-1 restriction by

preventing uncoating of viral particles and therefore inhibits entry of the PIC into the nucleus (267, 270-272).

The capsid-binding restriction factors Fv1, TRIM5α/TRIM-Cyp and MX2, affect their target viruses during early phase of infection: before integration of the viral cDNA into host chromosomes to form a provirus. The enJS56A1 is a late phase restriction factor that provides resistance against the Jaagsiekte sheep retrovirus (JSRV), a cause of lung cancer in sheep. Sheep genomes have 20 endogenous JSRV related sequences. The enJS56A1 locus encodes a trans-dominant JSRV Gag related protein that can thwart exogenous JSRV infection (273-275). Confocal microscopy experiments revealed than enJS56A1 proteins associate with JSRV Gag during viral assembly and disrupt pericentrosomal targeting and assembly of JSRV particles (276). This confers resistance by targeting viral proteins for degradation. Interestingly, p22 is derived from Ty1 Gag like Fv1 and enJS56A1 are derived from endogenous forms of viruses they restrict. Furthermore, p22 and enJS56A1 inhibit Ty1 and JRV particle assembly, respectively.

References

1. McClintock B. 1956. Controlling elements and the gene. Cold Spring Harb Symp Quant Biol 21:197-216.
2. Baucom RS, Estill JC, Chaparro C, Upshaw N, Jogi A, Deragon JM, Westerman RP, Sanmiguel PJ, Bennetzen JL. 2009. Exceptional diversity, non-random distribution, and rapid evolution of retroelements in the B73 maize genome. PLoS Genet 5:e1000732.
3. Kim JM, Vanguri S, Boeke JD, Gabriel A, Voytas DF. 1998. Transposable elements and genome organization: a comprehensive survey of retrotransposons revealed by the complete *Saccharomyces cerevisiae* genome sequence. Genome Res 8:464-478.
4. Carr M, Bensasson D, Bergman CM. 2012. Evolutionary genomics of transposable elements in *Saccharomyces cerevisiae*. PLoS One 7:e50978.
5. Smit AF. 1999. Interspersed repeats and other mementos of transposable elements in mammalian genomes. Curr Opin Genet Dev 9:657-663.
6. Smit AF. 1996. The origin of interspersed repeats in the human genome. Curr Opin Genet Dev 6:743-748.
7. Lander ES, Linton LM, Birren B, Nusbaum C, Zody MC, Baldwin J, Devon K, Dewar K, Doyle M, FitzHugh W, Funke R, Gage D, Harris K, Heaford A, Howland J, Kann L, Lehoczky J, LeVine R, McEwan P, McKernan K, Meldrim J, Mesirov JP, Miranda C, Morris W, Naylor J, Raymond C, Rosetti M, Santos R, Sheridan A, Sougnez C, Stange-Thomann Y, Stojanovic N, Subramanian A, Wyman D,

Rogers J, Sulston J, Ainscough R, Beck S, Bentley D, Burton J, Clee C, Carter N, Coulson A, Deadman R, Deloukas P, Dunham A, Dunham I, Durbin R, French L, Grafham D, et al. 2001. Initial sequencing and analysis of the human genome. Nature 409:860-921.

8. Lynch VJ, Nnamani MC, Kapusta A, Brayer K, Plaza SL, Mazur EC, Emera D, Sheikh SZ, Grutzner F, Bauersachs S, Graf A, Young SL, Lieb JD, DeMayo FJ, Feschotte C, Wagner GP. 2015. Ancient transposable elements transformed the uterine regulatory landscape and transcriptome during the evolution of mammalian pregnancy. Cell Rep 10:551-561.
9. Chuong EB, Elde NC, Feschotte C. 2016. Regulatory evolution of innate immunity through co-option of endogenous retroviruses. Science 351:1083-1087.
10. Richardson SR, Morell S, Faulkner GJ. 2014. L1 retrotransposons and somatic mosaicism in the brain. Annu Rev Genet 48:1-27.
11. Hancks DC, Kazazian HH, Jr. 2016. Roles for retrotransposon insertions in human disease. Mob DNA 7:9.
12. Drinnenberg IA, Weinberg DE, Xie KT, Mower JP, Wolfe KH, Fink GR, Bartel DP. 2009. RNAi in budding yeast. Science 326:544-550.
13. Capuano F, Mulleder M, Kok R, Blom HJ, Ralser M. 2014. Cytosine DNA methylation is found in Drosophila melanogaster but absent in *Saccharomyces cerevisiae*, *Schizosaccharomyces pombe*, and other yeast species. Anal Chem 86:3697-3702.
14. Proffitt JH, Davie JR, Swinton D, Hattman S. 1984. 5-Methylcytosine is not detectable in *Saccharomyces cerevisiae* DNA. Mol Cell Biol 4:985-988.

15. Garfinkel DJ, Boeke JD, Fink GR. 1985. Ty element transposition: reverse transcriptase and virus-like particles. Cell 42:507-517.

16. Boeke JD, Garfinkel DJ, Styles CA, Fink GR. 1985. Ty elements transpose through an RNA intermediate. Cell 40:491-500.

17. Ray DA, Feschotte C, Pagan HJ, Smith JD, Pritham EJ, Arensburger P, Atkinson PW, Craig NL. 2008. Multiple waves of recent DNA transposon activity in the bat, Myotis lucifugus. Genome Res 18:717-728.

18. Mitra R, Li X, Kapusta A, Mayhew D, Mitra RD, Feschotte C, Craig NL. 2013. Functional characterization of piggyBat from the bat Myotis lucifugus unveils an active mammalian DNA transposon. Proc Natl Acad Sci U S A 110:234-239.

19. Christensen T. 2016. Human endogenous retroviruses in neurologic disease. APMIS 124:116-126.

20. Ostertag EM, Kazazian HH, Jr. 2001. Biology of mammalian L1 retrotransposons. Annu Rev Genet 35:501-538.

21. Scott AF, Schmeckpeper BJ, Abdelrazik M, Comey CT, O'Hara B, Rossiter JP, Cooley T, Heath P, Smith KD, Margolet L. 1987. Origin of the human L1 elements: proposed progenitor genes deduced from a consensus DNA sequence. Genomics 1:113-125.

22. Szak ST, Pickeral OK, Makalowski W, Boguski MS, Landsman D, Boeke JD. 2002. Molecular archeology of L1 insertions in the human genome. Genome Biol 3:research0052.

23. Brouha B, Schustak J, Badge RM, Lutz-Prigge S, Farley AH, Moran JV, Kazazian HH, Jr. 2003. Hot L1s account for the bulk of retrotransposition in the human population. Proc Natl Acad Sci U S A 100:5280-5285.

24. Sassaman DM, Dombroski BA, Moran JV, Kimberland ML, Naas TP, DeBerardinis RJ, Gabriel A, Swergold GD, Kazazian HH, Jr. 1997. Many human L1 elements are capable of retrotransposition. Nat Genet 16:37-43.

25. Dombroski BA, Mathias SL, Nanthakumar E, Scott AF, Kazazian HH, Jr. 1991. Isolation of an active human transposable element. Science 254:1805-1808.

26. Khazina E, Weichenrieder O. 2009. Non-LTR retrotransposons encode noncanonical RRM domains in their first open reading frame. Proc Natl Acad Sci U S A 106:731-736.

27. Kolosha VO, Martin SL. 2003. High-affinity, non-sequence-specific RNA binding by the open reading frame 1 (ORF1) protein from long interspersed nuclear element 1 (LINE-1). J Biol Chem 278:8112-8117.

28. Martin SL, Bushman FD. 2001. Nucleic acid chaperone activity of the ORF1 protein from the mouse LINE-1 retrotransposon. Mol Cell Biol 21:467-475.

29. Feng Q, Moran JV, Kazazian HH, Jr., Boeke JD. 1996. Human L1 retrotransposon encodes a conserved endonuclease required for retrotransposition. Cell 87:905-916.

30. Mathias SL, Scott AF, Kazazian HH, Jr., Boeke JD, Gabriel A. 1991. Reverse transcriptase encoded by a human transposable element. Science 254:1808-1810.

31. Moran JV, Holmes SE, Naas TP, DeBerardinis RJ, Boeke JD, Kazazian HH, Jr. 1996. High frequency retrotransposition in cultured mammalian cells. Cell 87:917-927.

32. Grimaldi G, Skowronski J, Singer MF. 1984. Defining the beginning and end of KpnI family segments. EMBO J 3:1753-1759.

33. Denli AM, Narvaiza I, Kerman BE, Pena M, Benner C, Marchetto MC, Diedrich JK, Aslanian A, Ma J, Moresco JJ, Moore L, Hunter T, Saghatelian A, Gage FH. 2015. Primate-specific ORF0 contributes to retrotransposon-mediated diversity. Cell 163:583-593.

34. Swergold GD. 1990. Identification, characterization, and cell specificity of a human LINE-1 promoter. Mol Cell Biol 10:6718-6729.

35. Lavie L, Maldener E, Brouha B, Meese EU, Mayer J. 2004. The human L1 promoter: variable transcription initiation sites and a major impact of upstream flanking sequence on promoter activity. Genome Res 14:2253-2260.

36. Moran JV, DeBerardinis RJ, Kazazian HH, Jr. 1999. Exon shuffling by L1 retrotransposition. Science 283:1530-1534.

37. Goodier JL, Ostertag EM, Kazazian HH, Jr. 2000. Transduction of 3'-flanking sequences is common in L1 retrotransposition. Hum Mol Genet 9:653-657.

38. Pickeral OK, Makalowski W, Boguski MS, Boeke JD. 2000. Frequent human genomic DNA transduction driven by LINE-1 retrotransposition. Genome Res 10:411-415.

39. Xing J, Wang H, Belancio VP, Cordaux R, Deininger PL, Batzer MA. 2006. Emergence of primate genes by retrotransposon-mediated sequence transduction. Proc Natl Acad Sci U S A 103:17608-17613.

40. Alisch RS, Garcia-Perez JL, Muotri AR, Gage FH, Moran JV. 2006. Unconventional translation of mammalian LINE-1 retrotransposons. Genes Dev 20:210-224.

41. Hohjoh H, Singer MF. 1997. Ribonuclease and high salt sensitivity of the ribonucleoprotein complex formed by the human LINE-1 retrotransposon. J Mol Biol 271:7-12.

42. Kulpa DA, Moran JV. 2006. Cis-preferential LINE-1 reverse transcriptase activity in ribonucleoprotein particles. Nat Struct Mol Biol 13:655-660.

43. Moldovan JB, Moran JV. 2015. The Zinc-Finger Antiviral Protein ZAP Inhibits LINE and Alu Retrotransposition. PLoS Genet 11:e1005121.

44. Goodier JL, Cheung LE, Kazazian HH, Jr. 2013. Mapping the LINE1 ORF1 protein interactome reveals associated inhibitors of human retrotransposition. Nucleic Acids Res 41:7401-7419.

45. Dai L, Taylor MS, O'Donnell KA, Boeke JD. 2012. Poly(A) binding protein C1 is essential for efficient L1 retrotransposition and affects L1 RNP formation. Mol Cell Biol 32:4323-4336.

46. Luan DD, Korman MH, Jakubczak JL, Eickbush TH. 1993. Reverse transcription of R2Bm RNA is primed by a nick at the chromosomal target site: a mechanism for non-LTR retrotransposition. Cell 72:595-605.

47. Cost GJ, Feng Q, Jacquier A, Boeke JD. 2002. Human L1 element target-primed reverse transcription in vitro. EMBO J 21:5899-5910.

48. Xiong YE, Eickbush TH. 1988. Functional expression of a sequence-specific endonuclease encoded by the retrotransposon R2Bm. Cell 55:235-246.

49. Symer DE, Connelly C, Szak ST, Caputo EM, Cost GJ, Parmigiani G, Boeke JD. 2002. Human l1 retrotransposition is associated with genetic instability in vivo. Cell 110:327-338.

50. Gilbert N, Lutz-Prigge S, Moran JV. 2002. Genomic deletions created upon LINE-1 retrotransposition. Cell 110:315-325.

51. Gilbert N, Lutz S, Morrish TA, Moran JV. 2005. Multiple fates of L1 retrotransposition intermediates in cultured human cells. Mol Cell Biol 25:7780-7795.

52. Piskareva O, Schmatchenko V. 2006. DNA polymerization by the reverse transcriptase of the human L1 retrotransposon on its own template in vitro. FEBS Lett 580:661-668.

53. Cost GJ, Golding A, Schlissel MS, Boeke JD. 2001. Target DNA chromatinization modulates nicking by L1 endonuclease. Nucleic Acids Res 29:573-577.

54. Weiss RA. 2016. Human endogenous retroviruses: friend or foe? APMIS 124:4-10.

55. Sverdlov ED. 2000. Retroviruses and primate evolution. Bioessays 22:161-171.

56. Mayer J, Meese E. 2005. Human endogenous retroviruses in the primate lineage and their influence on host genomes. Cytogenet Genome Res 110:448-456.

57. Slokar G, Hasler G. 2015. Human Endogenous Retroviruses as Pathogenic Factors in the Development of Schizophrenia. Front Psychiatry 6:183.

58. Lee A, Nolan A, Watson J, Tristem M. 2013. Identification of an ancient endogenous retrovirus, predating the divergence of the placental mammals. Philos Trans R Soc Lond B Biol Sci 368:20120503.

59. Turner G, Barbulescu M, Su M, Jensen-Seaman MI, Kidd KK, Lenz J. 2001. Insertional polymorphisms of full-length endogenous retroviruses in humans. Curr Biol 11:1531-1535.

60. Gifford R, Tristem M. 2003. The evolution, distribution and diversity of endogenous retroviruses. Virus Genes 26:291-315.

61. Malik HS, Eickbush TH. 2001. Phylogenetic analysis of ribonuclease H domains suggests a late, chimeric origin of LTR retrotransposable elements and retroviruses. Genome Res 11:1187-1197.

62. Kassiotis G. 2014. Endogenous retroviruses and the development of cancer. J Immunol 192:1343-1349.

63. Perron H, Garson JA, Bedin F, Beseme F, Paranhos-Baccala G, Komurian-Pradel F, Mallet F, Tuke PW, Voisset C, Blond JL, Lalande B, Seigneurin JM, Mandrand B. 1997. Molecular identification of a novel retrovirus repeatedly isolated from patients with multiple sclerosis. The Collaborative Research Group on Multiple Sclerosis. Proc Natl Acad Sci U S A 94:7583-7588.

64. Garson JA, Tuke PW, Giraud P, Paranhos-Baccala G, Perron H. 1998. Detection of virion-associated MSRV-RNA in serum of patients with multiple sclerosis. Lancet 351:33.

65. Deb-Rinker P, Klempan TA, O'Reilly RL, Torrey EF, Singh SM. 1999. Molecular characterization of a MSRV-like sequence identified by RDA from monozygotic twin pairs discordant for schizophrenia. Genomics 61:133-144.

66. Deb-Rinker P, O'Reilly RL, Torrey EF, Singh SM. 2002. Molecular characterization of a 2.7-kb, 12q13-specific, retroviral-related sequence isolated by RDA from monozygotic twin pairs discordant for schizophrenia. Genome 45:381-390.

67. Nelson P, Rylance P, Roden D, Trela M, Tugnet N. 2014. Viruses as potential pathogenic agents in systemic lupus erythematosus. Lupus 23:596-605.

68. Chen JM, Stenson PD, Cooper DN, Ferec C. 2005. A systematic analysis of LINE-1 endonuclease-dependent retrotranspositional events causing human genetic disease. Hum Genet 117:411-427.

69. Belancio VP, Hedges DJ, Deininger P. 2008. Mammalian non-LTR retrotransposons: for better or worse, in sickness and in health. Genome Res 18:343-358.

70. de Boer M, van Leeuwen K, Geissler J, Weemaes CM, van den Berg TK, Kuijpers TW, Warris A, Roos D. 2014. Primary immunodeficiency caused by an exonized retroposed gene copy inserted in the CYBB gene. Hum Mutat 35:486-496.

71. Vogt J, Bengesser K, Claes KB, Wimmer K, Mautner VF, van Minkelen R, Legius E, Brems H, Upadhyaya M, Hogel J, Lazaro C, Rosenbaum T, Bammert S, Messiaen L, Cooper DN, Kehrer-Sawatzki H. 2014. SVA retrotransposon insertion-associated deletion represents a novel mutational mechanism

underlying large genomic copy number changes with non-recurrent breakpoints. Genome Biol 15:R80.

72. van der Klift HM, Tops CM, Hes FJ, Devilee P, Wijnen JT. 2012. Insertion of an SVA element, a nonautonomous retrotransposon, in PMS2 intron 7 as a novel cause of Lynch syndrome. Hum Mutat 33:1051-1055.

73. Nazaryan-Petersen L, Bertelsen B, Bak M, Jonson L, Tommerup N, Hancks DC, Tumer Z. 2016. Germline Chromothripsis Driven by L1-Mediated Retrotransposition and Alu/Alu Homologous Recombination. Hum Mutat 37:385-395.

74. Masson E, Hammel P, Garceau C, Benech C, Quemener-Redon S, Chen JM, Ferec C. 2013. Characterization of two deletions of the CTRC locus. Mol Genet Metab 109:296-300.

75. Peixoto A, Pinheiro M, Massena L, Santos C, Pinto P, Rocha P, Pinto C, Teixeira MR. 2013. Genomic characterization of two large Alu-mediated rearrangements of the BRCA1 gene. J Hum Genet 58:78-83.

76. Miki Y, Nishisho I, Horii A, Miyoshi Y, Utsunomiya J, Kinzler KW, Vogelstein B, Nakamura Y. 1992. Disruption of the APC gene by a retrotransposal insertion of L1 sequence in a colon cancer. Cancer Res 52:643-645.

77. Helman E, Lawrence MS, Stewart C, Sougnez C, Getz G, Meyerson M. 2014. Somatic retrotransposition in human cancer revealed by whole-genome and exome sequencing. Genome Res 24:1053-1063.

78. Rodriguez-Martin C, Cidre F, Fernandez-Teijeiro A, Gomez-Mariano G, de la Vega L, Ramos P, Zaballos A, Monzon S, Alonso J. 2016. Familial

retinoblastoma due to intronic LINE-1 insertion causes aberrant and noncanonical mRNA splicing of the RB1 gene. J Hum Genet 61:463-466.

79. Stacey SN, Kehr B, Gudmundsson J, Zink F, Jonasdottir A, Gudjonsson SA, Sigurdsson A, Halldorsson BV, Agnarsson BA, Benediktsdottir KR, Aben KK, Vermeulen SH, Cremers RG, Panadero A, Helfand BT, Cooper PR, Donovan JL, Hamdy FC, Jinga V, Okamoto I, Jonasson JG, Tryggvadottir L, Johannsdottir H, Kristinsdottir AM, Masson G, Magnusson OT, Iordache PD, Helgason A, Helgason H, Sulem P, Gudbjartsson DF, Kong A, Jonsson E, Barkardottir RB, Einarsson GV, Rafnar T, Thorsteinsdottir U, Mates IN, Neal DE, Catalona WJ, Mayordomo JI, Kiemeney LA, Thorleifsson G, Stefansson K. 2016. Insertion of an SVA-E retrotransposon into the CASP8 gene is associated with protection against prostate cancer. Hum Mol Genet 25:1008-1018.

80. Tubio JM, Li Y, Ju YS, Martincorena I, Cooke SL, Tojo M, Gundem G, Pipinikas CP, Zamora J, Raine K, Menzies A, Roman-Garcia P, Fullam A, Gerstung M, Shlien A, Tarpey PS, Papaemmanuil E, Knappskog S, Van Loo P, Ramakrishna M, Davies HR, Marshall J, Wedge DC, Teague JW, Butler AP, Nik-Zainal S, Alexandrov L, Behjati S, Yates LR, Bolli N, Mudie L, Hardy C, Martin S, McLaren S, O'Meara S, Anderson E, Maddison M, Gamble S, Group IBC, Group IBC, Group IPC, Foster C, Warren AY, Whitaker H, Brewer D, Eeles R, Cooper C, Neal D, Lynch AG, Visakorpi T, et al. 2014. Mobile DNA in cancer. Extensive transduction of nonrepetitive DNA mediated by L1 retrotransposition in cancer genomes. Science 345:1251343.

81. Ewing AD, Gacita A, Wood LD, Ma F, Xing D, Kim MS, Manda SS, Abril G, Pereira G, Makohon-Moore A, Looijenga LH, Gillis AJ, Hruban RH, Anders RA, Romans KE, Pandey A, Iacobuzio-Donahue CA, Vogelstein B, Kinzler KW, Kazazian HH, Jr., Solyom S. 2015. Widespread somatic L1 retrotransposition occurs early during gastrointestinal cancer evolution. Genome Res 25:1536-1545.

82. Rodic N, Steranka JP, Makohon-Moore A, Moyer A, Shen P, Sharma R, Kohutek ZA, Huang CR, Ahn D, Mita P, Taylor MS, Barker NJ, Hruban RH, Iacobuzio-Donahue CA, Boeke JD, Burns KH. 2015. Retrotransposon insertions in the clonal evolution of pancreatic ductal adenocarcinoma. Nat Med 21:1060-1064.

83. Doucet-O'Hare TT, Rodic N, Sharma R, Darbari I, Abril G, Choi JA, Young Ahn J, Cheng Y, Anders RA, Burns KH, Meltzer SJ, Kazazian HH, Jr. 2015. LINE-1 expression and retrotransposition in Barrett's esophagus and esophageal carcinoma. Proc Natl Acad Sci U S A 112:E4894-4900.

84. Hu S, Li J, Xu F, Mei S, Le Duff Y, Yin L, Pang X, Cen S, Jin Q, Liang C, Guo F. 2015. SAMHD1 Inhibits LINE-1 Retrotransposition by Promoting Stress Granule Formation. PLoS Genet 11:e1005367.

85. Zhao K, Du J, Han X, Goodier JL, Li P, Zhou X, Wei W, Evans SL, Li L, Zhang W, Cheung LE, Wang G, Kazazian HH, Jr., Yu XF. 2013. Modulation of LINE-1 and Alu/SVA retrotransposition by Aicardi-Goutieres syndrome-related SAMHD1. Cell Rep 4:1108-1115.

86. Stetson DB, Ko JS, Heidmann T, Medzhitov R. 2008. Trex1 prevents cell-intrinsic initiation of autoimmunity. Cell 134:587-598.

87. Hung T, Pratt GA, Sundararaman B, Townsend MJ, Chaivorapol C, Bhangale T, Graham RR, Ortmann W, Criswell LA, Yeo GW, Behrens TW. 2015. The Ro60 autoantigen binds endogenous retroelements and regulates inflammatory gene expression. Science 350:455-459.

88. Kaneko H, Dridi S, Tarallo V, Gelfand BD, Fowler BJ, Cho WG, Kleinman ME, Ponicsan SL, Hauswirth WW, Chiodo VA, Kariko K, Yoo JW, Lee DK, Hadziahmetovic M, Song Y, Misra S, Chaudhuri G, Buaas FW, Braun RE, Hinton DR, Zhang Q, Grossniklaus HE, Provis JM, Madigan MC, Milam AH, Justice NL, Albuquerque RJ, Blandford AD, Bogdanovich S, Hirano Y, Witta J, Fuchs E, Littman DR, Ambati BK, Rudin CM, Chong MM, Provost P, Kugel JF, Goodrich JA, Dunaief JL, Baffi JZ, Ambati J. 2011. DICER1 deficit induces Alu RNA toxicity in age-related macular degeneration. Nature 471:325-330.

89. Stoye JP. 2012. Studies of endogenous retroviruses reveal a continuing evolutionary saga. Nat Rev Microbiol 10:395-406.

90. Conley AB, Piriyapongsa J, Jordan IK. 2008. Retroviral promoters in the human genome. Bioinformatics 24:1563-1567.

91. Oja M, Peltonen J, Blomberg J, Kaski S. 2007. Methods for estimating human endogenous retrovirus activities from EST databases. BMC Bioinformatics 8 Suppl 2:S11.

92. Perot P, Mugnier N, Montgiraud C, Gimenez J, Jaillard M, Bonnaud B, Mallet F. 2012. Microarray-based sketches of the HERV transcriptome landscape. PLoS One 7:e40194.

93. Antony JM, van Marle G, Opii W, Butterfield DA, Mallet F, Yong VW, Wallace JL, Deacon RM, Warren K, Power C. 2004. Human endogenous retrovirus glycoprotein-mediated induction of redox reactants causes oligodendrocyte death and demyelination. Nat Neurosci 7:1088-1095.

94. Lower R, Lower J, Frank H, Harzmann R, Kurth R. 1984. Human teratocarcinomas cultured in vitro produce unique retrovirus-like viruses. J Gen Virol 65 (Pt 5):887-898.

95. Buscher K, Trefzer U, Hofmann M, Sterry W, Kurth R, Denner J. 2005. Expression of human endogenous retrovirus K in melanomas and melanoma cell lines. Cancer Res 65:4172-4180.

96. Kurth R, Bannert N. 2010. Beneficial and detrimental effects of human endogenous retroviruses. Int J Cancer 126:306-314.

97. Wang-Johanning F, Liu J, Rycaj K, Huang M, Tsai K, Rosen DG, Chen DT, Lu DW, Barnhart KF, Johanning GL. 2007. Expression of multiple human endogenous retrovirus surface envelope proteins in ovarian cancer. Int J Cancer 120:81-90.

98. Buscher K, Hahn S, Hofmann M, Trefzer U, Ozel M, Sterry W, Lower J, Lower R, Kurth R, Denner J. 2006. Expression of the human endogenous retrovirus-K transmembrane envelope, Rec and Np9 proteins in melanomas and melanoma cell lines. Melanoma Res 16:223-234.

99. Herbst H, Sauter M, Mueller-Lantzsch N. 1996. Expression of human endogenous retrovirus K elements in germ cell and trophoblastic tumors. Am J Pathol 149:1727-1735.

100. Muster T, Waltenberger A, Grassauer A, Hirschl S, Caucig P, Romirer I, Fodinger D, Seppele H, Schanab O, Magin-Lachmann C, Lower R, Jansen B, Pehamberger H, Wolff K. 2003. An endogenous retrovirus derived from human melanoma cells. Cancer Res 63:8735-8741.

101. Kaneko-Ishino T, Ishino F. 2012. The role of genes domesticated from LTR retrotransposons and retroviruses in mammals. Front Microbiol 3:262.

102. Henke C, Strissel PL, Schubert MT, Mitchell M, Stolt CC, Faschingbauer F, Beckmann MW, Strick R. 2015. Selective expression of sense and antisense transcripts of the sushi-ichi-related retrotransposon--derived family during mouse placentogenesis. Retrovirology 12:9.

103. Dupressoir A, Vernochet C, Bawa O, Harper F, Pierron G, Opolon P, Heidmann T. 2009. Syncytin-A knockout mice demonstrate the critical role in placentation of a fusogenic, endogenous retrovirus-derived, envelope gene. Proc Natl Acad Sci U S A 106:12127-12132.

104. Mi S, Lee X, Li X, Veldman GM, Finnerty H, Racie L, LaVallie E, Tang XY, Edouard P, Howes S, Keith JC, Jr., McCoy JM. 2000. Syncytin is a captive retroviral envelope protein involved in human placental morphogenesis. Nature 403:785-789.

105. Blaise S, de Parseval N, Benit L, Heidmann T. 2003. Genomewide screening for fusogenic human endogenous retrovirus envelopes identifies syncytin 2, a gene conserved on primate evolution. Proc Natl Acad Sci U S A 100:13013-13018.

106. Clausen J. 2003. Endogenous retroviruses and MS: using ERVs as disease markers. Int MS J 10:22-28.

107. Spencer TE, Mura M, Gray CA, Griebel PJ, Palmarini M. 2003. Receptor usage and fetal expression of ovine endogenous betaretroviruses: implications for coevolution of endogenous and exogenous retroviruses. J Virol 77:749-753.

108. Arnaud F, Caporale M, Varela M, Biek R, Chessa B, Alberti A, Golder M, Mura M, Zhang YP, Yu L, Pereira F, Demartini JC, Leymaster K, Spencer TE, Palmarini M. 2007. A paradigm for virus-host coevolution: sequential counter-adaptations between endogenous and exogenous retroviruses. PLoS Pathog 3:e170.

109. Best S, Le Tissier P, Towers G, Stoye JP. 1996. Positional cloning of the mouse retrovirus restriction gene Fv1. Nature 382:826-829.

110. Havecker ER, Gao X, Voytas DF. 2004. The diversity of LTR retrotransposons. Genome Biol 5:225.

111. Llorens C, Fares MA, Moya A. 2008. Relationships of gag-pol diversity between Ty3/Gypsy and Retroviridae LTR retroelements and the three kings hypothesis. BMC Evol Biol 8:276.

112. Zou S, Ke N, Kim JM, Voytas DF. 1996. The *Saccharomyces* retrotransposon Ty5 integrates preferentially into regions of silent chromatin at the telomeres and mating loci. Genes Dev 10:634-645.

113. Cristofari G, Ficheux D, Darlix JL. 2000. The GAG-like protein of the yeast Ty1 retrotransposon contains a nucleic acid chaperone domain analogous to retroviral nucleocapsid proteins. J Biol Chem 275:19210-19217.

114. Nishida Y, Pachulska-Wieczorek K, Blaszczyk L, Saha A, Gumna J, Garfinkel DJ, Purzycka KJ. 2015. Ty1 retrovirus-like element Gag contains overlapping

restriction factor and nucleic acid chaperone functions. Nucleic Acids Res 43:7414-7431.

115. Onafuwa-Nuga A, Telesnitsky A. 2009. The remarkable frequency of human immunodeficiency virus type 1 genetic recombination. Microbiol Mol Biol Rev 73:451-480, Table of Contents.

116. Wilhelm M, Wilhelm FX. 2005. Role of integrase in reverse transcription of the *Saccharomyces cerevisiae* retrotransposon Ty1. Eukaryot Cell 4:1057-1065.

117. Devine SE, Boeke JD. 1996. Integration of the yeast retrotransposon Ty1 is targeted to regions upstream of genes transcribed by RNA polymerase III. Genes Dev 10:620-633.

118. Cheung S, Ma L, Chan PH, Hu HL, Mayor T, Chen HT, Measday V. 2016. 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 291:6396-6411.

119. Baller JA, Gao J, Stamenova R, Curcio MJ, Voytas DF. 2012. A nucleosomal surface defines an integration hotspot for the *Saccharomyces cerevisiae* Ty1 retrotransposon. Genome Res 22:704-713.

120. Eigel A, Feldmann H. 1982. Ty1 and delta elements occur adjacent to several tRNA genes in yeast. EMBO J 1:1245-1250.

121. Bridier-Nahmias A, Tchalikian-Cosson A, Baller JA, Menouni R, Fayol H, Flores A, Saib A, Werner M, Voytas DF, Lesage P. 2015. Retrotransposons. An RNA polymerase III subunit determines sites of retrotransposon integration. Science 348:585-588.

122. Curcio MJ, Hedge AM, Boeke JD, Garfinkel DJ. 1990. Ty RNA levels determine the spectrum of retrotransposition events that activate gene expression in *Saccharomyces cerevisiae*. Mol Gen Genet 220:213-221.

123. Elder RT, St John TP, Stinchcomb DT, Scherer S, Davis RW. 1981. 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 45 Pt 2:581-591.

124. Huang Q, Purzycka KJ, Lusvarghi S, Li D, Legrice SF, Boeke JD. 2013. Retrotransposon Ty1 RNA contains a 5'-terminal long-range pseudoknot required for efficient reverse transcription. RNA 19:320-332.

125. Munchel SE, Shultzaberger RK, Takizawa N, Weis K. 2011. Dynamic profiling of mRNA turnover reveals gene-specific and system-wide regulation of mRNA decay. Mol Biol Cell 22:2787-2795.

126. Checkley MA, Mitchell JA, Eizenstat LD, Lockett SJ, Garfinkel DJ. 2013. Ty1 gag enhances the stability and nuclear export of Ty1 mRNA. Traffic 14:57-69.

127. Malagon F, Jensen TH. 2008. The T body, a new cytoplasmic RNA granule in *Saccharomyces cerevisiae*. Mol Cell Biol 28:6022-6032.

128. Morillon A, Benard L, Springer M, Lesage P. 2002. Differential effects of chromatin and Gcn4 on the 50-fold range of expression among individual yeast Ty1 retrotransposons. Mol Cell Biol 22:2078-2088.

129. Elder RT, Loh EY, Davis RW. 1983. 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 80:2432-2436.

130. Hou W, Russnak R, Platt T. 1994. Poly(A) site selection in the yeast Ty retroelement requires an upstream region and sequence-specific titratable factor(s) in vitro. EMBO J 13:446-452.

131. Dudley AM, Gansheroff LJ, Winston F. 1999. Specific components of the SAGA complex are required for Gcn4- and Gcr1-mediated activation of the his4-912delta promoter in *Saccharomyces cerevisiae*. Genetics 151:1365-1378.

132. Servant G, Pennetier C, Lesage P. 2008. Remodeling yeast gene transcription by activating the Ty1 long terminal repeat retrotransposon under severe adenine deficiency. Mol Cell Biol 28:5543-5554.

133. Curcio MJ, Lutz S, Lesage P. 2015. The Ty1 LTR-Retrotransposon of Budding Yeast, *Saccharomyces cerevisiae*. Microbiol Spectr 3:MDNA3-0053-2014.

134. Lesage P, Todeschini AL. 2005. Happy together: the life and times of Ty retrotransposons and their hosts. Cytogenet Genome Res 110:70-90.

135. Laloux I, Dubois E, Dewerchin M, Jacobs E. 1990. TEC1, a gene involved in the activation of Ty1 and Ty1-mediated gene expression in *Saccharomyces cerevisiae*: cloning and molecular analysis. Mol Cell Biol 10:3541-3550.

136. Morillon A, Springer M, Lesage P. 2000. Activation of the Kss1 invasive-filamentous growth pathway induces Ty1 transcription and retrotransposition in *Saccharomyces cerevisiae*. Mol Cell Biol 20:5766-5776.

137. Errede B, Company M, Hutchison CA, 3rd. 1987. Ty1 sequence with enhancer and mating-type-dependent regulatory activities. Mol Cell Biol 7:258-265.

138. Servant G, Pinson B, Tchalikian-Cosson A, Coulpier F, Lemoine S, Pennetier C, Bridier-Nahmias A, Todeschini AL, Fayol H, Daignan-Fornier B, Lesage P. 2012.

Tye7 regulates yeast Ty1 retrotransposon sense and antisense transcription in response to adenylic nucleotides stress. Nucleic Acids Res 40:5271-5282.

139. Ciriacy M, Freidel K, Lohning C. 1991. Characterization of trans-acting mutations affecting Ty and Ty-mediated transcription in *Saccharomyces cerevisiae*. Curr Genet 20:441-448.

140. Company M, Errede B. 1987. Cell-type-dependent gene activation by yeast transposon Ty1 involves multiple regulatory determinants. Mol Cell Biol 7:3205-3211.

141. Errede B. 1993. MCM1 binds to a transcriptional control element in Ty1. Mol Cell Biol 13:57-62.

142. Grant PA, Duggan L, Cote J, Roberts SM, Brownell JE, Candau R, Ohba R, Owen-Hughes T, Allis CD, Winston F, Berger SL, Workman JL. 1997. Yeast Gcn5 functions in two multisubunit complexes to acetylate nucleosomal histones: characterization of an Ada complex and the SAGA (Spt/Ada) complex. Genes Dev 11:1640-1650.

143. Gray WM, Fassler JS. 1996. Isolation and analysis of the yeast TEA1 gene, which encodes a zinc cluster Ty enhancer-binding protein. Mol Cell Biol 16:347-358.

144. Kent NA, Karabetsou N, Politis PK, Mellor J. 2001. In vivo chromatin remodeling by yeast ISWI homologs Isw1p and Isw2p. Genes Dev 15:619-626.

145. Madison JM, Dudley AM, Winston F. 1998. Identification and analysis of Mot3, a zinc finger protein that binds to the retrotransposon Ty long terminal repeat (delta) in *Saccharomyces cerevisiae*. Mol Cell Biol 18:1879-1890.

146. Pollard KJ, Peterson CL. 1997. Role for ADA/GCN5 products in antagonizing chromatin-mediated transcriptional repression. Mol Cell Biol 17:6212-6222.

147. Turkel S, Liao XB, Farabaugh PJ. 1997. GCR1-dependent transcriptional activation of yeast retrotransposon Ty2-917. Yeast 13:917-930.

148. Winston F, Durbin KJ, Fink GR. 1984. The *SPT3* gene is required for normal transcription of Ty elements in *S. cerevisiae*. Cell 39:675-682.

149. Yu K, Elder RT. 1989. Some of the signals for 3'-end formation in transcription of the *Saccharomyces cerevisiae* Ty-D15 element are immediately downstream of the initiation site. Mol Cell Biol 9:2431-2444.

150. Malagon F, Jensen TH. 2008. The T-body: a new cytoplasmic RNA granule in *Saccharomyces cerevisiae*. Mol Cell Biol doi:papers://3BFEA9DA-C419-4A26-B21C-2F05A923CFC0/Paper/p31.

151. Segref A, Sharma K, Doye V, Hellwig A, Huber J, Luhrmann R, Hurt E. 1997. Mex67p, a novel factor for nuclear mRNA export, binds to both poly(A)+ RNA and nuclear pores. EMBO J 16:3256-3271.

152. Rout MP, Aitchison JD, Suprapto A, Hjertaas K, Zhao Y, Chait BT. 2000. The yeast nuclear pore complex: composition, architecture, and transport mechanism. J Cell Biol 148:635-651.

153. Katahira J, Strasser K, Podtelejnikov A, Mann M, Jung JU, Hurt E. 1999. The Mex67p-mediated nuclear mRNA export pathway is conserved from yeast to human. EMBO J 18:2593-2609.

154. Gruter P, Tabernero C, von Kobbe C, Schmitt C, Saavedra C, Bachi A, Wilm M, Felber BK, Izaurralde E. 1998. TAP, the human homolog of Mex67p, mediates CTE-dependent RNA export from the nucleus. Mol Cell 1:649-659.

155. Weissenbach J, Dirheimer G, Falcoff R, Sanceau J, Falcoff E. 1977. Yeast tRNALeu (anticodon U--A--G) translates all six leucine codons in extracts from interferon treated cells. FEBS Lett 82:71-76.

156. Belcourt MF, Farabaugh PJ. 1990. Ribosomal frameshifting in the yeast retrotransposon Ty: tRNAs induce slippage on a 7 nucleotide minimal site. Cell 62:339-352.

157. Kawakami K, Pande S, Faiola B, Moore DP, Boeke JD, Farabaugh PJ, Strathern JN, Nakamura Y, Garfinkel DJ. 1993. A rare tRNA-Arg(CCU) that regulates Ty1 element ribosomal frameshifting is essential for Ty1 retrotransposition in *Saccharomyces cerevisiae*. Genetics 135:309-320.

158. Harger JW, Dinman JD. 2003. An in vivo dual-luciferase assay system for studying translational recoding in the yeast *Saccharomyces cerevisiae*. RNA 9:1019-1024.

159. Jacks T, Power MD, Masiarz FR, Luciw PA, Barr PJ, Varmus HE. 1988. Characterization of ribosomal frameshifting in HIV-1 gag-pol expression. Nature 331:280-283.

160. Dinman JD, Icho T, Wickner RB. 1991. A -1 ribosomal frameshift in a double-stranded RNA virus of yeast forms a gag-pol fusion protein. Proc Natl Acad Sci U S A 88:174-178.

161. Hizi A, Henderson LE, Copeland TD, Sowder RC, Hixson CV, Oroszlan S. 1987. Characterization of mouse mammary tumor virus gag-pro gene products and the ribosomal frameshift site by protein sequencing. Proc Natl Acad Sci U S A 84:7041-7045.

162. Farabaugh PJ, Zhao H, Vimaladithan A. 1993. A novel programed frameshift expresses the POL3 gene of retrotransposon Ty3 of yeast: frameshifting without tRNA slippage. Cell 74:93-103.

163. Dinman JD, Wickner RB. 1992. Ribosomal frameshifting efficiency and gag/gag-pol ratio are critical for yeast M1 double-stranded RNA virus propagation. J Virol 66:3669-3676.

164. Karacostas V, Wolffe EJ, Nagashima K, Gonda MA, Moss B. 1993. Overexpression of the HIV-1 gag-pol polyprotein results in intracellular activation of HIV-1 protease and inhibition of assembly and budding of virus-like particles. Virology 193:661-671.

165. Kirchner J, Sandmeyer SB, Forrest DB. 1992. Transposition of a Ty3 GAG3-POL3 fusion mutant is limited by availability of capsid protein. J Virol 66:6081-6092.

166. Xu H, Boeke JD. 1990. Host genes that influence transposition in yeast: the abundance of a rare tRNA regulates Ty1 transposition frequency. Proc Natl Acad Sci U S A 87:8360-8364.

167. Feng YX, Copeland TD, Oroszlan S, Rein A, Levin JG. 1990. Identification of amino acids inserted during suppression of UAA and UGA termination codons at

the gag-pol junction of Moloney murine leukemia virus. Proc Natl Acad Sci U S A 87:8860-8863.

168. Atwood A, Lin JH, Levin HL. 1996. The retrotransposon Tf1 assembles virus-like particles that contain excess Gag relative to integrase because of a regulated degradation process. Mol Cell Biol 16:338-346.

169. Doh JH, Lutz S, Curcio MJ. 2014. Co-translational localization of an LTR-retrotransposon RNA to the endoplasmic reticulum nucleates virus-like particle assembly sites. PLoS Genet 10:e1004219.

170. Dakshinamurthy A, Nyswaner KM, Farabaugh PJ, Garfinkel DJ. 2010. *BUD22* affects Ty1 retrotransposition and ribosome biogenesis in *Saccharomyces cerevisiae*. Genetics 185:1193-1205.

171. Dragon F, Gallagher JE, Compagnone-Post PA, Mitchell BM, Porwancher KA, Wehner KA, Wormsley S, Settlage RE, Shabanowitz J, Osheim Y, Beyer AL, Hunt DF, Baserga SJ. 2002. A large nucleolar U3 ribonucleoprotein required for 18S ribosomal RNA biogenesis. Nature 417:967-970.

172. Sandmeyer SB, Clemens KA. 2010. Function of a retrotransposon nucleocapsid protein. RNA Biol 7:642-654.

173. Checkley MA, Nagashima K, Lockett S, Nyswaner KM, Garfinkel DJ. 2010. P-bodies are required for Ty1 retrotransposition during assembly of retrotransposition-competent virus-like particles. Mol Cell Biol 30:382-398.

174. Garfinkel DJ, Boeke JD, Fink GR. 1985. Ty element transposition: reverse transcriptase and virus-like particles. Cell 42:507-517.

175. Decker CJ, Parker R. 2012. P-bodies and stress granules: possible roles in the control of translation and mRNA degradation. Cold Spring Harb Perspect Biol 4:a012286.

176. Coller J, Parker R. 2005. General translational repression by activators of mRNA decapping. Cell 122:875-886.

177. Bouveret E, Rigaut G, Shevchenko A, Wilm M, Seraphin B. 2000. A Sm-like protein complex that participates in mRNA degradation. EMBO J 19:1661-1671.

178. Tharun S, He W, Mayes AE, Lennertz P, Beggs JD, Parker R. 2000. Yeast Sm-like proteins function in mRNA decapping and decay. Nature 404:515-518.

179. Tharun S, Parker R. 2001. Targeting an mRNA for decapping: displacement of translation factors and association of the Lsm1p-7p complex on deadenylated yeast mRNAs. Mol Cell 8:1075-1083.

180. Stevens A. 1980. Purification and characterization of a *Saccharomyces cerevisiae* exoribonuclease which yields 5'-mononucleotides by a 5' leads to 3' mode of hydrolysis. J Biol Chem 255:3080-3085.

181. Beliakova-Bethell N, Beckham C, Giddings TH, Jr., Winey M, Parker R, Sandmeyer S. 2006. Virus-like particles of the Ty3 retrotransposon assemble in association with P-body components. RNA 12:94-101.

182. Dutko JA, Kenny AE, Gamache ER, Curcio MJ. 2010. 5' to 3' mRNA decay factors colocalize with Ty1 gag and human APOBEC3G and promote Ty1 retrotransposition. J Virol 84:5052-5066.

183. Curcio MJ, Garfinkel DJ. 1994. Heterogeneous functional Ty1 elements are abundant in the *Saccharomyces cerevisiae* genome. Genetics 136:1245-1259.

184. Adams SE, Mellor J, Gull K, Sim RB, Tuite MF, Kingsman SM, Kingsman AJ. 1987. The functions and relationships of Ty-VLP proteins in yeast reflect those of mammalian retroviral proteins. Cell 49:111-119.

185. Luschnig C, Bachmair A. 1997. RNA packaging of yeast retrotransposon Ty1 in the heterologous host, Escherichia coli. Biol Chem 378:39-46.

186. Luschnig C, Hess M, Pusch O, Brookman J, Bachmair A. 1995. The gag homologue of retrotransposon Ty1 assembles into spherical particles in Escherichia coli. Eur J Biochem 228:739-744.

187. Al-Khayat HA, Bhella D, Kenney JM, Roth JF, Kingsman AJ, Martin-Rendon E, Saibil HR. 1999. Yeast Ty retrotransposons assemble into virus-like particles whose T- numbers depend on the C-terminal length of the capsid protein. J Mol Biol 292:65-73.

188. Burns NR, Saibil HR, White NS, Pardon JF, Timmins PA, Richardson SM, Richards BM, Adams SE, Kingsman SM, Kingsman AJ. 1992. Symmetry, flexibility and permeability in the structure of yeast retrotransposon virus-like particles. Embo J 11:1155-1164.

189. Palmer KJ, Tichelaar W, Myers N, Burns NR, Butcher SJ, Kingsman AJ, Fuller SD, Saibil HR. 1997. Cryo-electron microscopy structure of yeast Ty retrotransposon virus- like particles. J Virol 71:6863-6868.

190. Brookman J, Stott A, Cheeseman P, Adamson C, Holmes D, Cole J, Burns N. 1999. Analysis of TyA protein regions necessary for formation of the Ty1 virus-like particle structure. Virology 212:69-76.

191. Roth JF. 2000. The yeast Ty virus-like particles. Yeast 16:785-795.

192. Martin-Rendon E, Marfany G, Wilson S, Ferguson DJ, Kingsman SM, Kingsman AJ. 1996. Structural determinants within the subunit protein of Ty1 virus-like particles. Mol Microbiol 22:667-679.

193. Garfinkel DJ, Hedge AM, Youngren SD, Copeland TD. 1991. Proteolytic processing of pol-TYB proteins from the yeast retrotransposon Ty1. J Virol 65:4573-4581.

194. Merkulov GV, Lawler JF, Eby Y, Boeke JD. 2001. Ty1 proteolytic cleavage sites are required for transposition: all sites are not created equal. J Virol 75:638-644.

195. Youngren SD, Boeke JD, Sanders NJ, Garfinkel DJ. 1988. Functional organization of the retrotransposon Ty from *Saccharomyces cerevisiae*: Ty protease is required for transposition. Mol Cell Biol 8:1421-1431.

196. Feng YX, Moore SP, Garfinkel DJ, Rein A. 2000. The genomic RNA in Ty1 virus-like particles is dimeric. J Virol 74:10819-10821.

197. Chapman KB, Bystrom AS, Boeke JD. 1992. Initiator methionine tRNA is essential for Ty1 transposition in yeast. Proc Natl Acad Sci U S A 89:3236-3240.

198. Keeney JB, Chapman KB, Lauermann V, Voytas DF, Astrom SU, von Pawel-Rammingen U, Bystrom A, Boeke JD. 1995. Multiple molecular determinants for retrotransposition in a primer tRNA. Mol Cell Biol 15:217-226.

199. Wilhelm M, Wilhelm FX, Keith G, Agoutin B, Heyman T. 1994. 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 22:4560-4565.

200. Friant S, Heyman T, Wilhelm ML, Wilhelm FX. 1996. Extended interactions between the primer tRNAi(Met) and genomic RNA of the yeast Ty1 retrotransposon. Nucleic Acids Res 24:441-449.

201. Cristofari G, Darlix J. 2009. A 5'-3' long range interaction in Ty1 RNA controls its reverse transcription and retrotransposition. EMBO 21:4368-4379.

202. Xu H, Boeke JD. 1990. Localization of sequences required in cis for yeast Ty1 element transposition near the long terminal repeats: analysis of mini-Ty1 elements. Mol Cell Biol 10:2695-2702.

203. Lange A, McLane LM, Mills RE, Devine SE, Corbett AH. 2010. Expanding the definition of the classical bipartite nuclear localization signal. Traffic 11:311-323.

204. Wilke CM, Maimer E, Adams J. 1992. The population biology and evolutionary significance of Ty elements in *Saccharomyces cerevisiae*. Genetica 86:155-173.

205. Bleykasten-Grosshans C, Friedrich A, Schacherer J. 2013. Genome-wide analysis of intraspecific transposon diversity in yeast. BMC Genomics 14:399.

206. Liti G, Peruffo A, James SA, Roberts IN, Louis EJ. 2005. Inferences of evolutionary relationships from a population survey of LTR-retrotransposons and telomeric-associated sequences in the *Saccharomyces* sensu stricto complex. Yeast 22:177-192.

207. Liti G, Carter DM, Moses AM, Warringer J, Parts L, James SA, Davey RP, Roberts IN, Burt A, Koufopanou V, Tsai IJ, Bergman CM, Bensasson D, O'Kelly MJ, van Oudenaarden A, Barton DB, Bailes E, Nguyen AN, Jones M, Quail MA, Goodhead I, Sims S, Smith F, Blomberg A, Durbin R, Louis EJ. 2009. Population genomics of domestic and wild yeasts. Nature 458:337-341.

208. Moore SP, Liti G, Stefanisko KM, Nyswaner KM, Chang C, Louis EJ, Garfinkel DJ. 2004. Analysis of a Ty1-less variant of *Saccharomyces paradoxus*: the gain and loss of Ty1 elements. Yeast 21:649-660.

209. Curcio MJ, Garfinkel DJ. 1991. Single-step selection for Ty1 element retrotransposition. Proc Natl Acad Sci U S A 88:936-940.

210. Curcio MJ, Hedge AM, Boeke JD, Garfinkel DJ. 1990. Ty RNA levels determine the spectrum of retrotransposition events that activate gene expression in *Saccharomyces cerevisiae*. Mol Gen Genet 220:213-221.

211. Elder RT, St John TP, Stinchcomb DT, Davis RW, Scherer S, Davis RW. 1981. 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 45 Pt 2:581-591.

212. Garfinkel DJ, Nyswaner K, Wang J, Cho JY. 2003. Post-transcriptional cosuppression of Ty1 retrotransposition. Genetics 165:83-99.

213. Matsuda E, Garfinkel DJ. 2009. Posttranslational interference of Ty1 retrotransposition by antisense RNAs. Proc Natl Acad Sci U S A 106:15657-15662.

214. Morillon A. 2008. A cryptic unstable transcript mediates transcriptional silencing of the Ty1 retrotransposon in *S. cerevisiae*. Genes and Development doi:papers://3BFEA9DA-C419-4A26-B21C-2F05A923CFC0/Paper/p1336:615-626.

215. Saha A, Mitchell JA, Nishida Y, Hildreth JE, Ariberre JA, Gilbert WV, Garfinkel DJ. 2015. A trans-dominant form of Gag restricts Ty1 retrotransposition and mediates copy number control. J Virol 89:3922-3938.

216. Blaszczyk L, Biesiada M, Saha A, Garfinkel DJ, Purzycka KJ. 2017. Structure of Ty1 Internally Initiated RNA Influences Restriction Factor Expression. Viruses 9.

217. Koning FA, Newman EN, Kim EY, Kunstman KJ, Wolinsky SM, Malim MH. 2009. Defining APOBEC3 expression patterns in human tissues and hematopoietic cell subsets. J Virol 83:9474-9485.

218. Refsland EW, Stenglein MD, Shindo K, Albin JS, Brown WL, Harris RS. 2010. Quantitative profiling of the full APOBEC3 mRNA repertoire in lymphocytes and tissues: implications for HIV-1 restriction. Nucleic Acids Res 38:4274-4284.

219. Bogerd HP, Cullen BR. 2008. Single-stranded RNA facilitates nucleocapsid: APOBEC3G complex formation. RNA 14:1228-1236.

220. Harris RS, Bishop KN, Sheehy AM, Craig HM, Petersen-Mahrt SK, Watt IN, Neuberger MS, Malim MH. 2003. DNA deamination mediates innate immunity to retroviral infection. Cell 113:803-809.

221. Mangeat B, Turelli P, Caron G, Friedli M, Perrin L, Trono D. 2003. Broad antiretroviral defence by human APOBEC3G through lethal editing of nascent reverse transcripts. Nature 424:99-103.

222. Zhang H, Yang B, Pomerantz RJ, Zhang C, Arunachalam SC, Gao L. 2003. The cytidine deaminase CEM15 induces hypermutation in newly synthesized HIV-1 DNA. Nature 424:94-98.

223. Yu X, Yu Y, Liu B, Luo K, Kong W, Mao P, Yu XF. 2003. Induction of APOBEC3G ubiquitination and degradation by an HIV-1 Vif-Cul5-SCF complex. Science 302:1056-1060.

224. Iwatani Y, Chan DS, Wang F, Maynard KS, Sugiura W, Gronenborn AM, Rouzina I, Williams MC, Musier-Forsyth K, Levin JG. 2007. Deaminase-independent inhibition of HIV-1 reverse transcription by APOBEC3G. Nucleic Acids Res 35:7096-7108.

225. Bishop KN, Verma M, Kim EY, Wolinsky SM, Malim MH. 2008. APOBEC3G inhibits elongation of HIV-1 reverse transcripts. PLoS Pathog 4:e1000231.

226. Marin M, Rose KM, Kozak SL, Kabat D. 2003. HIV-1 Vif protein binds the editing enzyme APOBEC3G and induces its degradation. Nat Med 9:1398-1403.

227. Sheehy AM, Gaddis NC, Malim MH. 2003. The antiretroviral enzyme APOBEC3G is degraded by the proteasome in response to HIV-1 Vif. Nat Med 9:1404-1407.

228. Stopak K, de Noronha C, Yonemoto W, Greene WC. 2003. HIV-1 Vif blocks the antiviral activity of APOBEC3G by impairing both its translation and intracellular stability. Mol Cell 12:591-601.

229. Schumacher AJ, Nissley DV, Harris RS. 2005. APOBEC3G hypermutates genomic DNA and inhibits Ty1 retrotransposition in yeast. Proc Natl Acad Sci U S A 102:9854-9859.

230. Schumacher AJ, Hache G, Macduff DA, Brown WL, Harris RS. 2008. The DNA deaminase activity of human APOBEC3G is required for Ty1, MusD, and human immunodeficiency virus type 1 restriction. J Virol 82:2652-2660.

231. Dutko JA, Schafer A, Kenny AE, Cullen BR, Curcio MJ. 2005. Inhibition of a yeast LTR retrotransposon by human APOBEC3 cytidine deaminases. Curr Biol 15:661-666.

232. Richardson SR, Narvaiza I, Planegger RA, Weitzman MD, Moran JV. 2014. APOBEC3A deaminates transiently exposed single-strand DNA during LINE-1 retrotransposition. Elife 3:e02008.

233. Aravind L, Koonin EV. 1998. The HD domain defines a new superfamily of metal-dependent phosphohydrolases. Trends Biochem Sci 23:469-472.

234. Li N, Zhang W, Cao X. 2000. Identification of human homologue of mouse IFN-gamma induced protein from human dendritic cells. Immunol Lett 74:221-224.

235. Berger A, Sommer AF, Zwarg J, Hamdorf M, Welzel K, Esly N, Panitz S, Reuter A, Ramos I, Jatiani A, Mulder LC, Fernandez-Sesma A, Rutsch F, Simon V, Konig R, Flory E. 2011. SAMHD1-deficient CD14+ cells from individuals with Aicardi-Goutieres syndrome are highly susceptible to HIV-1 infection. PLoS Pathog 7:e1002425.

236. Li S, Wang L, Berman M, Kong YY, Dorf ME. 2011. Mapping a dynamic innate immunity protein interaction network regulating type I interferon production. Immunity 35:426-440.

237. Powell RD, Holland PJ, Hollis T, Perrino FW. 2011. Aicardi-Goutieres syndrome gene and HIV-1 restriction factor SAMHD1 is a dGTP-regulated deoxynucleotide triphosphohydrolase. J Biol Chem 286:43596-43600.

238. Goldstone DC, Ennis-Adeniran V, Hedden JJ, Groom HC, Rice GI, Christodoulou E, Walker PA, Kelly G, Haire LF, Yap MW, de Carvalho LP, Stoye JP, Crow YJ,

Taylor IA, Webb M. 2011. HIV-1 restriction factor SAMHD1 is a deoxynucleoside triphosphate triphosphohydrolase. Nature 480:379-382.

239. Lahouassa H, Daddacha W, Hofmann H, Ayinde D, Logue EC, Dragin L, Bloch N, Maudet C, Bertrand M, Gramberg T, Pancino G, Priet S, Canard B, Laguette N, Benkirane M, Transy C, Landau NR, Kim B, Margottin-Goguet F. 2012. SAMHD1 restricts the replication of human immunodeficiency virus type 1 by depleting the intracellular pool of deoxynucleoside triphosphates. Nat Immunol 13:223-228.

240. Diamond TL, Roshal M, Jamburuthugoda VK, Reynolds HM, Merriam AR, Lee KY, Balakrishnan M, Bambara RA, Planelles V, Dewhurst S, Kim B. 2004. Macrophage tropism of HIV-1 depends on efficient cellular dNTP utilization by reverse transcriptase. J Biol Chem 279:51545-51553.

241. Mangeat B, Gers-Huber G, Lehmann M, Zufferey M, Luban J, Piguet V. 2009. HIV-1 Vpu neutralizes the antiviral factor Tetherin/BST-2 by binding it and directing its beta-TrCP2-dependent degradation. PLoS Pathog 5:e1000574.

242. Belshan M, Kimata JT, Brown C, Cheng X, McCulley A, Larsen A, Thippeshappa R, Hodara V, Giavedoni L, Hirsch V, Ratner L. 2012. Vpx is critical for SIVmne infection of pigtail macaques. Retrovirology 9:32.

243. Gibbs JS, Lackner AA, Lang SM, Simon MA, Sehgal PK, Daniel MD, Desrosiers RC. 1995. Progression to AIDS in the absence of a gene for vpr or vpx. J Virol 69:2378-2383.

244. Hirsch VM, Sharkey ME, Brown CR, Brichacek B, Goldstein S, Wakefield J, Byrum R, Elkins WR, Hahn BH, Lifson JD, Stevenson M. 1998. Vpx is required

for dissemination and pathogenesis of SIV(SM) PBj: evidence of macrophage-dependent viral amplification. Nat Med 4:1401-1408.

245. Brandariz-Nunez A, Valle-Casuso JC, White TE, Laguette N, Benkirane M, Brojatsch J, Diaz-Griffero F. 2012. Role of SAMHD1 nuclear localization in restriction of HIV-1 and SIVmac. Retrovirology 9:49.

246. Lilly F. 1970. Fv-2: identification and location of a second gene governing the spleen focus response to Friend leukemia virus in mice. J Natl Cancer Inst 45:163-169.

247. Hilditch L, Matadeen R, Goldstone DC, Rosenthal PB, Taylor IA, Stoye JP. 2011. Ordered assembly of murine leukemia virus capsid protein on lipid nanotubes directs specific binding by the restriction factor, Fv1. Proc Natl Acad Sci U S A 108:5771-5776.

248. Elis E, Ehrlich M, Prizan-Ravid A, Laham-Karam N, Bacharach E. 2012. p12 tethers the murine leukemia virus pre-integration complex to mitotic chromosomes. PLoS Pathog 8:e1003103.

249. Stremlau M, Owens CM, Perron MJ, Kiessling M, Autissier P, Sodroski J. 2004. The cytoplasmic body component TRIM5alpha restricts HIV-1 infection in Old World monkeys. Nature 427:848-853.

250. Sayah DM, Sokolskaja E, Berthoux L, Luban J. 2004. Cyclophilin A retrotransposition into TRIM5 explains owl monkey resistance to HIV-1. Nature 430:569-573.

251. Keckesova Z, Ylinen LM, Towers GJ. 2004. The human and African green monkey TRIM5alpha genes encode Ref1 and Lv1 retroviral restriction factor activities. Proc Natl Acad Sci U S A 101:10780-10785.

252. Perez-Caballero D, Hatziioannou T, Yang A, Cowan S, Bieniasz PD. 2005. Human tripartite motif 5alpha domains responsible for retrovirus restriction activity and specificity. J Virol 79:8969-8978.

253. Si Z, Vandegraaff N, O'Huigin C, Song B, Yuan W, Xu C, Perron M, Li X, Marasco WA, Engelman A, Dean M, Sodroski J. 2006. Evolution of a cytoplasmic tripartite motif (TRIM) protein in cows that restricts retroviral infection. Proc Natl Acad Sci U S A 103:7454-7459.

254. Ylinen LM, Keckesova Z, Webb BL, Gifford RJ, Smith TP, Towers GJ. 2006. Isolation of an active Lv1 gene from cattle indicates that tripartite motif protein-mediated innate immunity to retroviral infection is widespread among mammals. J Virol 80:7332-7338.

255. Fletcher AJ, Hue S, Schaller T, Pillay D, Towers GJ. 2010. Hare TRIM5alpha restricts divergent retroviruses and exhibits significant sequence variation from closely related lagomorpha TRIM5 genes. J Virol 84:12463-12468.

256. Schaller T, Hue S, Towers GJ. 2007. An active TRIM5 protein in rabbits indicates a common antiviral ancestor for mammalian TRIM5 proteins. J Virol 81:11713-11721.

257. Sawyer SL, Wu LI, Emerman M, Malik HS. 2005. Positive selection of primate TRIM5alpha identifies a critical species-specific retroviral restriction domain. Proc Natl Acad Sci U S A 102:2832-2837.

258. Song B, Gold B, O'Huigin C, Javanbakht H, Li X, Stremlau M, Winkler C, Dean M, Sodroski J. 2005. The B30.2(SPRY) domain of the retroviral restriction factor TRIM5alpha exhibits lineage-specific length and sequence variation in primates. J Virol 79:6111-6121.

259. Stremlau M, Perron M, Welikala S, Sodroski J. 2005. Species-specific variation in the B30.2(SPRY) domain of TRIM5alpha determines the potency of human immunodeficiency virus restriction. J Virol 79:3139-3145.

260. Ohkura S, Yap MW, Sheldon T, Stoye JP. 2006. All three variable regions of the TRIM5alpha B30.2 domain can contribute to the specificity of retrovirus restriction. J Virol 80:8554-8565.

261. Nisole S, Stoye JP, Saib A. 2005. TRIM family proteins: retroviral restriction and antiviral defence. Nat Rev Microbiol 3:799-808.

262. Colgan J, Yuan HE, Franke EK, Luban J. 1996. Binding of the human immunodeficiency virus type 1 Gag polyprotein to cyclophilin A is mediated by the central region of capsid and requires Gag dimerization. J Virol 70:4299-4310.

263. Franke EK, Yuan HE, Luban J. 1994. Specific incorporation of cyclophilin A into HIV-1 virions. Nature 372:359-362.

264. Perron MJ, Stremlau M, Lee M, Javanbakht H, Song B, Sodroski J. 2007. The human TRIM5alpha restriction factor mediates accelerated uncoating of the N-tropic murine leukemia virus capsid. J Virol 81:2138-2148.

265. Stremlau M, Perron M, Lee M, Li Y, Song B, Javanbakht H, Diaz-Griffero F, Anderson DJ, Sundquist WI, Sodroski J. 2006. Specific recognition and

accelerated uncoating of retroviral capsids by the TRIM5alpha restriction factor. Proc Natl Acad Sci U S A 103:5514-5519.

266. Fricke T, White TE, Schulte B, de Souza Aranha Vieira DA, Dharan A, Campbell EM, Brandariz-Nunez A, Diaz-Griffero F. 2014. MxB binds to the HIV-1 core and prevents the uncoating process of HIV-1. Retrovirology 11:68.

267. Fribourgh JL, Nguyen HC, Matreyek KA, Alvarez FJ, Summers BJ, Dewdney TG, Aiken C, Zhang P, Engelman A, Xiong Y. 2014. Structural insight into HIV-1 restriction by MxB. Cell Host Microbe 16:627-638.

268. Kong J, Xu B, Wei W, Wang X, Xie W, Yu XF. 2014. Characterization of the amino-terminal domain of Mx2/MxB-dependent interaction with the HIV-1 capsid. Protein Cell 5:954-957.

269. Haller O, Staeheli P, Schwemmle M, Kochs G. 2015. Mx GTPases: dynamin-like antiviral machines of innate immunity. Trends Microbiol 23:154-163.

270. Kane M, Yadav SS, Bitzegeio J, Kutluay SB, Zang T, Wilson SJ, Schoggins JW, Rice CM, Yamashita M, Hatziioannou T, Bieniasz PD. 2013. MX2 is an interferon-induced inhibitor of HIV-1 infection. Nature 502:563-566.

271. Goujon C, Moncorge O, Bauby H, Doyle T, Barclay WS, Malim MH. 2014. Transfer of the amino-terminal nuclear envelope targeting domain of human MX2 converts MX1 into an HIV-1 resistance factor. J Virol 88:9017-9026.

272. Liu Z, Pan Q, Ding S, Qian J, Xu F, Zhou J, Cen S, Guo F, Liang C. 2013. The interferon-inducible MxB protein inhibits HIV-1 infection. Cell Host Microbe 14:398-410.

273. Palmarini M, Hallwirth C, York D, Murgia C, de Oliveira T, Spencer T, Fan H. 2000. Molecular cloning and functional analysis of three type D endogenous retroviruses of sheep reveal a different cell tropism from that of the highly related exogenous jaagsiekte sheep retrovirus. J Virol 74:8065-8076.

274. Palmarini M, Sharp JM, de las Heras M, Fan H. 1999. Jaagsiekte sheep retrovirus is necessary and sufficient to induce a contagious lung cancer in sheep. J Virol 73:6964-6972.

275. Palmarini M, Sharp JM, Lee C, Fan H. 1999. In vitro infection of ovine cell lines by Jaagsiekte sheep retrovirus. J Virol 73:10070-10078.

276. Murcia PR, Arnaud F, Palmarini M. 2007. The transdominant endogenous retrovirus enJS56A1 associates with and blocks intracellular trafficking of Jaagsiekte sheep retrovirus Gag. J Virol 81:1762-1772.

Figures

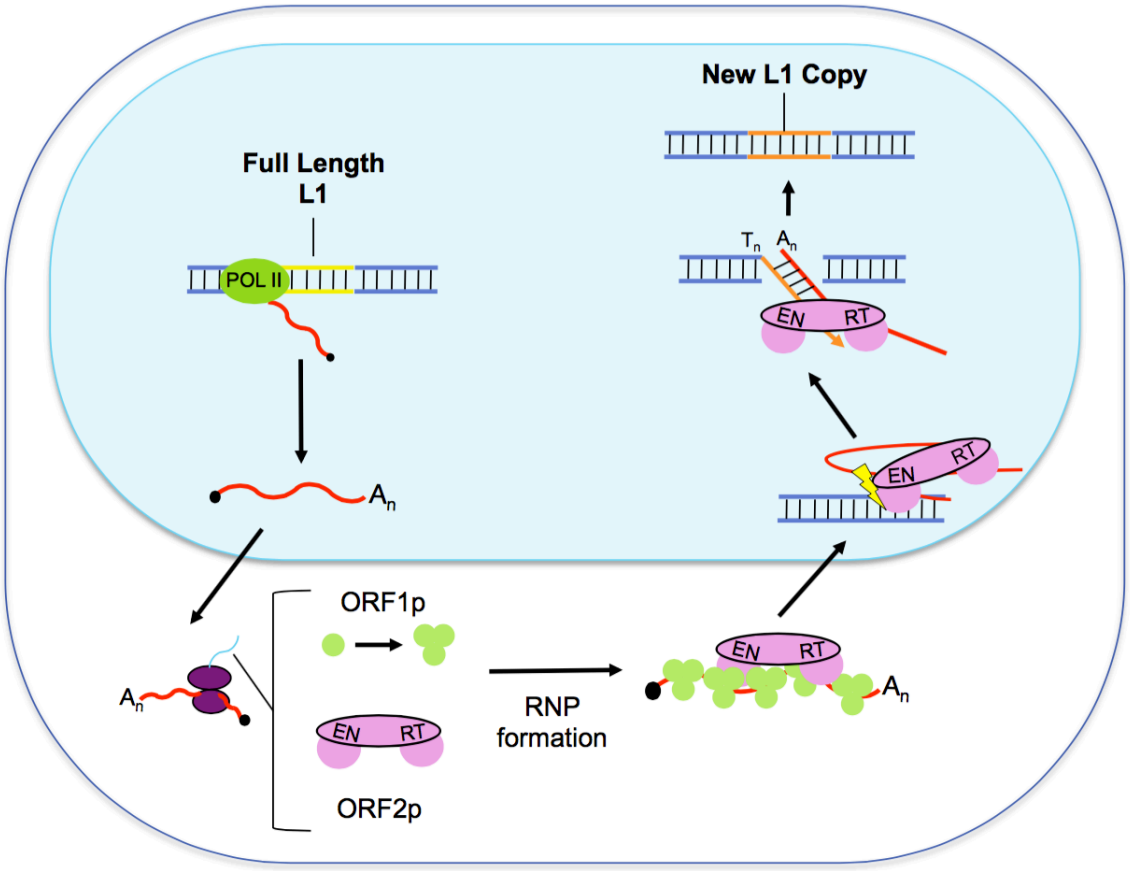

Figure 1-1. Replication of L1. RNA Pol II transcribes the L1 RNA which is then translated into several molecules of ORF1p and as few as one molecule of ORF2p. These proteins associate with the L1 RNA to form the L1 RNP. The RNP contains at least a ORF1p trimer and ORF2p monomer. The L1 RNP enters the nucleus where endonuclease activity of ORF2p cleaves target site genomic DNA to generate a free 3' end, which is used to prime reverse transcription. The L1 cDNA is then integrated in the host DNA to form a new copy of L1.

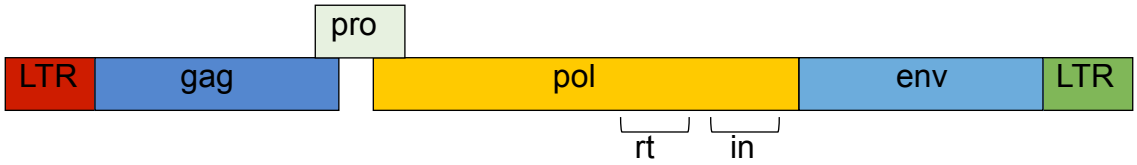

Figure 1-2. A HERV element.

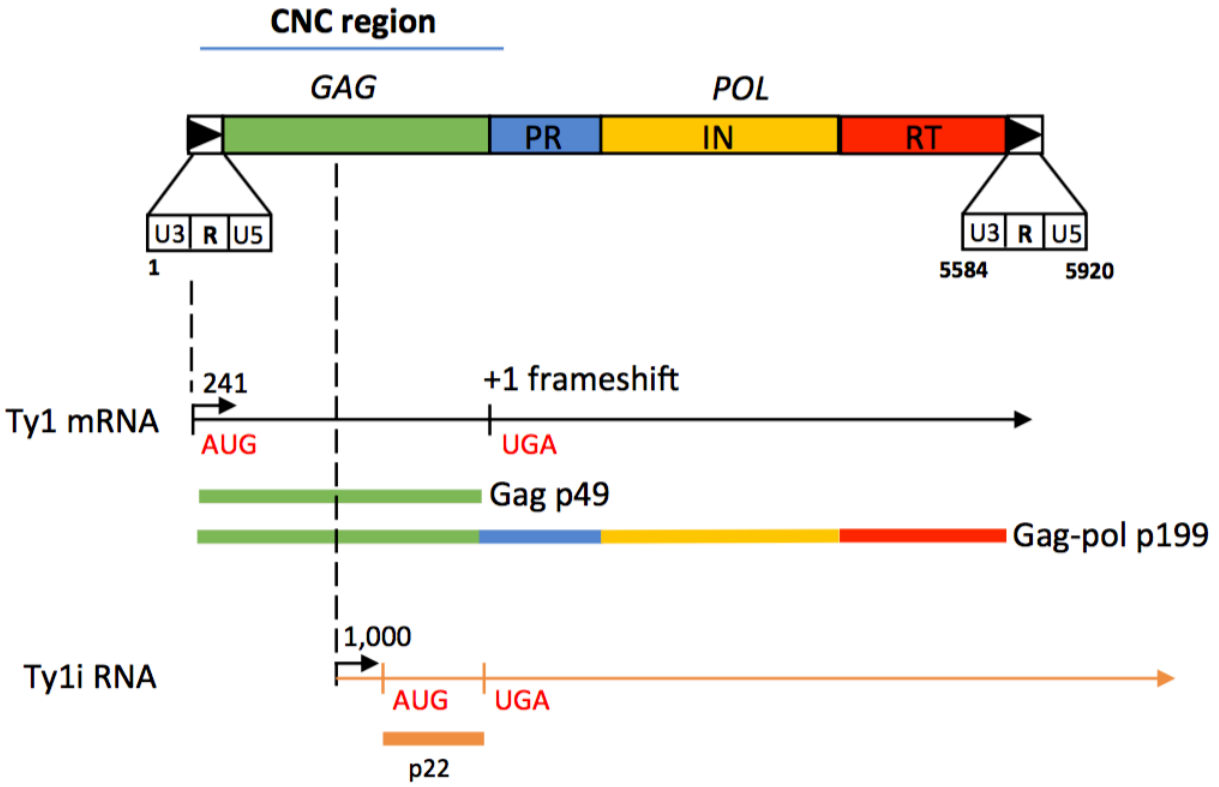

Figure 1-3. Functional organization of Ty1 and its gene products. Ty1 contains GAG and POL ORFs bracketed by LTRS (black triangles). The CNC region comprises of mostly the 5'LTR and GAG. U3, R and U5 are domains in the LTRs required for Ty1 mRNA expression and accurate reverse transcription. POL is comprised of protease (PR), integrase (IN) and reverse transcriptase (RT) proteins. Transcription of the Ty1 mRNA starts at nucleotide 241 in the R region of the 5'LTR and ends in R of the 3'LTR. Its main translation products are Gag-p49 (green) and Gag-Pol-p199, a fusion protein made of PR (blue), IN (orange) and RT (red). Transcription of Ty1i RNA initiates at position 1000 (orange), still in GAG. p22 is translated from Ty1i RNA and is identical to the C-terminal half of Gag-p49.

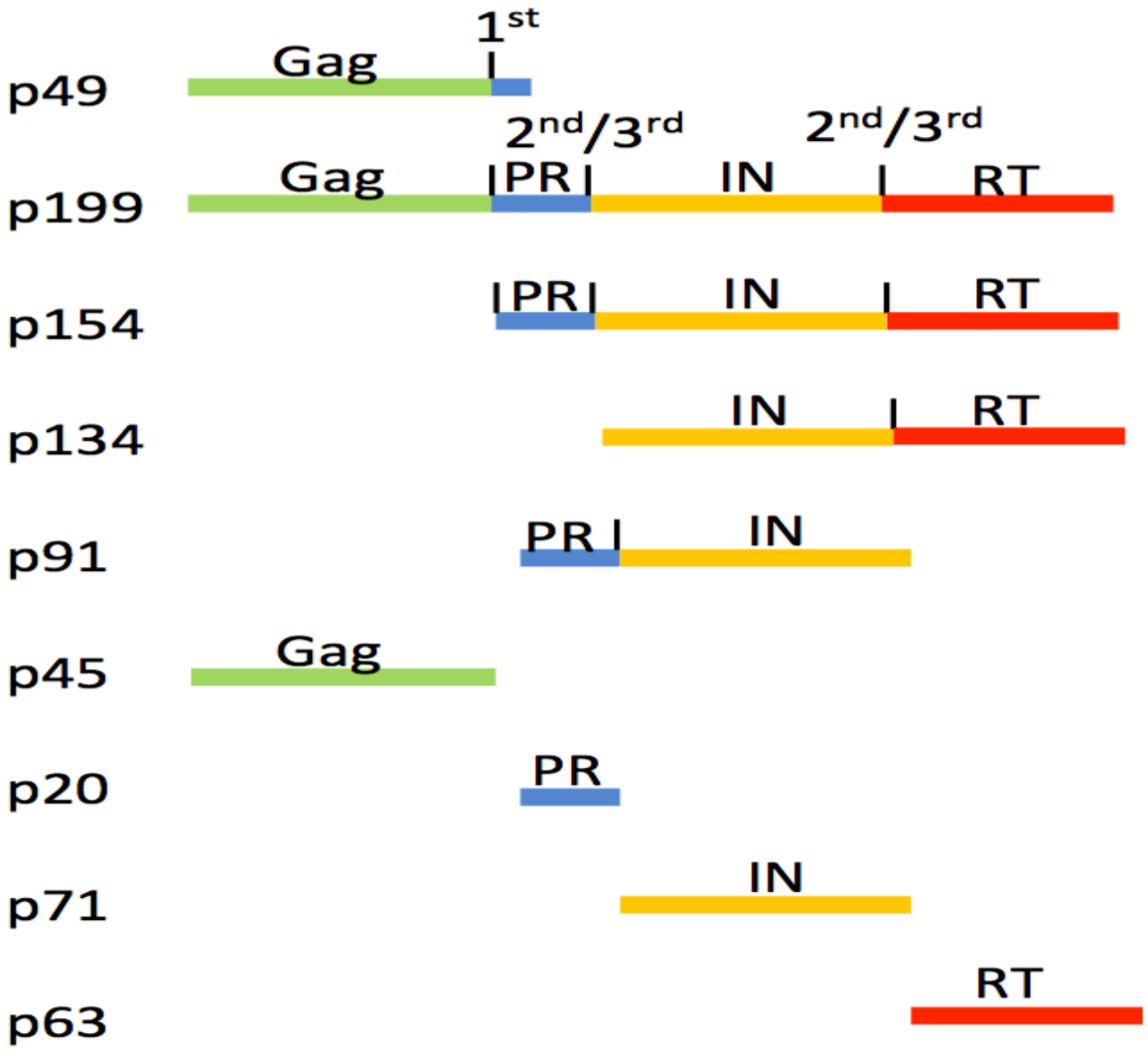

Figure 1-4. Ty1 protein processing during VLP maturation. Gag-p49 and Gag-Pol-p199 are processed by Ty1 protease (PR) at the 1st cleavage site to form Gag-p45 and PR-IN-RT- p154. The 2nd and 3rd sites are cleaved in no specific order and finally gives rise to mature Ty1 proteins PR-p20, IN-71 and RT-p63.

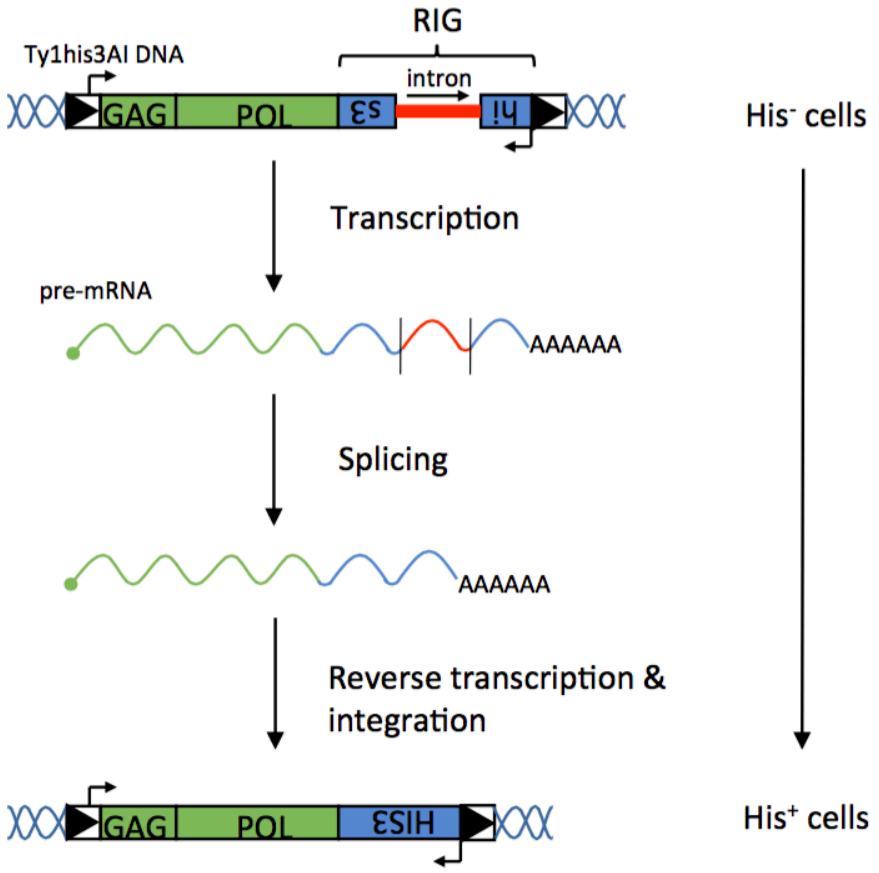

Figure 1-5. Ty1*his3-AI*. A Ty1 element tagged with a retrotransposition indicator gene (RIG) is shown. A HIS3 gene interrupted with an artificial intron (AI) is inserted in the POL sequence of Ty1. Following transcription of the Ty1*his3-AI* mRNA, the *AI* can be spliced out to generate Ty1HIS3 RNA. The *A*I in *HIS3-AI* transcripts made from the natural HIS3 promoter cannot be spliced because it is the wrong orientation. The Ty1HIS3 RNA can be packaged in VLPs, reverse transcribed, and the cDNA copy can be integrated into the host DNA. This restores a functional *HIS3* gene and the cells become prototrophs for histidine biosynthesis.

Table 1-1 Anti-retroviral restriction factors

Restriction factor	Organism	Target virus	Mechanism
APOBEC3G	Human	HIV-1	Deaminates Cytosine to Uracil in viral cDNA
SAMHD1	Human	HIV-1	Reduces cDNA synthesis by lowering dNTP pools
[a]Fv1	Mouse	MLV	Prevents viral integration into host genome
[a]TRIM5α	Non-human primates	HIV-1	Accelerates viral uncoating, inhibition of reverse transcription
[a]MX2	Human	HIV-1	Prevents viral uncoating, blocks PIC nuclear entry
[a]enJS56A1	Sheep	JSRV	Interferes with virus assembly

[a] capsid-binding restriction factors

CHAPTER 2

A *TRANS*-DOMINANT FORM OF GAG RESTRICTS TY1 RETROTRANSPOSITION AND MEDIATES COPY NUMBER CONTROL[1]

[1] Saha*, A., Mitchell*, J.A., Hildreth, J.E., Ariberre, J.A., Gilbert, W.V., and D.J. Garfinkel. 2015. *Journal of Virology*. 89(7):3922-38. doi: 10.1128/JVI.03060-14. *These authors contributed equally to this work. Reprinted here with permission from the publisher.

Abstract

Saccharomyces cerevisiae and *S. paradoxus* lack the conserved RNA interference pathway and utilize a novel form of copy number control (CNC) to inhibit Ty1 retrotransposition. Although noncoding transcripts have been implicated in CNC, here we present evidence that a truncated form of the Gag capsid protein (p22) or its processed form (p18) is necessary and sufficient for CNC and likely encoded by Ty1 internal transcripts. Coexpression of p22/p18 and Ty1 decreases mobility more than 30,000-fold. p22/p18 cofractionate with Ty1 virus-like particles (VLPs) and affect VLP yield, protein composition and morphology. Although p22/p18 and Gag colocalize in the cytoplasm, p22/p18 disrupt sites used for VLP assembly. GST affinity pull-downs also suggest that p18 and Gag interact. Therefore, this intrinsic Gag-like restriction factor confers CNC by interfering with VLP assembly and function, and expands the strategies used to limit retroelement propagation.

Importance

Retrotransposons dominate the chromosomal landscape in many eukaryotes, can cause mutations by insertion or genome rearrangement and are evolutionarily related to retroviruses such as HIV. Thus, understanding factors that limit transposition and retroviral replication are fundamentally important. The present manuscript describes a retrotransposon-encoded restriction protein derived from the capsid gene of the yeast Ty1 element that disrupts virus-like particle assembly in a dose dependent manner. This form of copy number control acts as a molecular rheostat, allowing high levels of

retrotransposition when few Ty1 elements are present and inhibiting transposition as copy number increases. Thus, yeast and Ty1 have co-evolved a form of copy number control that is beneficial to both “host and parasite”. To our knowledge, this is the first Gag-like retrotransposon restriction factor described in the literature and expands the ways restriction proteins modulate retroelement replication.

Introduction

Retrovirus-like retrotransposons and their long terminal repeat (LTR) derivatives inhabit the genomes of many organisms, including the budding yeast *Saccharomyces cerevisiae* and its closest relative *S. paradoxus*. The Ty1 family is active and related to the LTR retrotransposons Ty2-Ty5 in budding yeast (1). Variations in Ty1 copy number can be attributed to the relative rates of transposition, loss by LTR-LTR recombination or additional types of genome rearrangements, all of which can impact fitness (2-5). Ty1 resembles retroviruses in genome organization and replication (1). These elements consist of two overlapping open reading frames, *GAG* and *POL*, which are flanked by LTRs. Ty1 genomic RNA is translated or packaged as a dimer into virus-like particles (VLPs). The primary translation products are Gag (p49) and Gag-Pol (p199) precursors, the latter resulting from a +1 ribosomal frameshift during translation. Mature Gag (p45) is the major structural component of VLPs. *POL* encodes the enzymes required for proteolytic processing of Gag and Gag-Pol (protease; PR), cDNA integration (integrase; IN), and reverse transcription (reverse transcriptase; RT). Ty1 and Ty3 VLPs assemble within cytoplasmic foci, termed retrosomes or T-bodies, which contain Ty proteins and RNA (6-9). Once VLPs undergo maturation via the action of PR, Ty1 genomic RNA is

reverse transcribed to form a linear cDNA. A protein/DNA complex minimally containing Ty1 cDNA and IN are imported into the nucleus, where integration usually occurs near genes transcribed by RNA polymerase III.

S. cerevisiae and *S. paradoxus* laboratory strains and natural isolates contain fewer than 40 copies of Ty1 per haploid genome and several strains contain few if any elements (5, 10-14). Although budding yeast genomes characterized to date tend to have low Ty1 copy numbers, fertile *S. cerevisiae* strains containing more than 100 Ty1 insertions have been created artificially by numerous rounds of induction of a multi-copy plasmid containing an active Ty1 element (Ty1H3) fused to the *GAL1* promoter (pGTy1) (15, 16). Host cofactor and restriction genes involved in modulating Ty1 retrotransposition are diverse and encompass different steps in the replication cycle, ranging from transcription to integration site preference (17-21). For example, *SPT3* is required for transcription of full-length Ty1 mRNA (22) and encodes a component of the SAGA chromatin-remodeling complex (23), and *XRN1* is an important Ty1 cofactor implicated in transcription (24, 25), assembly of functional VLPs (7, 8), and encodes a 5'- 3' exonuclease required for mRNA turnover (26).

Transposon-derived regulatory factors are critically important for keeping transposition at a low level. Forms of RNA interference affect the level or utilization of transposon mRNA, and the source of the interfering RNAs can be the transposons themselves (27). A unique form of copy number control (CNC) minimizes Ty1 transposition in *S. cerevisiae* and *S. paradoxus* (28) in the absence of *dicer* and *argonaute* genes that comprise a functional RNAi pathway in a distant species *S. castellii* (29, 30). Ty1 CNC is defined by a copy number dependent decrease in Ty1

retrotransposition and is especially robust in a “Ty1-less” strain of *S. paradoxus* (28) that may have lost Ty1 elements by LTR-LTR recombination (13). Ty1 CNC acts posttranslationally and *in trans*, can be overcome by pGTy1 expression, and is characterized by lower levels of mature IN and reverse transcripts (28, 31, 32). Reduced levels of endogenous Ty1 IN, PR, cDNA, and VLPs are also present in *S. cerevisiae* (33-35), which displays CNC (28). These results suggest that Ty1 produces a titratable factor that inhibits transposition in a copy dependent manner.

Ty1 antisense (Ty1AS) RNAs have been implicated in silencing Ty1 expression by alterations in chromatin function (24) or when RNAi is reconstituted in *S. cerevisiae* (30). We reported evidence suggesting that Ty1AS RNAs interfere with Ty1 transposition posttranslationally (31). Inhibition occurs in a copy number dependent manner and the antisense transcripts map to a region within *GAG* that confers CNC (28). Deleting the common 3’ end of the antisense transcripts abolishes CNC and decreases the level of Ty1AS RNAs. However, ectopic expression of individual antisense transcripts does not restore CNC, suggesting that either multiple antisense transcripts or additional factors are required (24, 31). Also, nuclease protection and structural probing analyses suggest that although Ty1AS RNAs specifically associate with VLPs from CNC^+ strains, these transcripts are not packaged into VLPs and do not interact with Ty1 mRNA (32).

Here, further characterization of the minimal Ty1 sequence that confers CNC has led to the discovery of p22, an N-terminal truncated form of Gag that is likely encoded by an internally initiated Ty1 mRNA. Importantly, p22 is both necessary and sufficient for CNC. Coexpression of p22 and Ty1 interferes with assembly of functional VLPs,

which is conceptually similar to the inhibition displayed by Gag-like restriction factors derived from endogenous retroviruses in mammals (36, 37).

Materials and Methods

Genetic techniques, media and strain construction

Strains are listed in Table 2-1. Strains repopulated with Ty1 elements were obtained following pGTy1 induction as described previously (28). Standard yeast genetic and microbiological procedures were used in this work (38).

Plasmids

All nucleotide information used here corresponds to Ty1H3 sequence (39) (Genbank M18706.1). pGPOLΔ derivatives of pGTy1 were generated by digestion with *Bgl*II and ligation. pBJM78, pBJM79 and pBDG1595 were constructed by overlap PCR using flanking primers (Ty335F, 5'-TGGTAGCGCCTGTGCTTCGGTTAC-3'; TyRP1, 5'-CATTGATAGTCAATAGCACTAGACC-3') and overlapping primers (DELC1071b, 5'-GGTATCAGATTCATTTTTTCAATACTTTTGGAAAGAATTTTC-3'; DELC1071c, 5'-GTATTGAAAAAATGAATCTGATACCCAAGAGGCAAACGAC-3'; ADDA1303b, 5'-GAACAGTTCATGCGACTGTCATATTTAGATGTCGATGACGTG-3'; ADDA1303c, 5'-CTAAATATGACAGTCGCATGAACTGTTCTTAGATATCCATGC-3; B-AUG1Ala-R, 5'-AAAGAATTTTCGCGATATCCGTATAATCAACG-3'; C-AUG1Ala-F, 5'-GGATATCGCGAAAATTCTTTCCAAAAGTATTG-3'; B-AUG2Ala-R, 5'-TATCAGATTGCGCTTTTTCAATACTTTTGG-3'; C-AUG2Ala-F, 5'-TGAAAAAGCGCAATCTGATACCCAAGAGGC-3') and Ty1H3 as template. Final PCR products were subcloned into pGPOLΔ using *Bst*XI and *Bgl*II restriction sites. Plasmid

pBDG1534 was generated from plasmid pBDG606 (pGTy1*his3-AI*/Cen-*URA3*) (18) by replacing the *URA3* marker for *TRP1*. Briefly, *TRP1* was amplified from BY4742 with primers containing flanking *URA3* sequence (20718uratrpfwd, 5'-ATGTCGAAAGCTACATATAAGGAACGTGCTGCTACTCATCAATTCGGTCGAAAAAAGAAA-3'; 20916uratrprev, 5'-AGCTTTTTCTTTCCAATTTTTTTTTTTTCGTCATTATAATATGCTTGCTTTTCAAAAGGC-3') and the PCR product was cotransformed into yeast with pBDG606 linearized within *URA3* with *Apa*I. Transformants were selected on SC–Trp and plasmids were verified phenotypically and by restriction mapping. Plasmid pBDG1565 was created by PCR-amplifying Ty1 *GAG* coding sequence (nt 1038-1613; EcoRIstartF, 5'-CATGTTTCGAATTCATGAAAATTCTTTCCAAAAGTATTG-3'; Xho1stopR, 5'-CATGTTTCCTCGAGTTAGTAAGTTTCTGGCCTAAGATGAAG-3) using Ty1H3 as a template and cloning into pYES2 (Life Technologies, Carlsbad, CA) using *Eco*RI and *Xho*I. Plasmid pBDG1568 was made in a similar manner as pBDG1565, except an initial PCR step was performed to insert V5 coding sequence (underlined) in-frame between Ty1 *GAG* nt 1442 and 1443 (V51442b, 5'-CGTAGAATCGAGACCGAGGAGAGGGTTAGGGATAGGCTTACCTATAACTTTGGGTTTGGT-3'; V51442c, 5'-GGTAAGCCTATCCCTAACCCTCTCCTCGGTCTCGATTCTACGGCTCGGAATCCTCAAAAA-3'). For plasmid pBDG1571, *GAG* coding sequence cloned into pYES2 ended at nt 1496 (1496XhoI, 5'-CATGTTTCCTCGAGTTAGTGAGCCCTGGCTGTTTCG-3'). The GAG*PR mutation was created by mutating the Gag-PR cleavage site (RAHNVS) to AAGSAA (40) using overlapping primers (Gag*PRb, 5'-

AGCCGCTGCTGGATCCGCTGCTACATCTAATAACTCTCCCAGC-3'; Gag*PRc, 5'-GATGTAGCAGCGGATCCAGCAGCGGCTGTTTTCGATTTCGAAT-3'). To construct the *GAL1*-promoted GST-p18 protein fusion, the coding region for p18 (1038-1496) was amplified with *Xba*I and *Hind*III primer sets (1038XbaI, 5'-CTAGTCTAGACATGAAAATTCTTTCCAAAAGTATTG-3'; 1496XbaI, 5'-CCCAAGCTTTTAGTGAGCCCTGGCTGTTTTCG-3'). The PCR fragment was cloned into pEG(KT) (41) yielding pBDG1576. All plasmids generated by PCR cloning were verified by DNA sequencing. Phusion DNA polymerase, T4 DNA ligase, and restriction enzymes were obtained from New England BioLabs (Ipswich, MA).

Random mutagenesis and gap repair

The Ty1 CNC region was mutagenized by amplification with Taq DNA polymerase (ThermoFisher Scientific, Waltham, MA) using PCR forward primer FP1 (5'-CTCCGTGCGTCCTCGTCTTCACC-3') and reverse primer RP1 (5'-CATTGATAGTCAATAGCACTAGACC-3'). Gel purified PCR product was cotransformed into DG2196 along with a multicopy pGTy1 plasmid gapped with *Xho*I and *Bst*EII. Gap-repaired transformants were selected on SC-Ura medium. Plasmids recovered from the CNC^- strains were introduced into DG2196 to verify loss of CNC and then subjected to DNA sequencing. Aligning mutant sequences with Ty1H3 using ClustalW2 identified point mutations.

Ribosome footprint profiling analysis of chromosomal Ty1 elements

Samples were prepared and ribosome footprint profiling (Ribo-seq) was performed as previously described (42). Briefly, *S. cerevisiae* strain Sigma 1278b (YWG025; MATa *ura3 leu2 trp1 his3*) was grown to OD600 ~1.0-1.1 at 30˚C in YPD media, spun down,

and resuspended in pre-warmed YPA (no glucose) media. After 3 hr in YPA media, cycloheximide was added to a final concentration of 0.1mg/ml and cells harvested by centrifugation. Cells were lysed in 1xPLB (20 mM HEPES-KOH, pH 7.4, 2 mM Mg(OAc)2, 100 mM KOAC, 1% Triton-X 100, 0.1 mg/mL cycloheximide, 3 mM DTT), and libraries prepared essentially as described (43). Reads were mapped to Ty1H3 using the STAR RNA-seq alignment (44), allowing for zero mismatches. The Ty1 reads represent a composite of all Ty1 elements in the genome, including partial elements such as solo LTRs. No attempts were made to sort multiple mapping reads. The abundance of 5'end reads are displayed over Ty1 using custom scripts available upon request. Libraries used for these analyses include NCBI GEO accession numbers SRX264202 and SRX366898 (Sigma Ribo-seq).

Isolation of cDNA clones

A *S. cerevisiae* cDNA expression library fused to the *GAL1* promoter on a centromere-based *URA3* vector (45) was introduced into DG2196. Approximately 5,000 primary transformants were replica plated to SC-Ura + 2% galactose and incubated for 3 days at 30^{o}C. Colonies were then replica plated to SC-His-Ura and Ty1*HIS3* papillae were scored after incubation for 3 days at 30^{o}C. Galactose induction was performed at a suboptimal temperature for transposition to sensitize the screen, since induction at 22^{o}C resulted in too many Ty1*HIS3* papillae. Most transformants yielded about 5 Ty1*HIS3* mobility events/colony. Thirty-three transformants that had a lower level of Ty1 mobility were retested. Plasmids from 7 transformants were recovered in *E. coli* and sequenced from their 5' and 3' ends using *GAL1*- and vector-specific primers. *CCW12* (cell wall mannoprotein), *MSS4* (phosphatidylinositol-4-phosphate 5-kinase), *MRH1* (membrane

protein), *RGD1* (GTPase-activating protein), *TIR1* (cell wall mannoprotein), and *WHI5* (repressor of G1 transcription) were recovered as partial or complete cDNA clones, and were not studied further. One clone (pBDG1354) contained Ty1 sequences from nt 1042-5889 and conferred a strong *trans*-dominant negative inhibition of Ty1*his3-AI* mobility.

RNA isolation

Cultures were grown at 22^{o}C for 24 hr in SC or YEPD media. Total RNA was extracted using the MasterPure yeast RNA purification kit (Epicenter Biotechnologies, Madison, WI) following the manufacturer's protocol with minor modifications; 400 μl RNA extraction reagent and 200 μl of MPC protein reagent was used instead of 300 μl and 160 μl respectively. Poly(A)$^{+}$ RNA was isolated from ~ 250 μg total RNA using the NucleoTrap mRNA purification kit (Clontech, Mountainview, CA).

Northern blotting

RNA was resolved on a 1.2% agarose-formaldehyde gel at 120 V for 2 hr and blotted onto Hybond-XL N (GE Healthcare, Little Chalfont, United Kingdom). Riboprobes were transcribed *in vitro* from Ty1 *GAG* and *ACT1* coding sequence using a MAXIscript kit (Life Technologies) and uniformly labeled with [α-^{32}P] UTP (3,000 Ci/mmol; Perkin Elmer, Waltham MA). Hybridization and phosphorimage analysis was carried out as previously described (19, 28).

5' RACE

200 ng of poly(A)$^{+}$ RNA was used for synthesis of cDNA library using the SMARTer PCR cDNA Amplification Kit (Clontech). This method is 5' cap independent and the

library included cDNA from all poly(A)$^+$ transcripts. Ty1 specific cDNA was amplified with the gene specific primer

GSP1_3389 (5'-GACATGGGAGCAAGTAAAGGAAC-3') and the universal primer mix from the supplier. RACE products were resolved on a 1% agarose gel. Gel purified DNA fragments were TA-cloned into pCR2.1-TOPO vector (Life Technologies). Plasmid DNA was subjected to DNA sequencing using Ty1 specific sequencing primer (Ty1new2rev; 5'GAGAATCATTCTTCTCATCACTCG-3').

qPCR

The number of Ty1A1123G transposition events in strain DG3798 was estimated by qPCR. Strains DG2196 (Ty1*his3-AI*), DG2512 (Ty1*his3-AI* + 9 additional Ty1 elements), and DG2511 (Ty1*his3-AI* + 12 additional Ty1 elements) were used as standards, based on results from Southern analysis (28) (Ahn and Garfinkel, unpublished results). Duplicate samples were subjected to qPCR using IQ SYBR Green Supermix (Bio-Rad Laboratories, Hercules CA) and two different primer pairs from Ty1 *POL* (4681F, 5'-GAAATTCAATATGACATACTTGGC-3'; + 4851R, 5'-GTTCATCCTGGTCTATATATAAAGA-3'; 3251F, 5'-GAGAAGTTGACCCCAACATATCTG-3'; + 3480R, 5'-TGTATGATTAGTCTCATTTTCAC-3').

Ty1*his3-AI* mobility

The frequency of Ty1*his3-AI* mobility was determined as described previously (28, 46) with minor modifications. For transposition assays involving strains containing pGPOLΔ, a single colony from a SC-Ura plate incubated at 30^{o}C was resuspended in 1 ml of water and 5 µl of cells were added to quadruplicate 1 ml cultures of SC-Ura liquid

medium. The cultures were grown for 3 days at 22°C , washed, diluted and spread onto SC-Ura and SC-His-Ura plates to calculate Ty1 mobility. For mobility assays with strains repopulated with Ty1 elements, a single colony from a YEPD plate incubated at 30°C was diluted into 10 ml of water, and 1 μl of cell suspension was added to quadruplicate 1 ml YEPD cultures. The cultures were incubated for 2-3 days at 22°C, washed, diluted and then spread onto YEPD and SC-His plates. Plates were incubated for 4 days at 30°C. For mobility assays involving strains expressing pGTy1*his3-AI* and *GAL1*-p22 or related plasmids, a single colony was resuspended in 1 ml SC-Ura-Trp + 2% raffinose media and grown for 16 hr at 30°C, then diluted 25-fold into quadruplicate 1 ml cultures of SC-Ura-Trp + 2% galactose. Cultures were grown at 22°C for 2 days, washed, diluted, and spread onto SC-Ura-Trp and SC-Ura-Trp-His plates. Qualitative Ty1*his3-AI* mobility assays were performed as described previously (28, 46). For qualitative mobility assays involving strains containing pGPOLΔ, single colonies patched onto SC-Ura plates were incubated at 22°C for 2 days. To detect Ty1*HIS3* mobility events, cells were replica-plated onto SC-Ura-His plates and incubated at 30˚C for 3 days. For strains expressing pGTy1*his3-AI* and pGAL-p22 or related plasmids, single colonies patched onto SC-Ura-Trp plates were incubated for 2 days at 30°C. The resulting patches were replica-plated to SC-Ura-Trp + 2% galactose plates followed by incubation at 22°C for 2-4 days. To detect Ty1*HIS3* mobility events, galactose-induced cells were replica-plated to SC-Ura-Trp-His plates followed by incubation for 3 days at 30°C.

<u>p18 antiserum</u>

Ty1 (1068-1496) was amplified with primers (1068NdeI, 5'-CATGTTCCATATGCAATCTGATACCCAAGAGGCAA-3'; 1496XhoI, 5'-

CATGTTTCCTCGAGTTAGTGAGCCCTGGCTGTTTCG-3') and cloned into pET15bTEV vector (Novagen EMD, San Diego, CA). An 800 ml culture of *E. coli* BL21 (DE3) cells containing the expression plasmid in LB + 100 μg/ml ampicillin (Sigma-Aldrich, St. Louis, MO) were induced by 0.15 mM IPTG (Sigma-Aldrich) at 37°C. When cells reached an OD_{600} of 0.6-0.8, the temperature was reduced to 16°C, and incubated for an additional 24 hr. The cells were resuspended in 50 ml lysis buffer A (50 mM Phosphate Buffer pH 7.8, 1 M NaCl) and harvested by sonication. The His-tagged Ty1 product was purified with Talon affinity resin (Clontech) and eluted with 300 mM imidazole (Sigma-Aldrich). The elution product was dialyzed against storage buffer (10% glycerol, 1 M NaCl, and 25 mM Tris-HCl (pH 8.0)) overnight. A rabbit polyclonal antibody was raised against the truncated Ty1 Gag protein by Bio-synthesis Inc. (Lewisville, TX).

Protein isolation and immunoblotting

To detect protein expression from pGPOLΔ in the absence of galactose induction, 5 ml of SC-Ura medium was inoculated with a single colony and grown at 22°C for 24 hr. For coexpression of independent pGAL expression plasmids, 1 ml of SC-Ura-Trp + 2% raffinose was inoculated with a single colony and grown at overnight at 30°C. The overnight culture was diluted 25-fold into SC-Ura-Trp +2% galactose and grown for 2 days at 22°C. 5 ml of culture was processed by TCA extraction as described previously (47) except cells were broken by vortexing in the presence of glass beads, and 10 μl of the supernatant was separated by electrophoresis. For sucrose fractions, equal volumes of each fraction were analyzed. For P40 and VLP samples, 5 μg of P40 was used to detect p22/p18 and 10 μg of P40 was used for RT and IN. Samples were

separated on 10% (for RT and IN detection) or 15% (Gag p49/p45 and p22/p18 detection) SDS-PAGE gels. For optimal detection of p22/p18, proteins were transferred to PVDF membrane at 100 V for 90 min. The membranes were blocked in 5% milk/TBST (500 mM NaCl, 20 mM Tris-HCl, 0.1% Tween-20, pH 7.6) and then incubated with rabbit polyclonal antisera at the following dilutions: αp18; 1:5,000 in 2.5% milk/TBST, αRT/B8; 1:5,000 in TBST, and αIN/B2 1:2,500 in TBST (48). Immune complexes were detected with ECL reagent (GE Healthcare, Little Chalfont, United Kingdom).

VLP isolation

VLP purification from DG3739, DG3774 and DG3784 (Table 2-1) and reverse transcriptase assays were performed as described previously (49, 50) with the following modifications. Briefly, 40 ml SC-Ura-Trp + 2% raffinose cultures of strains used for VLP analysis were grown overnight at 30°C with shaking. Each culture was diluted 25-fold into 1 liter of SC-Ura-Trp + 3% galactose and grown at 21°C to OD_{600} 1-1.2. Cells were harvested by centrifuging at 6,000 rpm and homogenized with acid-washed glass beads in buffer B (15 mM KCl, 10 mM HEPES-KOH pH 7, 5 mM EDTA) containing protease inhibitor cocktail (0.125 mg/ml aprotinin, leupeptin, pepstatin A and 1.6 mg/ml PMSF). The crude lysate was centrifuged at 10,000 rpm and the supernatant loaded onto a step gradient of 20%, 30%, 45% and 75% sucrose in buffer B. The step gradient was centrifuged at 25,000 rpm in a SW28 rotor for 3 hr. Four ml of the gradient at the junction of the 30% and the 45% sucrose layers was withdrawn, diluted to 10% sucrose with buffer B and pelleted by centrifugation at 55,000 rpm in a Ti 70.1 rotor for 45 min. The resulting crude VLP pellet (P40) was suspended in buffer B and centrifuged

through a 20 - 60% continuous sucrose gradient in buffer B at 25,000 rpm in a SW41 rotor for 3 hr. The entire gradient was dripped into 19 equal fractions using an ISCO Foxy Jr. fraction collector (Lincoln, NE). All steps were carried out 4°C unless specified. Fractions were assayed for Ty1 reverse transcriptase activity as described previously (34, 49), except 10 μl samples were incubated with exogenous reverse transcriptase mix (50 mM Tris-HCl pH 8, 10 mM $MgCl_2$, 20 mM DTT, 15 μM dGTP, 10.7 μg poly(rC)/(dG)) and [α-^{32}P] dGTP (3,000 Ci/mmol; Perkin Elmer).

Electron microscopy

Three sucrose gradient fractions with the highest reverse transcriptase activity from DG3739 (fractions 5-7) and DG3774 (fractions 4-6) were pooled, diluted with buffer B and pelleted as described above. The sample was allowed to bind for 15 min to Formvar and carbon-coated 400-mesh copper grids. Grids were stained with 2% ammonium molybdate, pH 6.5 for 10 sec and visualized with a JEM-1210 Transmission Electron Microscope (JEOL USA Inc., Peabody, MA) equipped with an XR41C Bottom-Mount CCD Camera (Advanced Microscopy Techniques, Woburn, MA). Approximately 100 VLPs were analyzed to determine the percentage of closed versus open particles. VLP diameter was measured with closed VLPs only using ImageJ (51) and the two datasets were compared using an unpaired T test.

FISH/IF

Two ml SC-Ura-Trp + 3% raffinose cultures were inoculated with a single colony and grown 16 hr at 30°C. The overnight cultures were diluted 10-fold into SC-Ura-Trp + 3% galactose and grown at 22°C for 24-30 hr until an OD_{600} of 0.8-1.0 was reached. Formaldehyde was added directly to the culture at a final concentration of 4% and

allowed to fix for 1.75 hr. Processing of the cells for FISH/IF was performed as described previously (7). For Gag/p22-V5 colocalization experiments, primary antibodies were αVLP (rabbit polyclonal, 1:2,000, a kind gift from Alan and Susan Kingsman) and αV5 (Life Technologies, 1:4,000) and secondary antibodies used were α-rabbit-AF488 (Life Technologies, 1:200) and α-mouse-AF594 (Life Technologies, 1:400). Image acquisition was carried out using a Zeiss Axio Observer microscope equipped with an AxioCam HSm camera and analyzed with AxioVision v4.6 software (Carl Zeiss Microscopy, LLC, North America). Exposure times used to capture fluorescent and DAPI images were kept consistent throughout each experiment. Figures were constructed with Adobe Photoshop software (Adobe Systems, San Jose, CA).

GST-pulldown

One ml of SC-Ura + 2% raffinose was inoculated with a single colony at 30°C overnight and was then diluted 1:25 into 5 ml SC-Ura +2% galactose and grown for 2 days at 22°C. 2.5 ml of galactose-induced cells was suspended in 150 μl lysis buffer C (20 mM Tris-HCl pH 7.5, 100 mM NaCl, 10 mM $MgCl_2$, 1 mM EDTA, 1 mM DTT, 0.1% Triton X-100, 1 mM PMSF, 1 μg/ml aprotinin, 0.5 μg/ml leupeptin and 1 μg/ml pepstatin A) and homogenized with the same volume of acid washed glass beads. The crude lysate was centrifuged at 10,000 rpm at 4°C for 10 min. and 500 μl supernatant containing 300 μg of protein was gently mixed with 20 μl glutathione-coated resin (GenScript, Piscataway, NJ) at 4°C for 2 hr. The resin was washed three times with 1 ml lysis buffer C and then suspended in 40 μl SDS loading buffer. After boiling for 10 min, 5-8 μl per lane were loaded onto a 12% SDS-polyacrylamide gel. Immunoblotting was performed as described above and membranes were incubated with mouse monoclonal antibodies

αGST/B-14 (Santa Cruz Biotech) at 1:1,000 or αTY tag (a kind gift from Stephen Hajduk) at 1:50,000 in TBST.

Results

An internal Ty1 sense-strand RNA is required for CNC

The CNC region of Ty1 spans the 5' UTR and all of *GAG*, and a multicopy pGTy1 expression plasmid confers CNC *in trans* even when *GAL1*-promoted transcription is repressed (28, 31). To identify sequences necessary for CNC (Figure 2-1 (A)), a genetic screen for CNC^- mutations was performed in a Ty1-less *S. paradoxus* strain repopulated with a single chromosomal Ty1 insertion containing the selectable indicator gene *his3-AI* (46) (Table 2-1). Ty1*HIS3* insertions usually occur by retrotransposition following splicing of the artificial intron. Since Ty1*HIS3* cDNA can also undergo homologous recombination with genomic Ty1 elements or solo LTRs (52, 53), the term Ty1 "mobility" is used to describe both types of insertion. Ty1 mobility was followed using a qualitative papillation assay for His^+ cells in a Ty1-less test strain containing a chromosomal Ty1*his3-AI* element and an empty vector, wild-type pGTy1 plasmid or randomly mutagenized pGTy1. Cells were grown under repressive conditions for *GAL1* expression. To generate point mutations, the CNC region was amplified using Taq DNA polymerase and PCR products were recombined into pGTy1 *in vivo* by gap repair. Approximately 3,500 pGTy1 recombinants were screened for loss of Ty1 CNC and recovered plasmids were reintroduced to confirm the CNC^- phenotype. Although pGTy1 plasmids with one to four base-changes in the CNC region were identified, only

plasmids carrying single mutations (Figure 2-1(A) and Table 2-2(A)) were analyzed further.

To minimize the possibility that sequence changes outside of the gap-repaired region influence CNC and to facilitate molecular analyses, most of *POL* was deleted from the mutant pGTy1 plasmids to generate plasmid pGPOLΔ. Quantitative Ty1*his3-AI* mobility assays were performed with five mutants from the screen (Table 2-2 A). Mutations T399C, ΔA1456 and A1296G conferred moderate decreases in CNC when compared with the CNC^+ control, while T1108C and A1123G conferred low levels of CNC. Furthermore, the T1108C mutation affected CNC the most and was obtained from four independent isolates, suggesting T1108 is part of an important sequence motif involved in CNC. Since Ty1AS RNAs were reported to be necessary for CNC (31), Northern blotting was performed with total RNA from the five single mutants. All of the mutants except T399C contained a similar level of Ty1AS RNAs as that produced from a wild-type pGPOLΔ plasmid, when compared with the *ACT1* loading control (Figure 2-1 (B)). Surprisingly, four of the five CNC^- mutations do not map in the Ty1AS RNA transcription units and instead are located in an adjacent segment of the CNC region (Figure 2-1 (A)), and all change *GAG*'s coding potential (Table 2-2 (A)).

A 5' truncated Ty1 sense RNA can be detected in wild type cells and is enriched in an *spt3* mutant (22, 51). A similar observation was reported for an *xrn1* mutant, where the RNA was termed Ty1SL (Ty1 short length RNA) (24). Therefore, the point mutations identified in the screen could map in a shorter Ty1 sense RNA that initiates in *GAG* independently of normal Ty1 transcription and this transcript could be involved in CNC. To determine if a shorter Ty1 sense RNA was produced from the pGPOLΔ plasmids,

total RNA was subjected to Northern blotting using a strand-specific ^{32}P-labeled riboprobe from *GAG*. Cells containing pGPOLΔ and mutant derivatives were used in the Northern blotting since deleting *POL* results in the synthesis of Ty1 transcripts that are clearly distinguishable from Ty1*his3-AI* RNA. All point mutants except T399C made a shorter sense-strand Ty1 RNA, termed Ty1 internal (Ty1i) RNA, whereas cells containing an empty vector control lacked this transcript (Figure 2-1(C)). Two additional mutants, Δ238-281 and Δ238-353 were derived in the pGPOLΔ context and included in this analysis. Originally described in Matsuda and Garfinkel (2009), pGTy1 plasmids with short deletions in the 5' LTR abolished CNC. The loss of CNC was attributed to a decrease in the level of the Ty1AS RNAs due to deletion of their 3' ends. However, the lack of detectable Ty1i RNA in the Δ238-281 and Δ238-353 mutants may now explain their CNC$^-$ phenotype. These results also suggest that sequences near the 5' LTR, which contains the enhancer required for Ty1 transcription (5), may also be important for synthesizing Ty1i RNA.

Chromosomal Ty1A1123G insertions fail to confer CNC

The genetic screen identified several missense mutations in *GAG* that weakened CNC and were present on both Ty1 mRNA and the Ty1i transcript. To determine if CNC$^-$ mutations impacted Gag function, full-length pGTy1 plasmids containing A1123G (Tyr277Cys) or A1296G (Thr335Ala) mutations were compared with wild-type pGTy1 for their ability to stimulate or trans-activate movement of a chromosomal Ty1*his3-AI* element (33). The A1296G mutation likely affects both CNC and transposition, since pGTy1A1296G expression did not stimulate Ty1 mobility. However, induction of pGTy1A1123G increased Ty1*his3-AI* mobility *in trans* to similar levels observed with

wild-type pGTy1, suggesting that pGTy1A1123G encodes functional Gag yet is defective for CNC (data not shown). To determine if Ty1A1123G conferred CNC in a natural chromosomal context (Figure 2-2), a strain with a single chromosomal Ty1*his3-AI* (Figure 2-2 (A)) was repopulated with wild type (Figure 2-2 (B)) or Ty1A1123G elements (Figure 2-2 (C)). As expected, Ty1 mobility decreased 33-fold in a strain repopulated with 12 wild-type Ty1 elements when compared with the starting strain (Figure 2-2 (B), Table 2-2(B)). However, Ty1 mobility increased almost 5-fold in a strain containing 7 copies of Ty1A1123G, indicating that A1123G abolishes CNC without disrupting the function of Gag (Figure 2-2(C), Table 2-2(B)). The separation of function phenotype displayed by Ty1A1123G raised the possibility that an altered form of Gag encoded by Ty1i RNA mediates CNC.

Expression of Ty1i RNA

Since multicopy Ty1 plasmids were used as the source of *trans*-acting factors required for CNC, it was important to determine if chromosomal elements also synthesized Ty1i RNA and truncated forms of Gag (Figure 2-3). To detect Ty1i RNA in repopulated *S. paradoxus* as well as *S. cerevisiae* strains, poly(A)$^+$ RNA was subjected to Northern blotting using a ^{32}P-labeled riboprobe from *GAG-POL* (nt 1266-1601) (Figure 2-3(A)). Three *S. paradoxus* strains were analyzed: the Ty1-less strain (lane C), a derivative repopulated with 38 Ty1 elements (lane 1) and an isogenic *spt3Δ* mutant (lane 2). Five *S. cerevisiae* strains were also analyzed: GRF167 (lane 3) and an isogenic *spt3Δ* mutant (lane 4), BY4742 (lane 5) and isogenic *spt3Δ* (lane 6) and *xrn1Δ* (lane 7) mutant derivatives. A discrete subgenomic Ty1 RNA of 4.9 kb was detected below the full-length transcript (5.7 kb) in all strains except repopulated *S. paradoxus*

(Figure 2-3 (A), lane 1). The failure to detect a distinct transcript in this strain was unexpected, but may result from 5' heterogeneity of the 4.9 kb transcript. The 4.9 kb Ty1 RNA comigrated with the truncated transcripts detected in *spt3Δ* and *xrn1Δ* mutants. To determine the 5' end of the 4.9 kb transcript in BY4742 and an isogenic *spt3Δ* mutant, poly(A)$^+$ RNA was subjected to cap-independent 5'-RACE (Figure 2-4). In both strains, the majority of the 5' ends from the 4.9 kb transcript mapped to nucleotide 1000 of Ty1H3 (Figure 2-4 (A)). These results indicate that the 4.9 kb RNA observed in wild-type and an *spt3Δ* mutant share the same 5' ends. However, 5'-RACE analysis of the wild-type repopulated *S. paradoxus* strain showed heterogeneous amplification products (Figure 2-4 (B)) rather than discrete bands, supporting the results from Northern blotting (Figure 2-3). Although our results suggest that the 4.9 kb Ty1 RNA contains the Ty1i transcript in *S. cerevisiae*, other truncated forms of Ty1 RNA may be present (54, 55).

Ty1i RNA encodes Gag proteins, p22 and p18

Two closely spaced AUG codons are present 38 (AUG1) and 68 (AUG2) nucleotides downstream of the transcription start site for Ty1i RNA and one or both may be utilized to initiate synthesis of a truncated form of Gag (Figure 2-1 (A)). However, neither the predicted 22-kD (p22) Gag-like protein nor its processed product (p18), if p22 is cleaved by Ty1 PR, have been reported to date, and a commonly used VLP antiserum (56) failed to detect p22/p18 reproducibly (data not shown). Therefore, we purified recombinant p18 and generated a new antiserum to determine if Ty1i RNA is translated to produce an N-terminal truncated form of Gag (Figure 2-3 (B)). Whole cell extracts from the strains described above (Figure 2-3 (A)) were immunoblotted with p18

antiserum to detect endogenous Gag and additional Gag-related proteins. As expected, normal levels of Ty1 Gag p49/p45 were detected in wild-type strains (Figure 2-3 (B), lanes 1, 3, and 5) while reduced levels were observed in *spt3Δ* (lanes 2, 4, and 6) mutants. Importantly, p22 was detected in the *spt3Δ* mutants (lanes 2, 4 and 6), whereas p18 was only detected in the wild-type strains (lanes 1, 3 and 5) and the *xrn1Δ* mutant (lane 7). The increase in p18 observed in the *xrn1Δ* mutant likely results from an increase in Ty1i RNA level, since Xrn1 is the major 5'-3' exonuclease involved in RNA decay. Taken together, our results not only suggest that Ty1i RNA encodes p22, but the striking relationship between expression of full-length Ty1 mRNA, and hence Ty1 PR, and detection of p22 versus p18 suggests that p22 is cleaved by Ty1 PR to form p18. Furthermore, processing of p22 to p18 raises the possibility that p22 associates with VLPs to gain access to PR. As expected, Gag proteins were not detected in the Ty1-less strain (lane C).

Ribosome footprint profiling reveals an internal AUG as a potential translation start for p22

To determine if the candidate AUG1 or AUG2 translation start sites on Ty1i RNA (Figure 2-1 (A) were present in genomic sequencing analyses, we turned to ribosome footprint profiling (Ribo-seq) (Figure 2-5). In Ribo-seq, ribosomes in the act of translating an mRNA are treated with RNase I, leaving a ~ 28 nt ribosome footprint, which is harvested for high throughput sequencing to provide a snapshot of the abundance and distribution of ribosomes on mRNAs (43). Yeast starved for glucose for 3 hr accumulate as much as 10% of Ribo-seq reads at the start codon of ORFs (42), providing a sensitive method for detecting initiation codons *in vivo*. We utilized a

published dataset to analyze the Ribo-seq read distribution at the 5' end of Ty1i RNA (57). The most abundant read in this region corresponded to a ribosome footprint located on AUG1, which is the first start codon downstream of the Ty1i transcription start site (Figure 2-4 (A)). Also, the density of Ribo-seq reads increased downstream of AUG1, consistent with translation of the downstream ORF under glucose starvation. Additional mutational analysis of AUG1 and AUG2 will be required to verify the translation start of p22.

p22/p18 encoded by Ty1i RNA is necessary for CNC

To establish that p22/p18, rather than the Ty1i transcript itself, is responsible for CNC, we analyzed frameshift mutations in the pGPOLΔ construct (Figure 2-6 (A)) for alterations in Ty1*his3-AI* mobility (Figure 2-6 (B), Table 2-2 (C-1)), Ty1i RNA levels, (Figure 2-6 (C)) and protein levels (Figure 2-6 (D)). Cells containing pGPOLΔ decreased the mobility of a chromosomal Ty1*his3-AI* element up to 74-fold (Figure 2-6(B), Table 2-2 C-1) compared to an empty vector control, and produced Ty1i RNA (Figure 2-6 (C)) and p22 (Figure 2-6 (D)). Two frameshift mutations were placed downstream of AUG1 and AUG2 that introduce premature termination codons, ΔC1071 and +A1303. +A1303 was created to eliminate the possibility that downstream in-frame AUGs (AUG3 and AUG4, Figure 2-6 (A)) could be utilized to produce a *trans*-dominant factor. Both frameshift mutations caused an increase in Ty1*his3-AI* mobility to almost the same level as that obtained in a strain lacking CNC (Figure 2-6 (B), Table 2-2 C-1). Cells containing the mutant plasmids produced Ty1i RNA (Figure 2-6 (C)), but not wild-type p22 (Figure 2-6 (D)). The residual level of CNC conferred by the plasmids carrying the frameshift mutations may be caused by truncated protein synthesized prior to encountering the

mutations; however, immunoblotting using the p18 antiserum did not detect these smaller proteins (Figure 2-6 (A) and (D)). To fully eliminate protein production from AUG1 and AUG2, we replaced both initiation codons with the alanine codon GCG in pGPOLΔ. In cells carrying pGPOLΔ-GCG1GCG2, transposition frequency was fully restored and about 2-fold higher than the ΔC1071 frameshift (Table 2-2 C-2). These results show that AUG1 and/or AUG2 are necessary for CNC, and reinforces the observation that ΔC1071 confers a very low level of CNC. Taken together, our results identify p22 as a *trans*-dominant negative inhibitor of Ty1 retrotransposition and the intrinsic factor responsible for CNC.

Ectopic expression of p22/p18 is sufficient to inhibit Ty1 movement

To determine if p22/p18 reduces Ty1 transposition, a cDNA expression library (45) was screened for clones that inhibited Ty1*his3-AI* mobility, and p22/p18 was ectopically expressed from the *GAL1* promoter. One clone was obtained from the *GAL1*-driven cDNA library that contained Ty1 sequences 1042-5889 and inhibited chromosomal Ty1*his3-AI* mobility. The 5' end of the cDNA included AUG2 and 26 additional nucleotides upstream but did not contain AUG1. The 3' end terminated in the R region (3' LTR) of Ty1 RNA, which is the similar to the 3' ends mapped previously (15, 55). Therefore, an almost full length 4.9 kb Ty1i transcript from a chromosomal element was likely captured as this cDNA clone, and contains coding sequence for p22, as well as the *POL* coding sequence for PR, IN, and RT. When the cDNA clone and pGTy1*his3-AI* were coexpressed, Ty1 mobility decreased 570-fold when compared with a control strain expressing only pGTy1*his3-AI* (Table 2-2 D). These results support the idea that a truncated Gag protein likely utilizing AUG2 inhibits Ty1 mobility, although

initiation from AUG2 occurs less frequently than AUG1 based on ribosome profiling in the Sigma 1278b strain (Figure 2-5).

The following segments of *GAG* sequence starting with AUG1 were fused to the *GAL1* promoter on a multicopy expression plasmid and analyzed for *trans*-dominance (Table 2-2 E) or protein expression (Figures 2-7 and data not shown) in the Ty1-less *S. paradoxus* strain: p22, p22 containing an internal V5 epitope, p18, and $p22^{Gag*PR}$ containing a previously characterized mutation that disrupts Gag-PR cleavage by PR (40). Ty1*his3-AI* mobility decreased more than 32,000-fold in cells coexpressing pGTy1*his3-AI* and p22 or p22-V5 when compared with the control strain expressing pGTy1*his3-AI* (Table 2-2 E). Both p22 and p18 are present, again suggesting that some p22 is incorporated into VLPs and cleaved by Ty1 PR (Figure 2-7). Coexpression of p18 and pGTy1*his3-AI* reduced Ty1 mobility to levels similar to those observed with p22 (Figure 2-7). *GAL1*-promoted expression of p22 or p18 also inhibited pGTy1*his3-AI* mobility in *S. cerevisiae* strains BY4742 and GRF167 (data not shown).

To determine if p22 alone inhibited Ty1 mobility, we coexpressed $p22^{Gag*PR}$ and pGTy1*his3-AI* (Figure 2-7) in the Ty1-less *S. paradoxus* strain. Results from qualitative mobility assays indicated that $p22^{Gag*PR}$ retained most if not all of its inhibitory function when compared with wild-type p22 or p18 and the empty vector control, even though processing of p22 to p18 was blocked (Figure 2-7). The level of $p22^{Gag*PR}$ when expressed from the *GAL1* promoter was also comparable to that obtained with p22, p18, or full-length Gag. Together, these results show that p22 and p18 are potent *trans*-dominant inhibitors of Ty1 transposition.

p22/p18 cofractionates with VLPs and alters Ty1 proteins

One possibility to account for the dramatic decrease in Ty1 mobility is that p22 associates with assembling VLPs in the cell, leading to abnormal VLP function. Therefore, crude VLP preparations from Ty1-less strains expressing pGTy1*his3-AI* alone (Figure 2-8 (A)), p22 and pGTy1*his3-AI* (Figure 2-8 (B)) or p22 alone (Figure 2-8 (C)) were separated by centrifugation through 20-60% continuous sucrose gradients. Fractions were assayed for reverse transcriptase activity using an exogenous primer/template, and immunoblotted for Gag, IN, RT and p22/p18 (Figure 2-8 (A) and (B)) or p22/p18 alone (Figure 2-8 (C)). As expected, a peak of reverse transcriptase activity coincided with the highest concentrations of mature Gag, RT, and IN proteins in the strain expressing just pGTy1*his3-AI* (Figure 2-8 (A)). When pGTy1*his3-AI* and p22 were co-expressed, Gag and p22/p18 displayed a similar fractionation pattern across the gradient (Figure 2-8 (B). p18 appeared to be the predominant form present in crude VLP preparations, which is likely due to processing by Ty1 PR in VLPs. To further investigate if the cofractionation of Gag and p22/p18 resulted from an association between VLPs and p22, rather than comigration of a protein complex containing p22 that had a similar density as VLPs, an identical fractionation was performed in a strain expressing only *GAL1*-promoted p22. When expressed alone, p22 was detected near the top of the gradient (Figure 2-8 (C)), and therefore, had a different fractionation profile than that observed when pGTy1*his3-AI* and p22 were coexpressed (Figure 2-8 (B)). Furthermore, we detected p18 in the CNC^{+} VLPs (data not shown) used for structural probing of packaged Ty1 RNA (32). These results support an interaction between Ty1 VLPs and p22/p18.

When comparing the strain expressing pGTy1*his3-AI* (Figure 2-8 (A)) with one expressing pGTy1*his3-AI* and p22 (Figure 2-8 (B)), several differences in the fractionation patterns, protein composition and distribution, and reverse transcriptase activity were evident. First, cells expressing only pGTy1*his3-AI* yielded a higher concentration of Gag, IN, RT and reverse transcriptase-catalyzed incorporation of [α^{32}P]-dGTP in the peak fractions. Second, cells coexpressing pGTy1*his3-AI* and p22 showed a broader distribution of Ty1 proteins and reverse transcriptase activity. Third, the VLPs formed in the presence of p22 had a lower level of [α^{32}P]-dGTP incorporation throughout the gradient. Fourth, Ty1 protein processing or stability was altered when pGTy1*his3-AI* and p22 were coexpressed. There was an accumulation of the PR-IN precursor (p91) and much less mature IN (p71), which is similar to results obtained previously (31). Ty1 RT (p63) now appeared as a doublet with an additional higher molecular weight protein that reacted with the RT antibody (Figure 2-8 (B), denoted by an asterisk). Fifth, Ty1 Gag appeared to undergo more proteolysis overall when p22 was present, as evidenced by multiple lower molecular weight Gag-related proteins, which cofractionated with full-length Gag. These unusual Ty1 proteins may result from aberrant processing by Ty1 PR, cleavage by a cellular protease, or from differences in posttranslational modification of Ty1 proteins brought about by a VLP-p22 interaction. Therefore, the mechanism of CNC involves differences in the physical and biochemical properties of VLPs assembled in the presence of p22.

p22/p18 changes VLP morphology

Since p22/p18 affected the fractionation of Ty1 VLPs and appearance of Ty1 proteins (Figure 2-8), we examined the size and morphology of VLPs assembled in the presence or absence of p22 by electron microscopy (Figure 2-9). Equivalent sucrose gradient fractions with the highest level of [α^{32}P]-dGTP incorporation (Figure 2-8) were pooled, diluted, and concentrated by ultracentrifugation prior to staining with 2% ammonium molybdate. Ty1 VLPs formed in the absence of p22 (Figure 2-9 (A)) were mostly intact with an average diameter of 37.4 ± 2.7 nm, and only 13% of wild-type VLPs appeared malformed. In contrast, almost half of Ty1 VLPs formed in the presence of p22 (Figure 2-9 (B)) appeared open or incomplete, suggesting that these VLPs are either not formed properly or are less stable during sample preparation. The diameter of intact VLPs assembled in the presence of p22 was 39.2 ± 3.1 nm. Although the difference in diameters of the two batches of VLPs is statistically significant (P = 0.0005), further analyses will be required to determine if this difference is functionally relevant.

p22-V5 disrupts pGTy1-induced retrosomes and colocalizes with Gag

Since p22 altered the fractionation pattern and morphology of VLPs, and the processing or stability of Ty1 proteins, we examined whether p22 influenced the appearance of retrosomes, which are sites for VLP assembly. Ty1-less strains expressing p22-V5 and pGTy1*his3-AI* alone or together were subjected to indirect immunofluorescence (IF) and fluorescence *in situ* hybridization (FISH) to visualize retrosomes (Figure 2-10). VLP or V5 antibodies were used to detect Ty1 proteins and a *GAG*-DIG probe was used to detect full length Ty1 mRNA. The internal V5 tag in p22-

V5 did not disrupt *trans*-dominance (Table 2 E) and retrosome analysis of cells expressing untagged p22 was identical to that from strains expressing p22-V5. Three types of staining were observed: (1) large, distinct foci that co-stain for Ty1 mRNA and Gag were defined as retrosomes (R), (2) nondistinct, punctate staining for both Ty1 mRNA and Gag was termed “puncta” (P), and (3) lack of staining for Ty1 mRNA, Gag or both was designated as “none”. In cells containing puncta, colocalization between Ty1 mRNA and Gag could not be confidently determined in the majority of cells. In a control strain expressing pGTy1*his3-AI* alone, retrosomes were observed in 61% of cells, while only 7% of cells showed a punctate localization of Ty1 mRNA and Gag proteins (Figure 2-10 (A)). When p22-V5 and pGTy1*his3-AI* were coexpressed, the percentage of cells containing normal retrosomes decreased to 18% while Ty1 puncta was observed in 31% of cells. Thus, p22-V5 disrupts Ty1 retrosomes in a large fraction of cells. In addition, cells were analyzed for Ty1 Gag and p22-V5 colocalization using VLP and V5 antibodies, respectively (Figure 2-10 (C) and (D)). As expected, a similar percentage of cells exhibited retrosomes (61%) (Figure 2-10 (A)) and Gag foci in the absence of p22-V5 (62%) (Figure 2-10 (C)). In the presence of p22-V5, a comparable fraction of cells displaying Gag foci (28%) and puncta (42%) was observed (Figure 2-10 (D)) when compared to the staining observed using FISH/IF analysis (Figure 2-10 (B)). Interestingly, we detected colocalization of p22-V5 and Gag in almost 70% of Gag foci (Figure 2-10 (D) inset). p22-V5 colocalized with endogenous retrosomes in *S. cerevisiae*, suggesting the possibility that p22 can associate with VLP pre-assembly intermediates (Mitchell and Garfinkel, unpublished results).

GST-pull downs support an interaction between Gag and p18

To provide additional evidence for an interaction between p22/p18 and Gag, a fusion protein consisting of p18 tagged at its N-terminus with Glutathione-S-transferase (41) was expressed from the *GAL1* promoter in BY4742 or a Ty1-less strain (Figure 2-11). Free glutathione-S-transferase was expressed alone as a negative control. Protein extracts were immunoblotted using antisera specific for GST, Gag p49/p45, or Hts1 prior to mixing with the glutathione-coated resin (Input) or released from the GST complexes bound to resin after several washes with lysis buffer (Pull-down). Fusions between GST and full-length Gag were insoluble under a variety of conditions, and therefore, could not be analyzed further. The GST-p18 fusion protein was soluble under the conditions used for the pull-down; however, partial degradation of GST-p18 resulted in free GST protein. GST-p18 formed a complex containing Gag p45 and p49 encoded by the genomic Ty1 elements in BY4742, whereas GST expressed alone did not. Ty1 Gag-p18 complexes were also not detected in the Ty1-less strain. Hts1 was used to control for nonspecific trapping of cellular proteins in the Ty1 complexes, and as expected, was only detected in the input samples. Together, these results suggest that p18 and Gag interact.

Discussion

Here, we characterize a restriction factor derived from Ty1 *GAG* that confers CNC by perturbing VLP assembly and function. This unique form of transposon CNC (28) may have evolved after an ancestral *S. cerevisiae/paradoxus* lineage lost the evolutionarily conserved RNAi pathway used to silence Ty1 expression (29, 30).

Noncoding antisense transcripts from Ty1 have been implicated in repressing transcription (24), RNAi in budding yeast (30), and CNC (31). The identification of mutations that abrogated both CNC and Ty1AS RNA expression implicated Ty1AS RNAs in CNC. Additionally, the association of Ty1AS RNAs with VLPs further supported models of AS RNA-based CNC. Here, we show additional mutations in the CNC region of Ty1 fail to confer CNC, yet do not perturb Ty1AS RNA expression. One *GAG* mutation in particular abolished CNC but did not affect transposition or AS RNA production. The behavior of this separation of function mutation suggested that a Ty1 protein might contribute to CNC. Evaluation of these mutants, along with those previously reported helped reveal p22, a Gag-like restriction factor encoded by a 5' truncated sense RNA (Ty1i) that likely forms the basis of CNC. The role of Ty1AS RNAs in Ty1 CNC, if any, remains to be determined.

We detect differences in the transcripts encoding the p22 restriction factor and how these transcripts are utilized for protein synthesis. In *S. cerevisiae*, a 4.9 kb Ty1i RNA is detected in wild-type strains both in our work and in previous studies when poly(A)$^+$ RNA is subjected to Northern blotting (22, 54), but is rarely detected in numerous studies when total RNA is analyzed (7, 8, 18, 58-60). Perhaps the level of RNA degradation observed with the abundant 5.7 kb Ty1 genomic RNA obscures the 4.9 kb Ty1i transcript when total RNA is analyzed by Northern blotting, because we can detect a shorter Ty1i transcript produced from a pGPOLΔ plasmid with total RNA from *S. paradoxus*. Alternatively, it has been reported that only 15% of Ty1 mRNA transcripts are polyadenylated (9). Hence, it is possible that Ty1i RNA is readily detected by Northern blotting of poly(A$^+$) RNA because the majority of Ty1i transcripts are

polyadenylated, whereas the majority of Ty1 mRNA is not. However, in an isogenic repopulated *S. paradoxus* strain, a discrete Ty1i transcript is not detected from chromosomal Ty1 elements even when poly(A)$^+$ RNA is analyzed by Northern blotting or cap-independent 5'-RACE. Ty1i RNA is present in both species when full-length Ty1 transcription is altered by deleting the Spt3 subunit of SAGA and related complexes (22, 23). Spt3 helps modulate the recruitment of the TATA-binding protein to the TATA box of SAGA-dependent promoters (61-63), and therefore, can specify transcriptional initiation. However, the initiation site for Ty1i RNA within *GAG* predominates in an *spt3Δ* mutant, which is similar to the activation of cryptic intragenic promoters observed in a variety of chromatin and transcription-related mutants (64). Although our results are consistent with the idea that transcription of Ty1i RNA responds differently to the complexes containing Spt3, such as SAGA, in *S. cerevisiae* versus *S. paradoxus*, detailed functional comparisons between Spt3/SAGA from these species will be required to resolve this question.

Surprisingly, appreciable levels of p22/p18 are present in wild-type *S. paradoxus* repopulated with Ty1H3 in the absence of detectable 4.9 kb Ty1i RNA. This result raises the possibility that full-length Ty1 and Ty1i transcripts may utilize an internal ribosome entry site (65) upstream of AUG1 or AUG2 to drive synthesis of p22. Other mechanisms by which p22 could be translated from full-length Ty1 mRNA are leaky scanning, where scanning ribosomes sometimes initiate translation from an alternate AUG codon (66-69) or translation reinitiation in which translation starts at a downstream AUG after translation of an ORF situated upstream (70, 71). Although leaky scanning and translation reinitiation remain possible mechanisms, both require closely spaced

AUGs. However, seven in-frame and seven out-of-frame AUGs are present in the 745 bases between the Gag initiation codon (nt 293) and p22 AUG1 (nt 1038), making leaky scanning or translation reinitiation unlikely. Alternatively, exceptional forms of translation initiation may not be required to synthesize p22 if heterogeneous Ty1i transcripts that contain AUG1 or AUG2 in the repopulated *S. paradoxus* strain remain translatable. In support of this view, we show that Ty1i RNA is a functional template for translation of p22 in *S. paradoxus* and *S. cerevisiae spt3Δ* mutants in the absence of full length Ty1 mRNA. Although it is possible that there are two modes of p22 production in yeast (Ty1 mRNA and Ty1i RNA mediated), production of p22 from internal Ty1 RNA products alone is an attractive idea.

Once synthesized, p22 profoundly inhibits retrotransposition by altering VLP assembly and function. Earlier work as well as our mutational analysis of the CNC region demonstrates that Ty1 produces a *trans*-dominant inhibitor, now identified as p22, that decreases Ty1*his3-AI* mobility 20 to >340-fold depending on the relative expression of Ty1 and p22 (28, 31, 32, 72). However, when a cDNA derived from Ty1i RNA or p22 and Ty1*his3-AI* are coexpressed from the *GAL1* promoter in a Ty1-less strain, mobility decreases 570- and 32,000-fold, respectively, indicating that p22 is necessary and sufficient for inhibition. The extreme inhibitory effect and broad dynamic range raises the possibility that the process of retrotransposition is very sensitive to the level of p22, with increasingly severe defects appearing as the level of p22 increases. Conversely, the relative amount of Ty1 verses p22 expression can likely saturate the inhibitor, as is evident from previous studies utilizing *GAL1*-promoted Ty1 induction (15, 16, 33, 34). In fact, Ty1 “transpositional dormancy,” which was described upon the

discovery of Ty1 retrotransposition (15, 34) may result from an inhibitor that is saturated or overcome when Ty1 is induced via the *GAL1* promoter (73-75). The work presented here supports this hypothesis and identifies p22 as the intrinsic inhibitor at least partly responsible for Ty1 dormancy.

When crude VLPs from the Ty1-less strain expressing Ty1 and p22 are analyzed by sucrose gradient sedimentation, both p22 and its processed product p18 cofractionate with Ty1 VLPs. p22 does not exhibit the same fractionation pattern in the absence of pGTy1 expression. Furthermore, analysis of a $p22^{Gag*PR}$ cleavage site mutant shows that p22 as well as p18 effectively inhibits transposition, and cleavage of p22 does not play a major role in CNC. The sucrose gradient fractions have also been assayed for reverse transcriptase activity and subjected to additional immunoblotting to detect Gag, IN, RT and p22/p18. Expression of p22 causes a moderate decrease in the level of reverse transcriptase activity when assayed using an exogenous primer/template, prevents the accumulation of mature IN, which reinforces previous work (31, 32), and broadens the peak containing VLP proteins. In addition, an overall degradation of Gag and the presence of aberrant RT proteins are indicative of proteolysis of the Gag-Pol precursor by Ty1 PR, increased proteolysis by cellular enzymes, or possible post-translational modifications. Furthermore, the excessive proteolysis of IN could explain the appearance of higher molecular weight, RT antibody-reactive proteins and the absence of mature IN. Our results suggest that p22 interacts with and inhibits VLP functionality during assembly or in association with fully formed VLPs and also is processed by Ty1 PR to form p18.

Since these results suggest that VLP structure may be altered by p22, peak sucrose gradient fractions have been concentrated and visualized by electron microscopy. Most of the VLPs (87%) isolated from the control strain lacking p22/p18 are completely spherical with similar curvatures, however, almost half (46%) of the VLPs formed in the presence of p22/p18 are aberrant and have an open or incomplete morphology. VLPs analyzed from CNC^{+} cells containing much less p22/p18 do not appear malformed, but when extracts containing these VLPs are treated with the endonuclease benzonase, less protection of packaged Ty1 mRNA is observed (32). Our results suggest that VLP integrity is compromised in the presence of higher levels of p22/p18, and that normal assembly of functional VLPs is inhibited by an interaction between Gag and p22.

To further investigate if Gag and p22 interact, cells expressing Ty1 and p22/p18 have been subjected to FISH/IF microscopy and GST-pull down analysis. The number of cells with aberrant retrosomes increases more than 3-fold when Ty1 and p22-V5 are coexpressed, and 70% of Gag foci also stain for p22-V5. In addition, GST pull down analysis suggests that endogenous Gag can interact with GST-p18. Although p22 engages Gag during active VLP assembly, p22 may also interact with Gag in endogenous retrosomes, which contain few if any VLPs (7). Ty1 *GAG* is necessary for retrosome formation (7, 76), and certain Ty3 *GAG* mutations alter retrosome appearance or location (77, 78). Interestingly, cellular mutations that alter retrograde movement of Gag from the endoplasmic reticulum (ER) destabilize Gag and abolish nucleation of retrosomes (79). Whether p22 enters the ER remains to be determined. Ty1 *GAG* mutations have been isolated that confer a *trans*-dominant negative

phenotype (80-82) or affect VLP assembly (83), and some of these mutations map in p22. A synthetic peptide containing sequences within p22 also displays RNA chaperone activity (84), which is required for specific RNA transactions during the retroviral life-cycle such as virion assembly, RNA packaging, primer annealing, and reverse transcription (85). Thus, p22 may inhibit multiple functions carried out by Gag.

Certain retroelement genes have undergone purifying selection in mammals, suggesting that these elements have been domesticated or exapted by their host (86). To date, domesticated *GAG* and *POL* genes have either evolved a new function used in normal cellular processes or have been incorporated into an innate defense pathway used to inhibit retroviral propagation. The prototypic Gag-like restriction factors Fv1 and enJS56A1 block replication of murine leukemia virus (MLV) and Jaagsiekte sheep retrovirus (JSRV), respectively, by interacting with viral proteins during infection (87-89), and share features in common with CNC of Ty1 by p22/p18. Fv1 is derived from the *GAG* gene of a member of the HERV-L family of human and murine endogenous retroviruses (87, 90, 91). Fv1 inhibits progression of the MLV life cycle following infection and reverse transcription, but prior to integration. Although the infecting viral Gag protein as well as Fv1 determines the level of restriction, an ordered assembly of Gag is required for efficient Fv1 binding (88, 92). Our results suggest that Ty1 Gag interacts with p22/p18; however, the polymerization state of Gag and p22 required for maximum restriction of retrotransposition remains an open question. In addition, p22/p18 affects VLP assembly and function, whereas Fv1 inhibits a different step in the replication cycle that occurs post-infection. Conceptually similar to MLV-Fv1 restriction, the sheep genome harbors about 20 copies of endogenous (en) JSRVs and these

sequences are homologous with exogenous JSRV that can cause lung cancer. Certain endogenous copies have evolved a *trans*-dominant Gag protein enJS56A that like Ty1-p22 blocks replication at a step soon after protein synthesis. The JSRV-enJS56A interaction prevents Gag from entering into an endosome trafficking pathway, and results in aggregation and turnover by the proteasome (89, 93).

The MLV-Fv1 and JSRV-enJS56A restriction systems contain two components, raising the possibility of an arms race between the infecting retrovirus and the domesticated chromosomal *GAG* gene (94). In contrast, the many retrotransposition-competent Ty1 elements inhabiting *Saccharomyces* genomes encode their own inhibitor, and therefore, must balance mutations altering p22 potency with those affecting *GAG* fitness. Since Ty1 *GAG* or p22 coding regions have not been detected as an exapted gene capable of inhibiting Ty1 movement, the graduated retrotransposition rate provided by CNC may benefit *Saccharomyces* and Ty1, as suggested by recent work relating increases in Ty1 copy number with longer chronological lifespan (95). The Ty1-p22 interaction appears to directly block assembly of functional VLPs in a dose dependent manner, and to our knowledge represents a novel and effective way to allow some but not rampant retroelement movement. Further understanding of the molecular events underlying Ty1 Gag-p22 interaction, including the characterization of CNC-resistant mutants and the role that cellular genes have in modulating p22 expression or function, should reveal additional similarities and differences between Ty1 and retroviral restriction factors.

Acknowledgements

We acknowledge Karen Stefanisko for technical help at the beginning of this project and Mary Ard (University of Georgia, CAUR) for assistance with electron microscopy. We thank Thomas Mason (Hts1), Stephen Hajduk (TY tag), and Alan and Susan Kingsman (Ty1 VLP) for providing antisera, and Emiko Matsuda and Mary Ann Checkley for yeast strains. We thank Claiborne Glover, Steven Hajduk, William Lanzilotta, Michael Terns, Zachary Wood, and Shaying Zhang for sharing equipment. We thank Hyo Won Ahn and Katarzyna Purzycka for useful discussions. This work was supported by NIH grants GM095622 (D.J.G.) and GM081399 (W.V.G.). J.A.M. was supported in part by the National Science Foundation Graduate Research Fellowship 1011RH25213, and J.A.A. was supported by an NRSA Postdoctoral Fellowship 1F32GM112474-01 through NIGMS.

References

1. Voytas DF, Boeke JD. 2002. Ty1 and Ty5 of *Sacharomyces cerevisiae*., p. 614-630. *In* Craig NL, Craigie R, Gellert M, Lambowitz AM (ed.), Mobile DNA II. ASM Press, Washington, DC.

2. Dunham M, Badrane H, Ferea T, Adams J, Brown P, Rosenzweig F, Botstein D. 2002. Characteristic genome rearrangements in experimental evolution of *Saccharomyces cerevisiae.* Proceedings of the National Academy of Sciences USA 99:16144-16149.

3. Garfinkel DJ. 2005. Genome evolution mediated by Ty elements in *Saccharomyces*. Cytogenetic and Genome Research 110:63-69.

4. Wilke C, Adams J. 1992. Fitness effects of Ty transposition in *Saccharomyces cerevisiae.* Genetics 131:31-42.

5. Wilke CM, Maimer E, Adams J. 1992. The population biology and evolutionary significance of Ty elements in *Saccharomyces cerevisiae.* Genetica 86:155-173.

6. Beliakova-Bethell N, Beckham C, Giddings TH, Winey M, Parker R, Sandmeyer S. 2006. Virus-like particles of the Ty3 retrotransposon assemble in association with P-body components. RNA 12:94-101.

7. Checkley MA, Nagashima K, Lockett SJ, Nyswaner KM, Garfinkel DJ. 2010. P-body components are required for Ty1 retrotransposition during assembly of retrotransposition-competent virus-like particles. Molecular and Cellular Biology 30:382-398.

8. Dutko JA, Kenny AE, Gamache ER, Curcio MJ. 2010. 5' and 3' mRNA decay factors colocalize with Ty1 Gag and human APOBEC3G and promote Ty1 retrotransposition. Journal of Virology 84:5052-5066.

9. Malagon F, Jensen TH. 2008. The T body, a new cytoplasmic RNA granule in *Saccharomyces cerevisiae*. Molecular and Cellular Biology 28:6022-6032.

10. Bleykasten-Grosshans C, Friedrich A, Schacherer J. 2013. Genome-wide analysis of intraspecific transposon diversity in yeast. BMC Genomics 14:399.

11. Carr M, Bensasson D, Bergman CM. 2012. Evolutionary genomics of transposable elements in *Saccharomyces cerevisiae.* PLoS ONE 7:e50978.

12. Liti G, Peruffo A, James SA, Roberts IN, Louis EJ. 2005. Inferences of evolutionary relationships from a population survey of LTR-retrotransposons and telomeric-associated sequences in the *Saccharomyces sensu stricto* complex. Yeast 22:177-192.

13. Moore SP, Liti G, Stefanisko KM, Nyswaner KM, Chang C, Louis EJ, Garfinkel DJ. 2004. Analysis of a Ty1-less variant of *Saccharomyces paradoxus*: the gain and loss of Ty1 elements. Yeast 21:649-660.

14. Liti G, Carter DM, Moses AM, Warringer J, Parts L, James SA, Davey RP, Roberts IN, Burt A, Koufopanou V, Tsai IJ, Bergman CM, Bensasson D, O'kelly MJT, Van Oudenaarden A, Barton DBH, Bailes E, Nguyen AN, Jones M, Quail MA, Goodhead I, Sims S, Smith F, Blomberg A, Durbin R, Louis EJ. 2009. Population genomics of domestic and wild yeasts. Nature 458:337-341.

15. Boeke JD, Garfinkel DJ, Styles CA, Fink GR. 1985. Ty elements transpose through an RNA intermediate. Cell 40:491-500.

16. Scheifele LZ, Cost GJ, Zupancic ML, Caputo EM, Boeke JD. 2009. Retrotransposon overdose and genome integrity. Proceedings of the National Academy of Sciences USA 106:13927-13932.

17. Maxwell PH, Curcio MJ. 2007. Host factors that control long terminal repeat retrotransposons in *Saccharomyces cerevisiae*: implications for regulation of mammalian retroviruses. Eukaryotic Cell 6:1069-1080.

18. Dakshinamurthy A, Nyswaner KM, Farabaugh PJ, Garfinkel DJ. 2010. *BUD22* affects Ty1 retrotransposition and ribosome biogenesis in *Saccharomyces cerevisiae.* Genetics 185:1193-1105.

19. Nyswaner KM, Checkley MA, Yi M, Stephens RM, Garfinkel DJ. 2008. Chromatin-associated genes protect the yeast genome from Ty1 insertional mutagenesis. Genetics 178:197-214.

20. Risler JK, Kenny AE, Palumbo RJ, Gamache ER, Curcio MJ. 2012. Host co-factors of the retrovirus-like transposon Ty1. Mobile DNA 3:12.

21. Suzuki K, Morimoto M, Kondo C, Ohsumi Y. 2011. Selective autophagy regulates insertional mutagenesis by the Ty1 retrotransposon in *Saccharomyces cerevisiae*. Developmental Cell 21:358-365.

22. Winston F, Durbin KJ, Fink GR. 1984. The *SPT3* gene is required for normal transcription of Ty elements in *S. cerevisiae*. Cell 39:675-682.

23. Grant PA, Duggan L, Côté J, Roberts SM, Brownell JE, Candau R, Ohba R, Owen-Hughes T, Allis CD, Winston F, Berger SL, Workman JL. 1997. Yeast Gcn5 functions in two multisubunit complexes to acetylate nucleosomal histones: characterization of an Ada complex and the SAGA (Spt/Ada) complex. Genes & Development 11:1640-1650.

24. Berretta J, Pinskaya M, Morillon A. 2008. A cryptic unstable transcript mediates transcriptional trans-silencing of the Ty1 retrotransposon in *S. cerevisiae*. Genes & Development 22:615-626.

25. Haimovich G, Medina DA, Causse SZ, Garber M, Millán-Zambrano G, Barkai O, Chávez S, Pérez-Ortín JE, Darzacq X, Choder M. 2013. Gene expression is circular: factors for mRNA degradation also foster mRNA synthesis. Cell 153:1000-1011.

26. Parker R. 2012. RNA degradation in *Saccharomyces cerevisae.* Genetics 191:671-702.

27. Dumesic PA, Madhani HD. 2014. Recognizing the enemy within: licensing RNA-guided genome defense. Trends in Biochemical Sciences 39:25-34.

28. Garfinkel DJ, Nyswaner K, Wang J, Cho J-Y. 2003. Post-transcriptional cosuppression of Ty1 retrotransposition. Genetics 165:83-99.

29. Drinnenberg IA, Fink GR, Bartel DP. 2011. Compatibility with killer explains the rise of RNAi-deficient fungi. Science 333:1592.

30. Drinnenberg IA, Weinberg DE, Xie KT, Mower JP, Wolfe KH, Fink GR, Bartel DP. 2009. RNAi in budding yeast. Science 326:544-550.

31. Matsuda E, Garfinkel DJ. 2009. Posttranslational interference of Ty1 retrotransposition by antisense RNAs. Proceedings of the National Academy of Sciences USA 106:15657-15662.

32. Purzycka KJ, Legiewicz M, Matsuda E, Eizentstat LD, Lusvarghi S, Saha A, Le Grice SFJ, Garfinkel DJ. 2013. Exploring Ty1 retrotransposon RNA structure within virus-like particles. Nucleic Acids Research 41:463-473.

33. Curcio MJ, Garfinkel DJ. 1992. Posttranslational control of Ty1 retrotransposition occurs at the level of protein processing. Molecular and Cellular Biology 12:2813-2825.

34. Garfinkel DJ, Boeke JD, Fink GR. 1985. Ty element transposition: Reverse transcriptase and virus-like particles. Cell 42:502-517.

35. Curcio MJ, Garfinkel DJ. 1999. New lines of host defense: inhibition of Ty1 retrotransposition by Fus3p and NER/TFIIH. Trends in Genetics 15:43-45.

36. Sanz-Ramos M, Stoye JP. 2013. Capsid-binding retrovirus restriction factors: discovery, restriction specificity and implications for the development of novel therapeutics. Journal of General Virology 94:2587-2598.

37. Spencer TE, Palmarini M. 2012. Endogenous retroviruses of sheep: a model system for understanding physiological adaptation to an evolving ruminant genome. Journal of Reproduction and Development 58:33-37.

38. Guthrie C, Fink GR. 1991. Guide to Yeast Genetics and Molecular Biology, p. 933, Methods in Enzymology, vol. 194. Academic Press Inc., San Diego, CA.

39. Boeke JD, Eichinger D, Castrillon D, Fink GR. 1988. The *Saccharomyces cerevisiae* genome contains functional and nonfunctional copies of transposon Ty1. Molecular and Cellular Biology 8:1432-1442.

40. Merkulov GV, Lawler JF, Eby Y, Boeke JD. 2001. Ty1 proteolytic cleavage sites are required for transposition: all sites are not created equal. Journal of Virology 75:638-644.

41. Mitchell DA, Marshall TK, Deschenes RJ. 1993. Vectors for the inducible overexpression of glutathione S-transferase fusion proteins in yeast. Yeast 9:715-722.

42. Vaidyanathan PP, Zinshteyn B, Thompson MK, Gilbert WV. 2014. Protein kinase A regulates gene-specific translational adaptation in differentiating yeast. RNA 20:912-922.

43. Ingolia NT, Ghaemmaghami S, Newman JR, Weissman JS. 2009. Genome-wide analysis *in vivo* of translation with nucleotide resolution using ribosome profiling. Science 324:218-223.

44. Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, Batut P, Chaisson M, Gingeras TR. 2012. STAR: ultrafast universal RNA-seq aligner. Bioinformatics 29:15-21.

45. Liu H, Krizek J, Bretscher A. 1992. Construction of a *GAL1*-regulated yeast cDNA expression library and its application to the identification of genes whose overexpression causes lethality in yeast. Genetics 132:665-673.

46. Curcio MJ, Garfinkel DJ. 1991. Single-step selection for Ty1 element retrotransposition. Proceedings of the National Academy of Sciences USA 88:936-940.

47. Lawler JF, Merkulov GV, Boeke JD. 2002. A nucleocapsid functionality contained within the amino terminus of the Ty1 protease that is distinct and separable from proteolytic activity. Journal of Virology 76:346-354.

48. Garfinkel DJ, Hedge AM, Youngren SD, Copeland TD. 1991. Proteolytic processing of Pol-TYB proteins from the yeast retrotransposon Ty1. Journal of Virology 65:4573-4581.

49. Eichinger DJ, Boeke JD. 1988. The DNA intermediate in yeast Ty1 element transposition copurifies with virus-like particles: cell-free Ty1 transposition. Cell 54:955-966.

50. Youngren SD, Boeke JD, Sanders NJ, Garfinkel DJ. 1988. Functional organization of the retrotransposon Ty from *Saccharomyces cerevisiae*: Ty protease is required for transposition. Molecular and Cellular Biology 8:1421-1431.

51. Schneider CA, Rasband WS, Eliceiri KW. 2012. NIH Image to ImageJ: 25 years of image analysis. Nature Methods 9:671-675.

52. Melamed C, Nevo Y, Kupiec M. 1992. Involvement of cDNA in homologous recombination between Ty elements in *Saccharomyces cerevisiae.* Molecular and Cellular Biology 12:1613-1620.

53. Sharon G, Burkett TJ, Garfinkel DJ. 1994. Efficient homologous recombination of Ty1 element cDNA when integration is blocked. Molecular and Cellular Biology 14:6540-6551.

54. Winston F, Dollard C, Malone EA, Clare J, Kapakos JG, Farabaugh P, Minehart PL. 1987. Three genes are required for trans-activation of Ty transcription in yeast. Genetics 115:649-656.

55. Elder RT, Loh EY, Davis RW. 1983. RNA from the yeast transposable element Ty1 has both ends in the direct repeats, a structure similar to retrovirus RNA. Proceedings of the National Academy of Sciences USA 80:2432-2436.

56. Adams SE, Mellor J, Gull K, Sim RB, Tuite MF, Kingsman SM, Kingsman AJ. 1987. The functions and relationships of Ty-VLP proteins in yeast reflect those of mammalian retroviral proteins. Cell 49:111-119.

57. Arribere JA, Gilbert WV. 2013. Roles for transcript leaders in translation and mRNA decay revealed by transcript leader sequencing. Genome Research 23:977-987.

58. Curcio MJ, Hedge AM, Boeke JD, Garfinkel DJ. 1990. Ty RNA levels determine the spectrum of retrotransposition events that activate gene expression in *Saccharomyces cerevisiae*. Molecular and General Genetics 220:213-221.

59. Servant G, Pennetier C, Lesage P. 2008. Remodeling yeast gene transcription by activating the Ty1 long terminal repeat retrotransposon under severe adenine deficiency. Molecular and Cellular Biology 28:5543-5554.

60. Servant G, Pinson B, Tchalikian-Cosson A, Coulpier F, Lemoine S, Pennetier C, Bridier-Nahmias A, Todeschini A-L, Fayol H, Daignan-Fornier B, Lesage P. 2012. Tye7 regulates yeast Ty1 retrotransposon sense and antisense transcription in response to adenylic nucleotides stress. Nucleic Acids Research.

61. Belotserkovskaya R, Sterner DE, Deng M, Sayre MH, Lieberman PM, Berger SL. 2000. Inhibition of TATA-binding protein function by SAGA subunits Spt3 and Spt8 at Gcn4-activated promoters. Molecular and Cellular Biology 20:634-647.

62. Eisenmann DM, Arndt KM, Ricupero SL, Rooney JW, Winston F. 1992. Spt3 interacts with TFIID to allow normal transcription in *Saccharomyces cerevisiae*. Genes & Development 6:1319-1331.

63. Mohibullah N, Hahn S. 2008. Site-specific cross-linking of TBP *in vivo* and *in vitro* reveals a direct functional interaction with the SAGA subunit Spt3. Genes & Development 22:2994-3006.

64. Cheung V, Chua G, Batada NN, Landry CR, Michnick SW, Hughes TR, Winston F. 2008. Chromatin and transcription-related factors repress transcription from within coding regions throughout the *Saccharomyces cerevisiae* genome. PLoS Biology 6:e277.

65. Gilbert WV. 2010. Alternative ways to think about cellular internal ribosome entry. Journal of Biological Chemistry 285:29033-29038.

66. Harrison PM, Kumar A, Lang N, Snyder M, Gerstein M. 2002. A question of size: the eukaryotic proteome and the problems in defining it. Nucleic Acids Research 30:1083-1090.

67. Kochetov AV. 2008. Alternative translation start sites and hidden coding potential of eukaryotic mRNAs. BioEssays : news and reviews in molecular, cellular and developmental biology 30:683-691.

68. Kochetov AV, Sarai A, Rogozin IB, Shumny VK, Kolchanov NA. 2005. The role of alternative translation start sites in the generation of human protein diversity. Molecular genetics and genomics : MGG 273:491-496.

69. Wang XQ, Rothnagel JA. 2004. 5'-untranslated regions with multiple upstream AUG codons can support low-level translation via leaky scanning and reinitiation. Nucleic Acids Research 32:1382-1391.

70. Kochetov AV. 2005. AUG codons at the beginning of protein coding sequences are frequent in eukaryotic mRNAs with a suboptimal start codon context. Bioinformatics 21:837-840.

71. Porras P, Padilla CA, Krayl M, Voos W, Barcena JA. 2006. One single in-frame AUG codon is responsible for a diversity of subcellular localizations of glutaredoxin 2 in *Saccharomyces cerevisiae.* Journal of Biological Chemistry 281:16551-16562.

72. Garfinkel DJ, Nyswaner KM, Stefanisko KM, Chang C, Moore SP. 2005. Ty1 copy number dynamics in *Saccharomyces*. Genetics 169:1845-1857.

73. Curcio MJ, Garfinkel DJ. 1994. Heterogeneous functional Ty1 elements are abundant in the *Saccharomyces cerevisiae* genome. Genetics 136:1245-1259.

74. Farabaugh P. 1995. Post-transcriptional regulation of transposition by Ty retrotransposons of *Saccharomyces cerevisiae*. The Journal of Biological Chemistry 270:10361-10364.

75. Fink G, Boeke J, Garfinkel D. 1986. The mechanism and consequences of retrotransposition. Trends in Genetics 2:118-123.

76. Malagon F, Jensen TH. 2011. T-body formation precedes virus-like particle maturation in *S. cerevisiae*. RNA Biology 8:184-189.

77. Clemens K, Larsen L, Zhang M, Kuznetsov Y, Bilanchone V, Randall A, Harned A, DaSilva R, Nagashima K, McPherson A, Baldi P, Sandmeyer S. 2011. The Ty3 Gag3 spacer controls intracellular condensation and uncoating. Journal of Virology 85:3055-3066.

78. Larsen LS, Beliakova-Bethell N, Bilanchone V, Zhang M, Lamsa A, Dasilva R, Hatfield GW, Nagashima K, Sandmeyer S. 2008. Ty3 nucleocapsid controls localization of particle assembly. Journal of Virology 82:2501-2514.

79. Doh JH, Lutz S, Curcio MJ. 2014. Co-translational localization of an LTR-retrotransposon RNA to the endoplasmic reticulum nucleates virus-like particle assembly sites. PLoS Genetics 10:e1004219.

80. Braiterman LT, Monokian GM, Eichinger DJ, Merbs SL, Gabriel A, Boeke JD. 1994. In-frame linker insertion mutagenesis of yeast transposon Ty1: phenotypic analysis. Gene 139:19-26.

81. Checkley MA, Mitchell JA, Eizenstat LD, Lockett SJ, Garfinkel DJ. 2013. Ty1 Gag enhances the stability and nuclear export of Ty1 mRNA. Traffic 14:57-69.

82. Monokian GM, Braiterman LT, Boeke JD. 1994. In-frame linker insertion mutagenesis of yeast transposon Ty1: mutations, transposition and dominance. Gene 139:9-18.

83. Martin-Rendon E, Marfany G, Wilson S, Ferguson DJ, Kingsman SM, Kingsman AJ. 1996. Structural determinants within the subunit protein of Ty1 virus-like particles. Molecular Microbiology 22:667-679.

84. Cristofari G, Ficheux D, Darlix J-L. 2000. The Gag-like protein of the yeast Ty1 retrotransposon contains a nucleic acid chaperone domain analogous to retroviral nucleocapsid proteins. Journal of Biological Chemistry 275:19210-19217.

85. Rein A, Datta SAK, Jones CP, Musier-Forsyth K. 2011. Diverse interactions of retroviral Gag proteins with RNAs. Trends in Biochemical Sciences 36:373-380.

86. Kaneko-Ishino T, Ishino F. 2012. The role of genes domesticated from LTR retrotransposons and retroviruses in mammals. Frontiers in Microbiology 3:262.

87. Best S, Le Tissier P, Towers G, Stoye JP. 1996. Positional cloning of the mouse retrovirus restriction gene Fv1. Nature 382:826-829.

88. Hilditch L, Matadeen R, Goldstone DC, Rosenthal PB, Taylor IA, Stoye JP. 2011. Ordered assembly of murine leukemia virus capsid protein on lipid nanotubes directs specific binding by the restriction factor, Fv1. Proceedings of the National Academy of Sciences USA 108:5771-5776.

89. Murcia PR, Arnaud F, Palmarini M. 2007. The transdominant endogenous retrovirus enJS56A1 associates with and blocks intracellular trafficking of Jaagsiekte sheep retrovirus Gag. Journal of Virology 81:1762-1772.

90. Bénit L, De Parseval N, Casella JF, Callebaut I, Cordonnier A, Heidmann T. 1997. 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. Journal of Virology 71:5652-5657.

91. Qi CF, Bonhomme F, Buckler-White A, Buckler C, Orth A, Lander MR, Chattopadhyay SK, Morse HC. 1998. Molecular phylogeny of Fv1. Mammalian Genome 9:1049-1055.

92. Goldstone DC, Walker PA, Calder LJ, Coombs PJ, Kirkpatrick J, Ball NJ, Hilditch L, Yap MW, Rosenthal PB, Stoye JP, Taylor IA. 2014. Structural studies of postentry restriction factors reveal antiparallel dimers that enable avid binding to the HIV-1 capsid lattice. Proceedings of the National Academy of Sciences USA 111:9609-9614.

93. Arnaud F, Murcia PR, Palmarini M. 2007. Mechanisms of late restriction induced by an endogenous retrovirus. Journal of Virology 81:11441-11451.

94. Jern P, Coffin J. 2008. Effects of retroviruses on host genome function. Annual Review of Genetics 42:709-732.

95. VanHoute D, Maxwell PH. 2014. Extension of *Saccharomyces paradoxus* chronological lifespan by retrotransposons in certain media conditions is associated with changes in reactive oxygen species. Genetics 198:531-545.

96. Brachmann C, Davies A, Cost G, Caputo E, Li J, Hieter P, Boeke J. 1998. Designer deletion strains derived from *Saccharomyces cerevisiae* S288C: a useful set of strains and plasmids for PCR-mediated gene disruption and other applications. Yeast 14:115-132.

Figures

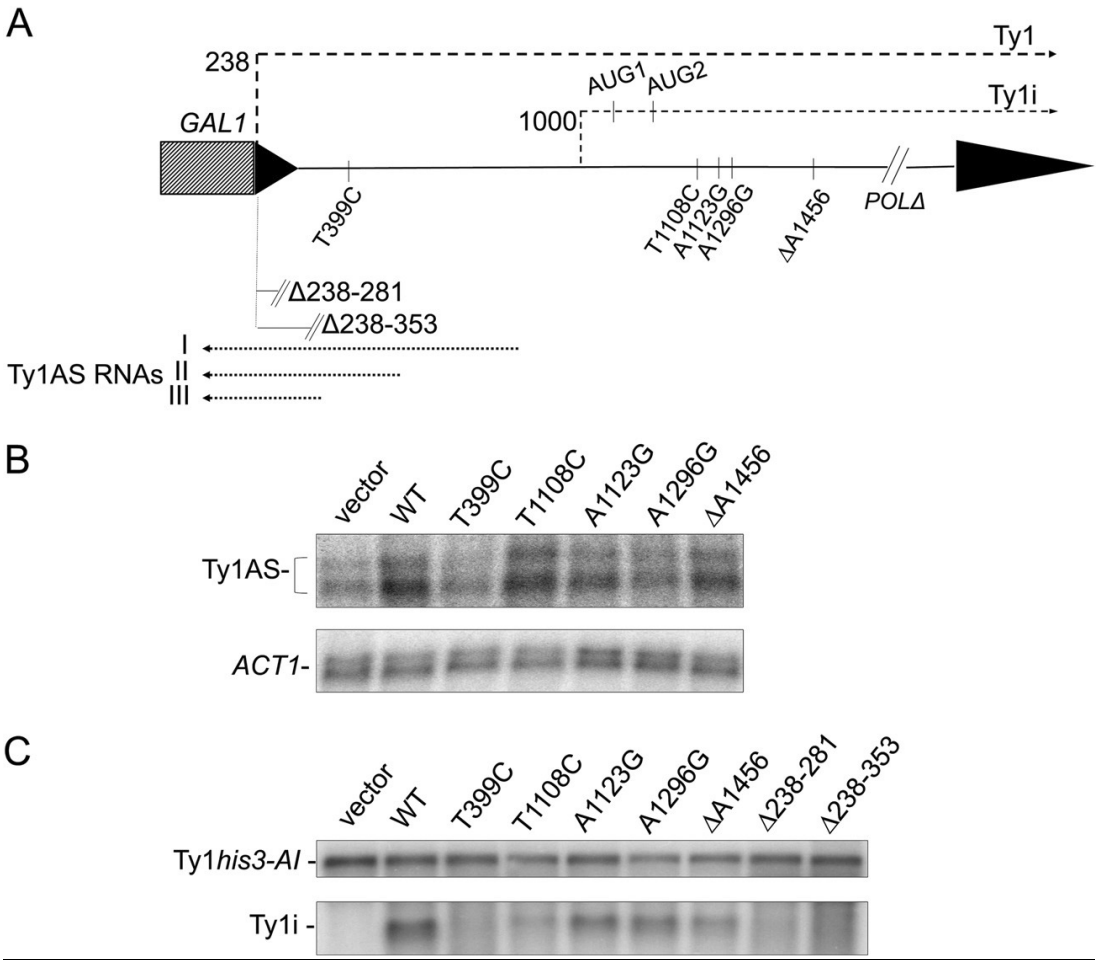

Figure 2-1 An internal Ty1i transcript is involved in CNC. (A) Functional organization of the Ty1 CNC region, which covers *GAG* and the beginning of *POL*. Location of the *GAL1* promoter (hatched rectangle), LTR (solid triangle), Ty1 transcripts (Figure 3.4), candidate initiation codons present on Ty1i RNA, and CNC$^-$ defective deletions and point mutations are noted. Ty1AS RNAs I, II and III are shown with dotted lines. Ty1AS RNAs share a common 3' end at nt 136 but have different 5' ends: nt 760 for II and 594 for III. The exact 5' end of Ty1AS RNA I has not been determined (31). (B) Total RNA from a Ty1-less strain with a single chromosomal Ty1*his3-AI* element containing empty vector, wild-type pGPOLΔ (DG2374), or mutant plasmids T399C (YAS73), T1108C

(YAS69), A1123G (YAS71), A1296G (YAS72), ΔA1456 (YAS70) was analyzed by Northern blotting to detect Ty1AS RNAs. Cells were grown in glucose and Ty1 strand-specific (nt 238-1702) and *ACT1* ^{32}P-labeled riboprobes were used. (C) Total RNA from the strains in (B), plus two additional strains containing mutant plasmids Δ238-281 (YAS74) and Δ238-353 (YAS75), was probed for Ty1i transcripts. Ty1*his3-AI* served as a loading control.

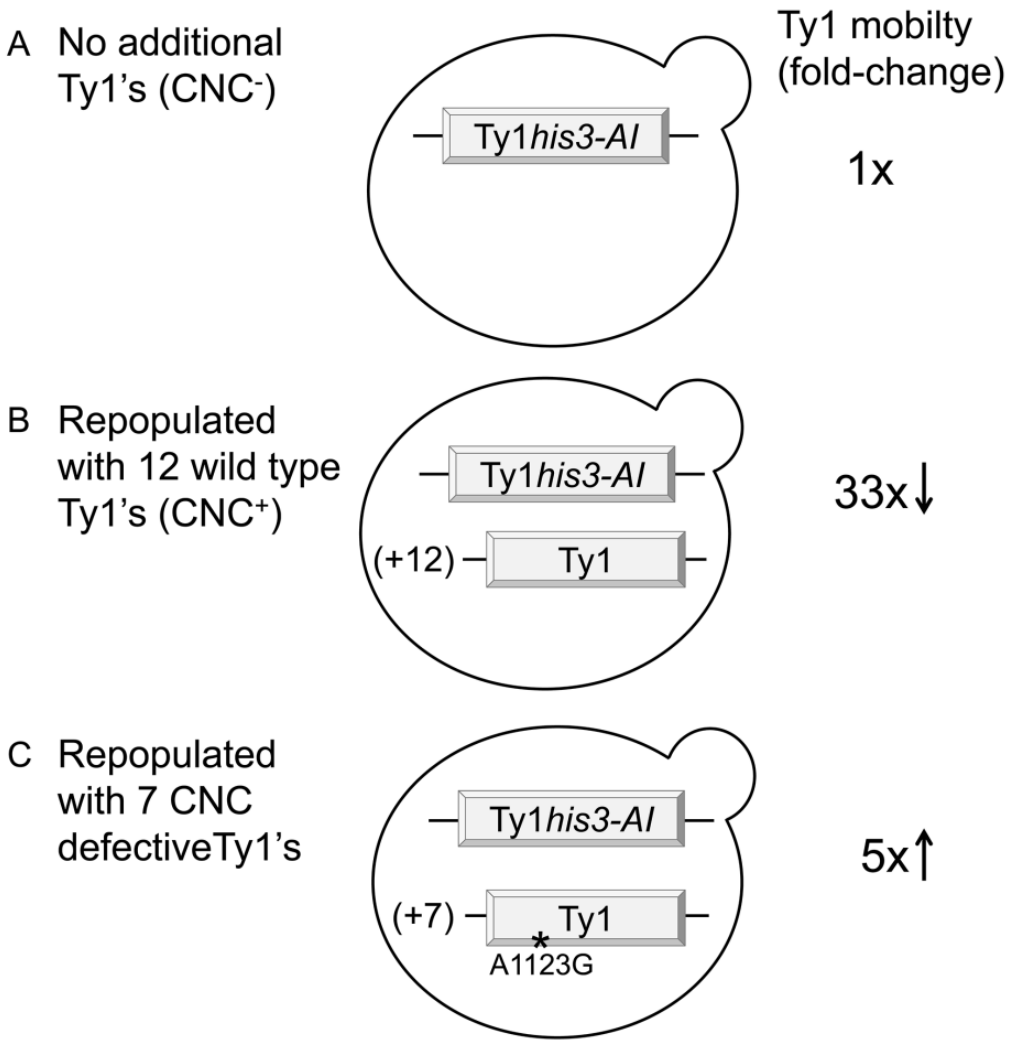

Figure 2-2 Chromosomal Ty1A1123G insertions do not confer CNC. Ty1-less *S. paradoxus* containing a single chromosomal Ty1*his3-AI* (A) was repopulated with unmarked, wild type (B) or A1123G (C) Ty1 elements. Genome repopulation with 12 wild type Ty1 elements resulted in an overall decrease in Ty1*his3-AI* mobility, while repopulation with 7 CNC^- mutant Ty1A1123G elements resulted in an overall increase in Ty1*his3-AI* mobility. Also refer to Table 2B.

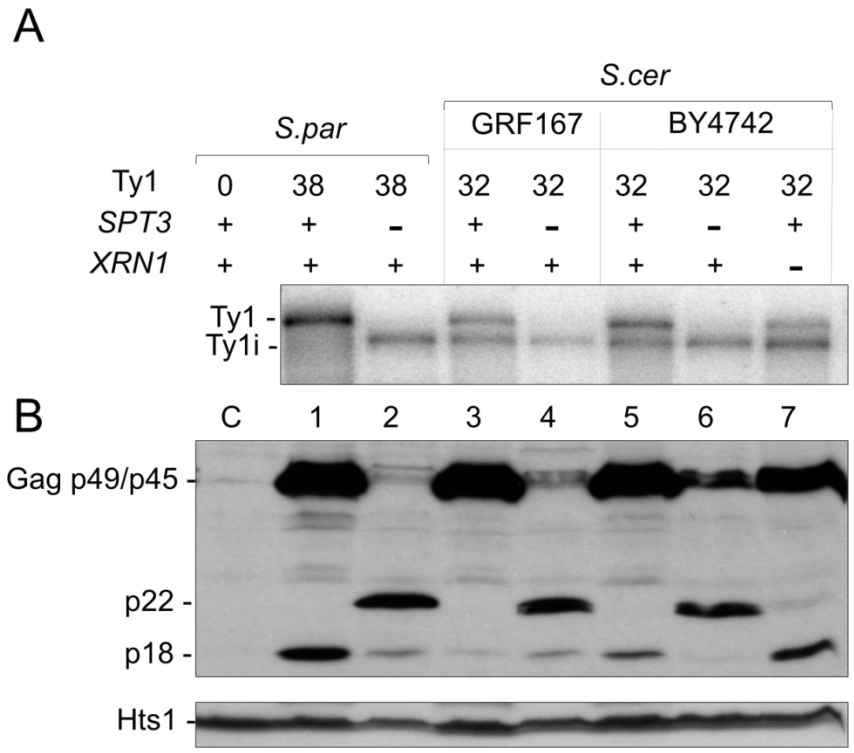

Figure 2-3 Detecting Ty1i RNA and p22/p18 from chromosomal Ty1 elements. (A) Northern blotting of poly(A)$^+$ RNA from *S. paradoxus* and *S. cerevisiae* (GRF167 and BY4742) wild-type, *spt3Δ* (DG789 and DG2247) and *xrn1Δ* (MAC103) mutant strains. A Ty1 ^{32}P-labeled riboprobe (nt 1266-1601) hybridized with full-length Ty1 and Ty1i transcripts. (B) Total protein extracts were immunoblotted with the p18 antiserum to detect full-length Gag p49/p45 and p22/18. A Ty1-less *S. paradoxus* strain (DG1768)

and cellular histidyl tRNA synthetase (Hts1) served as negative (lane C) and loading controls, respectively.

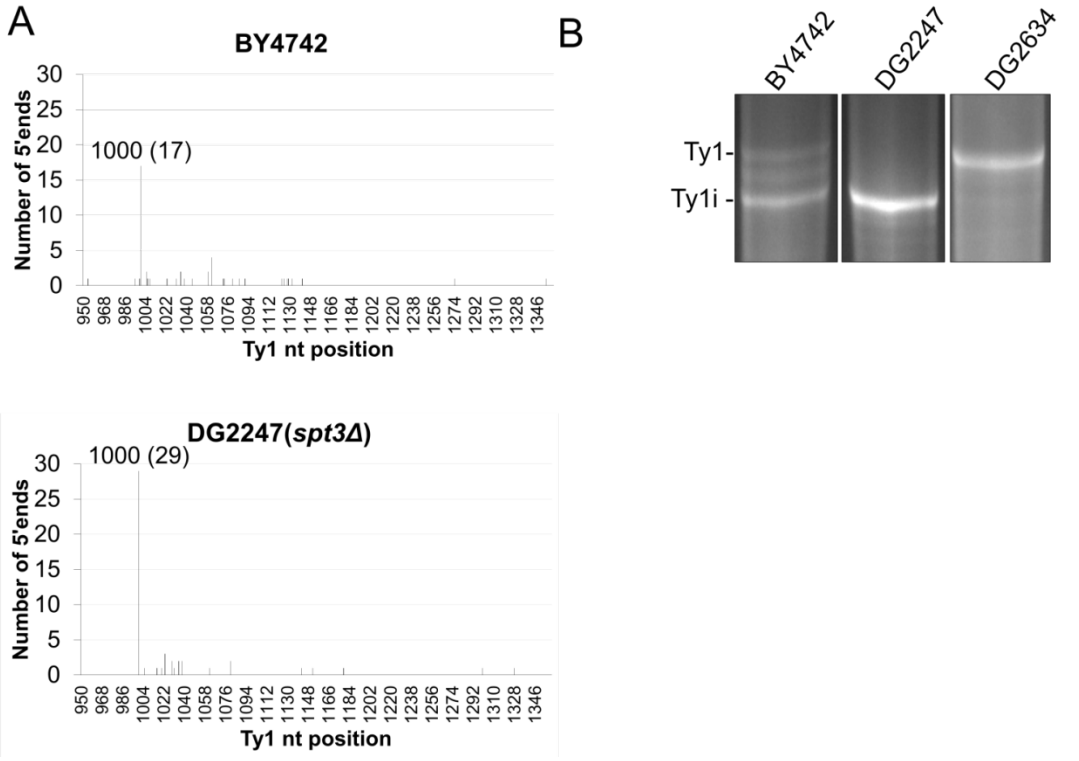

Figure 2-4 The major 5' end of the 4.9 kb Ty1i RNA maps to nt 1000. (A) Cap-independent 5'- RACE was performed with poly(A)$^+$ RNA from wild-type BY4742 and an isogenic *spt3Δ* mutant (DG2247). The number of 5' termini was plotted against the Ty1H3 sequence and the distribution of the termini are on the X and Y-axis respectively. The tallest peak represents the total number of 5' ends captured at nt 1000 and is shown in parentheses. (B) 5' RACE cDNA libraries from the wild type and *spt3*Δ strains mentioned above, and a repopulated *S. paradoxus* strain (DG2634) were amplified using a universal primer mix and a Ty1-specific primer GSP1_3389. The amplification reactions were separated by agarose gel electrophoresis to demonstrate the presence

of cDNA products corresponding to the 5' ends of the full-length (5.7 kb) Ty1 and the truncated (4.9 kb) Ty1i RNAs.

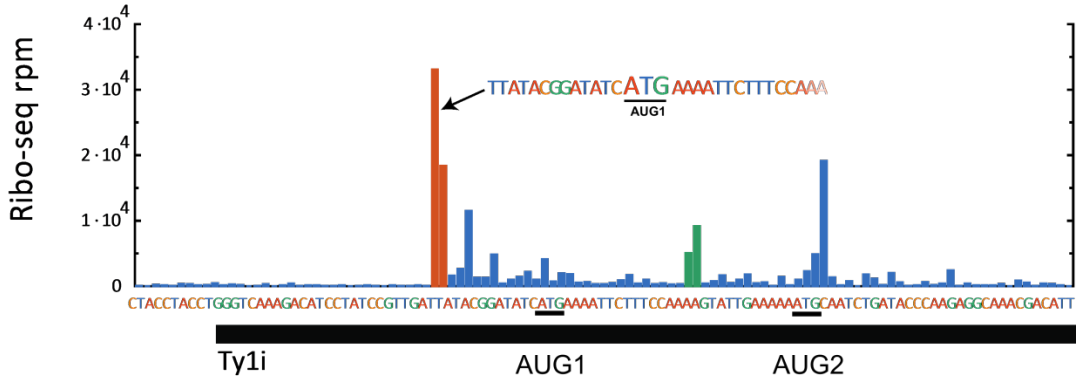

Figure 2-5 Whole genome analysis of internal translation initiation sites. Ribosome footprint profiling (Ribo-seq) was performed to detect translation initiation at internal AUG codons, two of which (AUG1 and AUG2, see Figure 3.1) are located immediately downstream of the Ty1i RNA transcription start site. Reads per million (rpm) Ty1-mapped reads were placed on the Ty1H3 sequence and the 5' end of ribosome footprints aligned downstream of the Ty1i transcription start are shown. Ribo-seq reads with 5' ends 12-13 nt upstream of AUG1 and AUG2 are highlighted in orange and green, respectively. ~12 nt downstream of the 5' end corresponds to the ribosomal P-site. Because these libraries were prepared with poly(A) tailing, the exact 3'-end of the footprint, and thus the footprint size at AUG1, is ambiguous but within the range of 26-30 nt, inclusive.

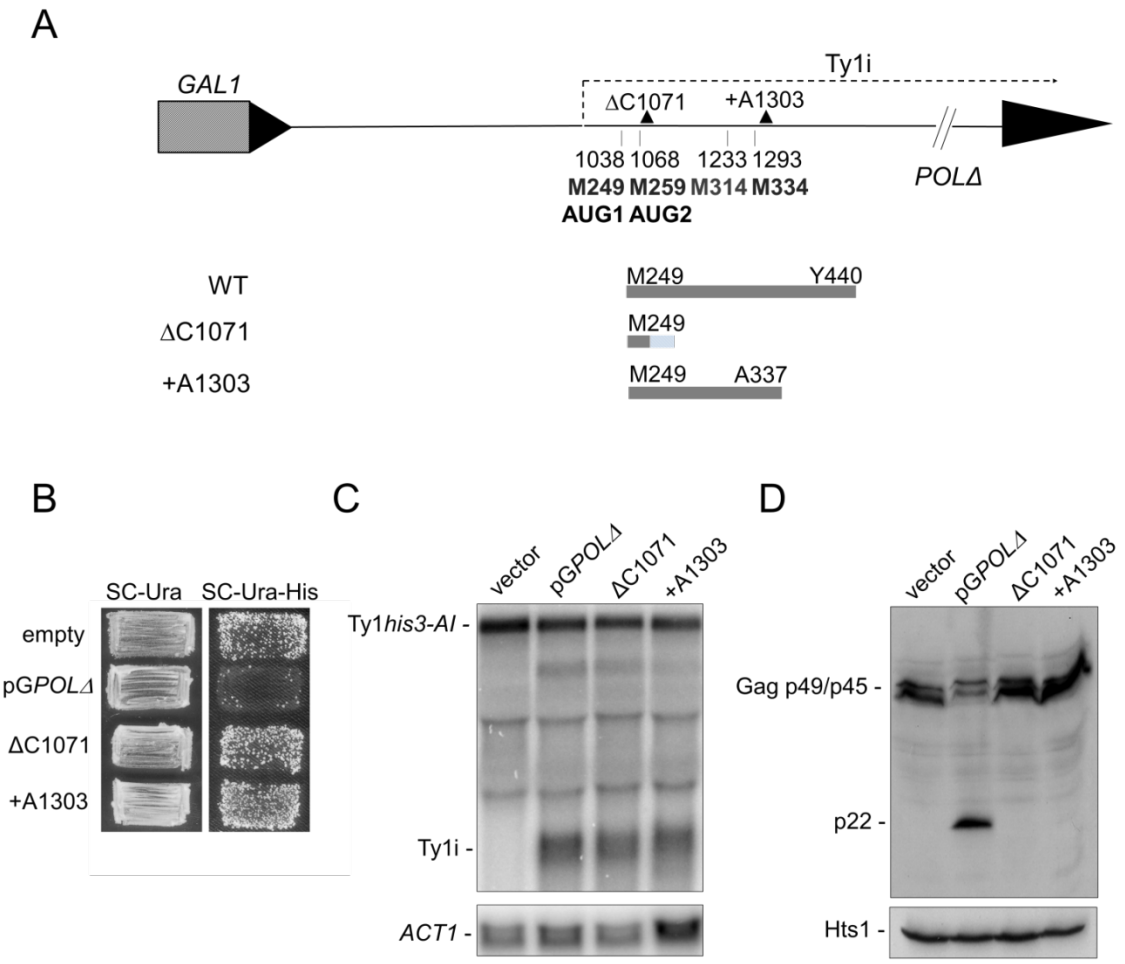

Figure 2-6 p22 is necessary for CNC. (A) Ty1 sequence present on pGPOLΔ illustrating the Ty1i RNA transcription start site (nt 1000), location of in-frame AUGs, and frameshift mutations (ΔC1071 and +A1303, black circles). Proteins encoded by wild-type (WT) or mutant plasmids are shown (wild-type sequence, solid; nonsense sequence, dashed) based on predicted usage of AUG1 by Ribo-seq (Figure 3.5). ΔC1071 and +A1303 are predicted to synthesize truncated p22 peptides of 11 and 89 residues, respectively, before encountering the frameshift mutation. (B) A *S. paradoxus* strain with a single chromosomal Ty1*his3-AI* carrying an empty vector (DG2411), pGPOLΔ (DG2374) or the mutant plasmids ΔC1071 (JM321) and +A1303 (JM320) were assessed for Ty1 mobility

using a qualitative assay. Cell patches grown on SC-Ura medium at 22^{o}C were replica plated to SC-Ura-His medium to select for cells that contain at least one Ty1*HIS3* insertion. The number of His^{+} papillae that grew on SC-Ura-His medium is a read-out for Ty1 mobility. Also refer to Table 2 C-1. (C) Total RNA from the strains described above was subjected to Northern blotting to detect Ty1*his3-AI* and Ty1i transcripts as described in Figure 3.1. The band labeled with an asterisk is a pervasive transcript approximately 4.5 kb in length, and contains both Ty1 and non-Ty1 sequences from the pGPOLΔ. The 'r' represents compression bands formed by two main species of ribosomal RNA in yeast, the 26S (3.8 kb) and 18S (2 kb) rRNAs. (D) Total cell extracts were analyzed for the presence of p22/p18 as described in Figure 2-3.

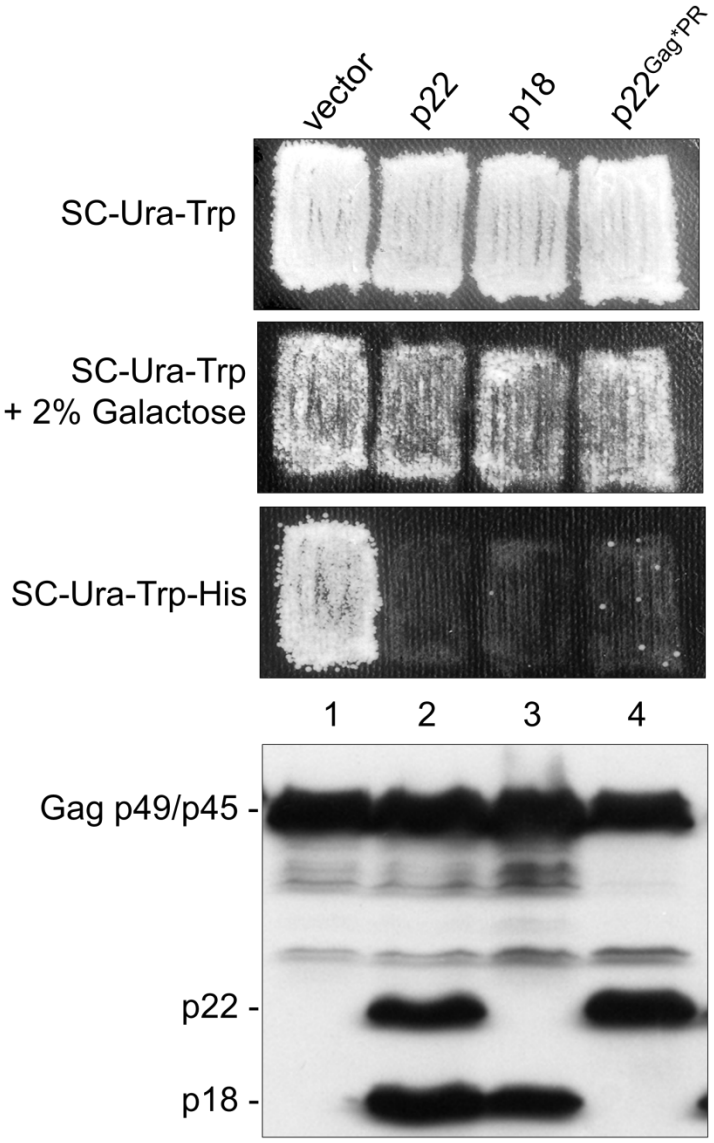

Figure 2-7 Cleavage of p22 to p18 does not disturb *trans*-dominant inhibition of Ty1 mobility. A mutant Gag-PR cleavage site, AAGSAA (Gag*PR) (40), was inserted into p22, replacing the normal Gag-PR cleavage site, RAHNVS. A Ty1-less strain containing pGTy1*his3-AI* and an empty vector (DG3739; lane 1), or *GAL1*-p22 (DG3774, lane 2), *GAL1*-p18 (DG3791, lane 3), and *GAL1*-p22^{Gag*PR} (JM399, lane 4) were analyzed for Ty1*his3-AI* mobility using a qualitative assay. Cell patches from a single colony were induced for pGTy1 expression by replica plating from SC-Ura-Trp to SC-Ura-Trp + 2%

galactose medium for 2 days at 22^{o}C. To detect Ty1*his3-AI* mobility, galactose-induced cells were replica plated to SC-Ura-His medium. Below is an immunoblot using total cell extracts from the same strains and the p18 antiserum to detect Gag-p49/p45 and p22/p18.

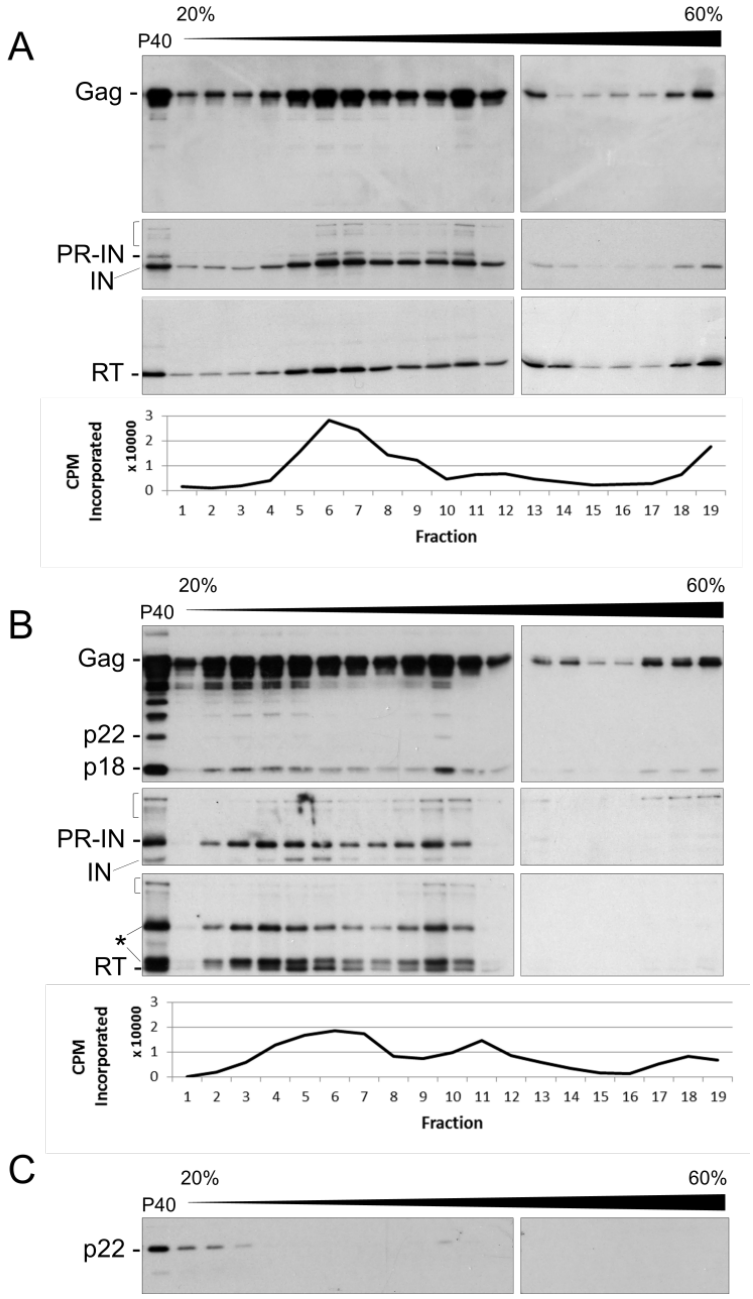

Figure 2-8 Cofractionation of p22/p18 with Ty1 VLPs. Crude VLP pellets (P40) prepared from galactose-induced Ty1-less strains expressing pGTy1*his3-AI* alone (A; DG3739), pGTy1*his3-AI* and p22 (B; DG3774) or p22 alone (C; DG3784) were fractionated through a 20-60% continuous sucrose gradient. VLP pellets (P40) and equal volumes from collected fractions were analyzed by immunoblotting with p18

antiserum, IN and RT antisera. Ty1 proteins are labeled, brackets indicate known Ty1 processing intermediates, and asterisks indicate aberrant Ty1 proteins (estimated size: 65 and 90 kD). Reverse transcriptase activity was detected using an exogenous poly(rC)-oligo/(dG) template and [α^{32}P]-dGTP.

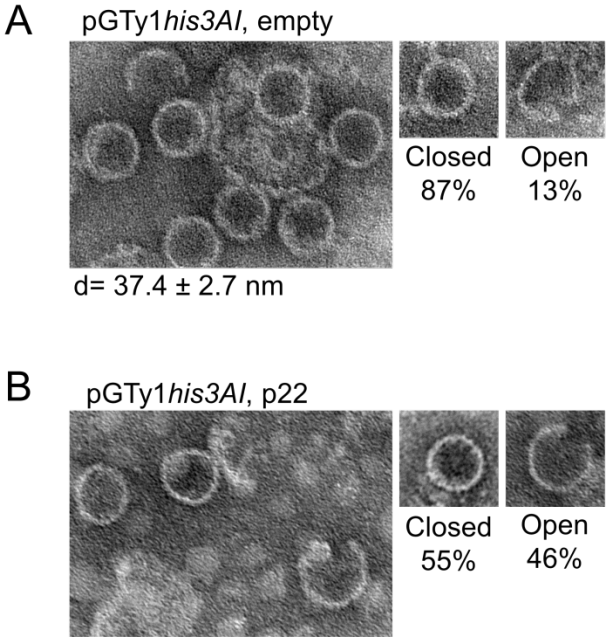

Figure 2-9 Electron microscopy of Ty1 VLPs assembled in the presence of p22/p18. VLP pellets were collected from sucrose gradient fractions with peak reverse transcriptase activity from experiments similar to those shown in Figure 3.8. VLPs from pGTy1*his3-A*I alone (A; DG3739) or pGTy1*his3-AI* and p22 (B; DG3774) were stained with 2% ammonium molybdate and examined by transmission electron microscopy. Approximately 100 VLPs were analyzed for closed versus open particles and representative images are shown. The diameter (d) was measured with closed VLPs only.

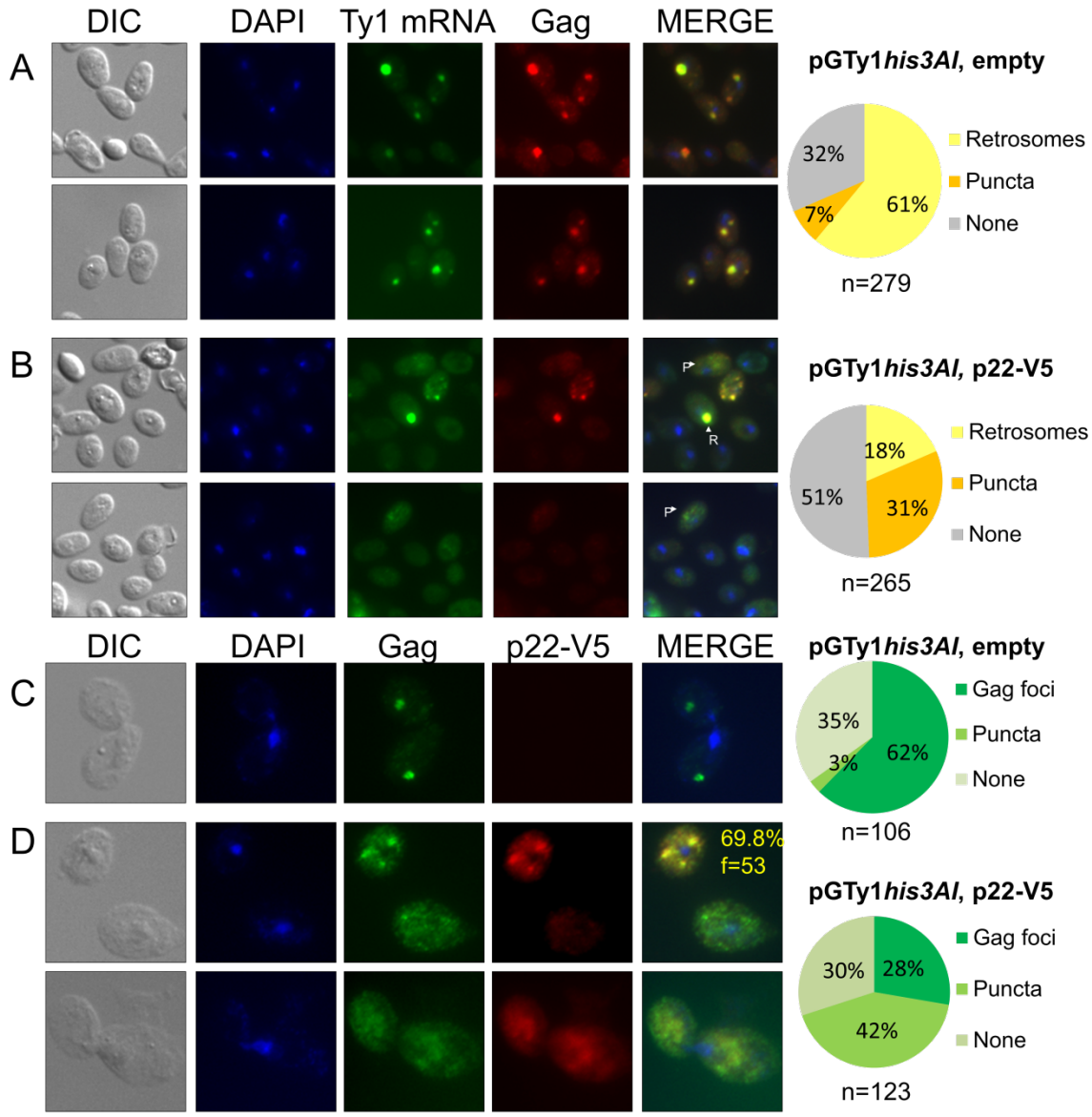

Figure 2-10 p22-V5 disrupts retrosomes and colocalizes with Gag. Ty1-less strains expressing pGTy1*his3-AI* alone (A, C; DG3739) or pGTy1*his3-AI* and p22-V5 together (B, D; JM367) were galactose-induced and analyzed for Ty1 mRNA and Gag colocalization via FISH/IF (panels A and B). Pie charts depict cells examined for the appearance of retrosomes (R), puncta (P), or no staining (None). Refer to the text for

additional details. (C and D) In a separate experiment, cells were analyzed for Ty1 Gag and p22-V5 colocalization via IF using VLP and V5 antibodies, respectively. The experiment in panel D was additionally analyzed for the percentage of Gag foci that colocalize with p22-V5 (yellow, f = total Gag foci analyzed). For both experiments, DNA was stained with DAPI and representative images are shown (n= number of cells analyzed).

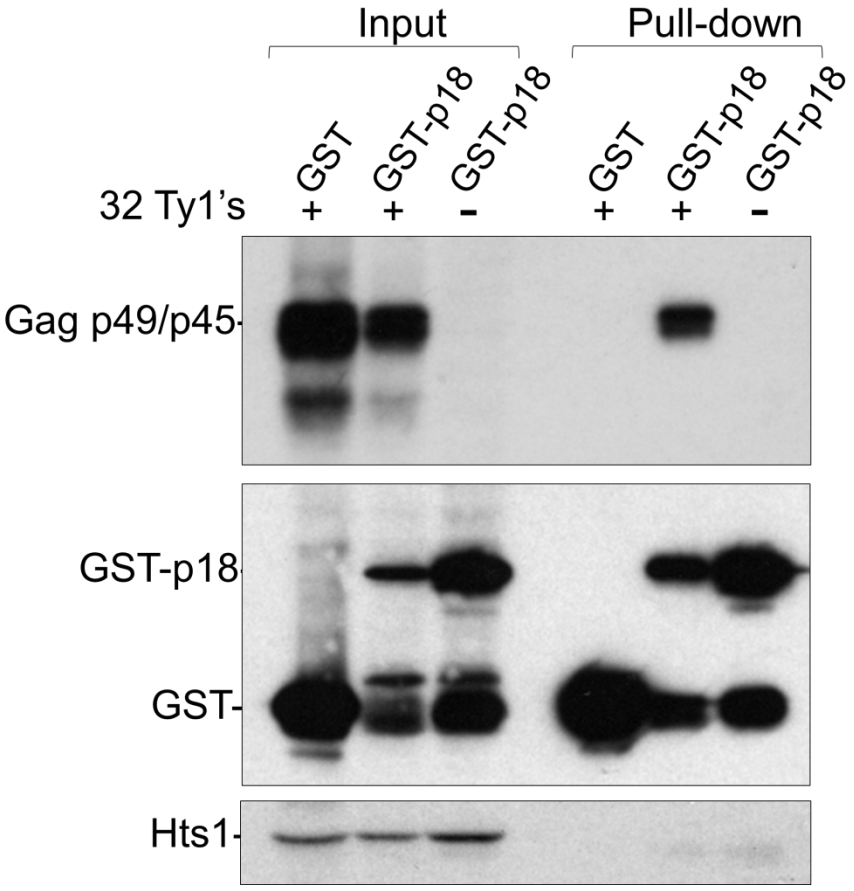

Figure 2-11 GST-p18 interacts with endogenous Ty1 Gag. Protein extracts (Input) from BY4742 induced for expression of GST (DG3808) or GST-p18 (DG3809) were incubated with glutathione-coated resin. Bound proteins were analyzed by

immunoblotting to detect Gag, GST-p18, and p18/Ty1 Gag complexes (Pull-down) after extensive washing with lysis buffer. A Ty1-less strain expressing GST-p18 (DG3810) and the presence of Hts1 served as negative controls. Gag was detected with TY tag monoclonal antibody, which recognizes p49/p45 but not p22/p18 due to the location of the epitope. GST proteins and Hts1 were detected with GST and Hts1 antibodies, respectively.

Tables

Table 2-1 Yeast strains

Strain	Genotype	Plasmid	Source
DG2196	*MATa*, *his3-Δ200hisG*, *ura3*, *trp1*, Ty1-less, Ty1*his3-AI(96)*	-	(28)
DG2254	DG2196	pGAL/2μ-*URA3*	(28)
DG2255	DG2196	pGTy1/2μ-*URA3*	(28)
DG2411	DG2196	empty/2μ-*URA3*	(28)
DG2374	DG2196	pBDG1130 (pGPOLΔ/2μ-*URA3*)	This study
YAS73	DG2196	pBAS39 (pGPOLΔ-T399C)	This study
YAS69	DG2196	pBAS38 (pGPOLΔ-T1108C)	This study
YAS71	DG2196	pBAS35 (pGPOLΔ-A1123G)	This study
YAS72	DG2196	pBAS36 (pGPOLΔ-A1296G)	This study

YAS70	DG2196	pBAS34 (pGPOLΔ-ΔA1456)	This study
YAS74	DG2196	pBAS43 (pGPOLΔ-Δ238-281)	This study
YAS75	DG2196	pBAS44 (pGPOLΔ-Δ238-353)	This study
JM321	DG2196	pBJM79 (pGPOLΔ-ΔC1071)	This study
JM320	DG2196	pBJM78 (pGPOLΔ- +A1303)	This study
DG2511	DG2196 +12 Ty1		This study
DG3856	DG2196	pBDG1595 (pGPOLΔ-GCG1GCG2)	This study
DG2512	DG2196 +9 Ty1		This study
DG3798	DG2196 +7 Ty1-A1123G		This study
DG1768	*MATα*, *his3-Δ200hisG*, *ura3*, Ty1-less		(28)
DG2533	DG1768, *Ty1–4253his3-AI*		(3)
DG2634	DG1768, *Ty1–4253his3-AI*, +37 Ty1		(31)
YEM515	DG2634, *spt3-ΔKanMX4*		E. Matsuda
GRF167	*MATα*, *ura3-167, his3-Δ200*		(15)

DG789	GRF167, *spt3-101*		(50)
BY4742	*MAT*α, *his3-Δ1*, *leu2-Δ0*, *lys2-Δ0*, *ura3-Δ0*		(96)
DG2247	BY4742, *spt3-ΔKanMX4*		(31)
MAC103	BY4742, *xrn1/kem1-*Δ*KanMX4*		(7)
DG3582	DG1768, *trp1*		This study
DG3753	DG3582	p*GAL*/Cen-*URA3*, pBDG1534 (pGTy1*his3-AI*/Cen-*TRP1*)	This study
DG3751	DG3582	pBDG1354 (p*GAL*:1042-5889/Cen-*URA3*), pBDG1534	This study

Strain	Genotype	Plasmid	Source
DG3739	DG3582	pGAL-Yes2/2μ-*URA3*, pBDG1534	This study
DG3774	DG3582	pBDG1565 (p*GAL*-Yes2:1038-1616), pBDG1534	This study
DG3784	DG3582	pBDG1565, empty/Cen-*TRP1*	This study
JM367	DG3582	pBDG1568 (p*GAL*-Yes2:1038-V5-1616), pBDG1534	This study
DG3791	DG3582	pBDG1571 (p*GAL*-Yes2:1038-1496), pBDG1534	This study

JM399	DG3582	pBJM90 (p*GAL*-Yes2:1038-1616^{Gag*PR}), pBDG1534	This study
DG3808	BY4742	pEG(KT): p*GAL*:GST/2μ-*URA3*	This study
DG3809	BY4742	pBDG1576 (p*GAL*:GST-1038-1496)	This study
DG3810	DG3582	pBDG1576	This study

Table 2-2 Ty1*his3-AI* mobility

	Strain	Relevant genotype	Ty1*his3-AI* mobility x 10^{-6} (SD)	Fold decrease
A	DG2254	Ty1-less, Ty1*his3-AI*, p*GAL*/2μ	220 (69)	1
	DG2374	pGPOLΔ	7 (1.8)	31
	YAS73	pGPOLΔ-T399C (Gag: Ser36Pro)	24 (5.7)	9.2
	YAS69	pGPOLΔ-T1108C (Gag: Leu272Pro)	100 (27)	2.2
	YAS71	pGPOLΔ-A1123G (Gag: Tyr277Cys)	82 (11)	2.7
	YAS72	pGPOLΔ-A1296G (Gag: Thr335Ala)	42 (3.4)	5.2
	YAS70	pGPOLΔ-ΔA1456	31 (9.7)	7
B	DG2196	Ty1-less, Ty1*his3-AI*	120 (14)	1
	DG2511	+12 Ty1	3.6 (0.63)	33
	DG3798	+7 Ty1A1123G	570 (150)	4.8↑
C-1	DG2411	Ty1-less, Ty1*his3-AI*, empty/2μ	140 (30)	1
	DG2374	pGPOLΔ	2 (0.36)	74
	JM321	pGPOLΔ-ΔC1071	61 (39)	2.3
	JM320	pGPOLΔ- +A1303	58 (6.2)	2.4

C-2	JM321	pGPOLΔ-ΔC1071	65 (22)	2.1
	DG3856	pGPOLΔ-GCG1GCG2	137 (23)	1
D	DG3753	Ty1-less, pGTy1*his3-AI*/Cen, p*GAL*/Cen	21000 (2600)	1
	DG3751	pGAL:1042-5889	37 (8.8)	570
E	DG3739	Ty1-less, pGTy1*his3-AI*/Cen, pYES2/2μ	60000 (4700)	1
	DG3774	pYES2: 1038-1616	1.7 (0.57)	35,000
	JM367	pYES2: 1038-V5-1616	2 (0.65)	32,000

CHAPTER 3

CONCLUSIONS

Ty1 CNC was first demonstrated in a Ty1-less *S. paradoxus* strain which had lost its original complement of Ty1 elements by LTR-LTR recombination. In this strain, transposition of an introduced Ty1*his3-AI* element (1) decreased in response to repopulation by Ty1. (2) The dose-dependence of the response and mutant analyses suggested that Ty1 was producing a factor that limited retrotransposition. However, the factor that caused this copy number dependent Ty1 inhibition was unknown. An early report showed that a Ty1 antisense cryptic unstable transcript (CUT) silences Ty1 expression *in trans* by altering RNA Pol II occupancy and epigenetic marks on Ty1 chromatin (3). Ty1 antisense (AS) RNAs were also proposed to be responsible for loss of Ty1 mRNA when RNAi was reestablished in *S. cerevisiae* (4). Another study suggested that under severe adenine starvation, an increase in Ty1 mobility is coupled to transcriptional repression of Ty1 AS RNAs. The transcription factor Tye7 was shown to be responsible for repressing Ty1 AS RNAs resulting in increased Ty1 mobility (5). In 2009, Matsuda and Garfinkel provided evidence that Ty1 AS transcripts mediate CNC (6). Cells expressing Ty1 AS RNAs had low levels of transposition (CNC^+), AS RNAs associated with virus-like particles, and mutations in the 5' LTR that destabilized AS RNAs also rescued transposition (6). However, research presented here shows that Ty1 AS RNAs play little if any role in CNC (7). Random mutagenesis of the Ty1 CNC region revealed that mutations that abolish CNC frequently lie outside the AS RNA transcription units and do not affect AS RNA levels (Figure 2-1). All were missense

mutations leading to amino acid codon changes, suggesting a Ty1 protein was involved in CNC. We sought to discover the Ty1 RNA encoding this protein. Northern analysis of poly(A) RNA from wild type as well as mutant strains of *S. cerevisiae* (*spt3*Δ*, xrn1*Δ) and *S. paradoxus* (*spt3*Δ) revealed a subgenomic sense strand Ty1 RNA that was shorter that full-length Ty1 mRNA (Figure 2-3(A)). 5' RACE mapping revealed this subgenomic RNA initiated at Ty1 nucleotide 1000 (Figure 2-4), about 800 bases downstream of the Ty1 mRNA start site, in a region that encompassed the C-terminal half of *GAG* coding sequence. We named this transcript Ty1i RNA, for internally initiated RNA. Immunoblotting of whole cell extracts from the strains used in northern analysis, with a newly developed antibody (see Chapter 2, Materials and Methods) detected a novel 22 kD protein encoded by Ty1i RNA (Figure 2-3(B)). We also detected the processed form of p22 called p18, which is formed by a C-terminal cleavage by Ty1 PR at the same site used to process Gag-p49 to p45 (Figure 2-7). By introducing frameshift and nonsense mutations in this region of *GAG*, we showed that cells lose Ty1 CNC in the absence of p22 (Figure 2-6). Interestingly, Mutations in the R-U5 region or upstream of p22 coding sequence that destabilize Ty1 AS RNAs also destabilize Ty1i RNA (Figure 2-1). Finally, we demonstrated that p22 was sufficient for Ty1 CNC by ectopic expression of the restriction factor.

But how is Ty1i RNA transcription regulated? We showed that Ty1i RNA and p22 levels increase in mutants lacking the Spt3 subunit of the yeast transcription complex SAGA (7). Interestingly, Spt3 is required for Ty1 mRNA transcription (8) (Figure 2-3). The increase in Ty1i RNA in a *spt3*Δ mutant may be caused by increased accessibility of internal, cryptic promoters as observed in chromatin and transcription factor mutants

(9). Another interesting observation requiring further study is that we failed to detect a discrete Ty1i transcripts by northern or 5' RACE analyses in a Ty1-less *S. paradoxus* strain repopulated with 38 copies of Ty1 (Figure 2-2) even though p22/p18 was present. Instead of a discrete signal, we detected heterogeneous transcripts around 4.9 kb. However, the 5.7 kb Ty1 mRNA and therefore Gag was easily detectable in the same strain. Perhaps p22/p18 can be translated from Ty1 mRNA in *S. paradoxus* using an internal ribosome entry site (10), or by leaky scanning or by translation re-initiation (11-14). However, a more plausible explanation is that the heterogeneous population of Ty1i RNAs is translatable into p22 in the repopulated *S. paradoxus* strain as long as the transcripts contain a start codon for p22. *In vitro* translation studies showed that Ty1i RNA translation efficiency was related to the secondary structure at its 5' end and Gag may contribute to the Ty1i RNA stability and translation (15). Recent work showed that Ty1i RNA levels increase in several ribosomal protein mutants relative to Ty1 mRNA (16). More work will be required to understand how mutations in ribosomal proteins and resulting defects in translation lead to changes in Ty1i RNA levels and CNC. There is a lot more to learn about how Ty1i RNA transcription is regulated and what environmental factors, if any, control Ty1i RNA expression. Identifying promoter and enhancer sequences that are responsible for Ty1i RNA synthesis as well as identifying what transcription factors and accessory proteins bind to those sequences will shed more light on regulation of Ty1i transcription. Also, a systematic screen of published Ty1 restriction genes (17-21) for Ty1i RNA/p22 levels might lead to identification of a pathway or pathways that control Ty1i RNA/p22 expression and therefore Ty1 CNC.

To further study how p22 inhibited VLP assembly and function, we constructed an ectopic coexpression strain in the Ty1-less *S. paradoxus* that contained pGTy1*his3-AI* and pGp22 (7, 22). Following galactose induction, there was a dramatic 35,000-fold decrease in Ty1*his3-AI* transposition (Table 2-2). Sucrose gradient fractionation experiments showed that some p22/p18 cofractionated with VLPs. We also provided evidence that PR processing of Gag-p49, IN-p71 and RT-p63 was aberrant and resulted in novel proteins that do not correspond to known Gag-Pol-p199 processing intermediates (Figure 2-8). Transmission electron microscopy of these fractions revealed the presence of abnormally shaped VLPs that have an 'open' or 'incomplete' conformation (Figure 2-9). These results suggest that p22 association leads to defects in VLP protein processing, morphology and function. We also showed that p22 colocalized with retrosomes and disrupted their appearance. Since retrosomes contain VLP assembly intermediates, p22 may initially interact with Gag multimers or higher order structures of Gag. Although Gag and p22 co-immunoprecipitate the nature of the Gag multimer to which p22 binds remains to be determined. In order to understand how p22 association with VLPs alters their structure, it will be very informative to obtain high resolution 3-D reconstructions of VLPs bound by p22 using techniques like cryo-EM single particle analysis (23). This will help map p22-VLP interaction surfaces and reveal how p22 distorts VLP morphology.

Evolution of p22 mediated Ty1 CNC in *Saccharomyces*

RNA interference (RNAi) is an RNA based genome defense system important for maintaining genomic integrity against viruses and transposons. RNAi is brought about

by an endonuclease called Dicer, which cleaves dsRNA substrates into 20-30 nucleotide long RNAs that are bound by Argonaute proteins. This RNA induced silencing complex (RISC) binds to target mRNAs leading to loss of expression. It has been proposed that the last common ancestor to all eukaryotes had a form of RNAi (24). RNAi is present in a broad group of fungi including *Schizosaccharomyces pombe, Neurospora crassa, Cryptococcus neoformans* and *Mucor circineloides* (25-28). Interestingly, *S. cerevisiae* and *S. paradoxus* lost RNAi quite recently in evolutionary time (29). Perhaps p22 mediated control of Ty1 in RNAi-deficient yeast evolved as a result of this loss. In the absence of Dicer and Argonaute genes, Ty1i encoded p22 helps control transposition in these yeast species despite their lack of RNAi. But why did certain species of yeast lose RNAi in the first place? Exhaustive analyses of yeast that have lost or retained RNAi shed some light on what might have led to this loss (29, 30). The genomes of a large number of fungal species contain a dsRNA virus called killer, which encodes a toxin that kills other fungi that do not have the virus (30). All killer virus-containing yeasts have lost RNAi and closely related groups that did retain RNAi do not have killer, suggesting an inverse relationship between retaining RNAi and maintaining the killer virus. When RNAi was reconstituted in *S. cerevisiae* by introducing Dicer and Argonaute from closely related *S. castellii*, the killer ds RNA was processed into microRNAs resulting in loss of killer virus (30). Therefore, retaining RNAi leads to Dicer mediated cleavage of killer dsRNA, thereby rendering those cells susceptible to killing by strains that have intact killer. This provides a selective advantage to species that lost RNAi and retain the killer virus. However, losing RNAi could cause an increase in transposition that might lead to loss of genome integrity. *S. cerevisiae* and *S.*

paradoxus may have avoided these detrimental effects by evolving Ty1 CNC after losing RNAi.

References

1. Curcio MJ, Garfinkel DJ. 1991. Single-step selection for Ty1 element retrotransposition. Proc Natl Acad Sci U S A 88:936-940.
2. Garfinkel DJ, Nyswaner K, Wang J, Cho JY. 2003. Post-transcriptional cosuppression of Ty1 retrotransposition. Genetics 165:83-99.
3. Berretta J, Pinskaya M, Morillon A. 2008. A cryptic unstable transcript mediates transcriptional trans-silencing of the Ty1 retrotransposon in *S. cerevisiae*. Genes Dev 22:615-626.
4. Drinnenberg IA, Weinberg DE, Xie KT, Mower JP, Wolfe KH, Fink G, Bartel D. 2009. RNAi in budding yeast. Science 326:544-550.
5. Servant G, Pinson B, Tchalikian-Cosson A, Coulpier F, Lemoine S, Pennetier C, Bridier-Nahmias A, Todeschini AL, Fayol H, Daignan-Fornier B, Lesage P. 2012. Tye7 regulates yeast Ty1 retrotransposon sense and antisense transcription in response to adenylic nucleotides stress. Nucleic Acids Res 40:5271-5282.
6. Matsuda E, Garfinkel DJ. 2009. Posttranslational interference of Ty1 retrotransposition by antisense RNAs. Proc Natl Acad Sci U S A 106:15657-15662.
7. Saha A, Mitchell JA, Nishida Y, Hildreth JE, Ariberre JA, Gilbert WV, Garfinkel DJ. 2015. A trans-dominant form of Gag restricts Ty1 retrotransposition and mediates copy number control. J Virol 89:3922-3938.

8. Winston F, Durbin KJ, Fink GR. 1984. The *SPT3* gene is required for normal transcription of Ty elements in *S. cerevisiae*. Cell 39:675-682.

9. Cheung V, Chua G, Batada NN, Landry CR, Michnick SW, Hughes TR, Winston F. 2008. Chromatin- and transcription-related factors repress transcription from within coding regions throughout the *Saccharomyces cerevisiae* genome. PLoS Biol 6:e277.

10. Gilbert WV. 2010. Alternative ways to think about cellular internal ribosome entry. J Biol Chem 285:29033-29038.

11. Harrison PM, Kumar A, Lang N, Snyder M, Gerstein M. 2002. A question of size: the eukaryotic proteome and the problems in defining it. Nucleic Acids Res 30:1083-1090.

12. Kochetov AV. 2008. Alternative translation start sites and hidden coding potential of eukaryotic mRNAs. Bioessays 30:683-691.

13. Kochetov AV, Sarai A, Rogozin IB, Shumny VK, Kolchanov NA. 2005. The role of alternative translation start sites in the generation of human protein diversity. Mol Genet Genomics 273:491-496.

14. Wang XQ, Rothnagel JA. 2004. 5'-untranslated regions with multiple upstream AUG codons can support low-level translation via leaky scanning and reinitiation. Nucleic Acids Res 32:1382-1391.

15. Blaszczyk L, Biesiada M, Saha A, Garfinkel DJ, Purzycka KJ. 2017. Structure of Ty1 Internally Initiated RNA Influences Restriction Factor Expression. Viruses 9.

16. 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. Mob DNA 6:22.

17. Scholes DT, Banerjee M, Bowen B, Curcio MJ. 2001. Multiple regulators of Ty1 transposition in *Saccharomyces cerevisiae* have conserved roles in genome maintenance. Genetics 159:1449-1465.

18. Griffith JL, Coleman LE, Raymond AS, Goodson SG, Pittard WS, Tsui C, Devine SE. 2003. Functional genomics reveals relationships between the retrovirus-like Ty1 element and its host *Saccharomyces cerevisiae*. Genetics 164:867-879.

19. Nyswaner KM, Checkley MA, Yi M, Stephens RM, Garfinkel DJ. 2008. Chromatin-associated genes protect the yeast genome from Ty1 insertional mutagenesis. Genetics 178:197-214.

20. Dakshinamurthy A, Nyswaner KM, Farabaugh PJ, Garfinkel DJ. 2010. *BUD22* affects Ty1 retrotransposition and ribosome biogenesis in *Saccharomyces cerevisiae*. Genetics 185:1193-1205.

21. Risler JK, Kenny AE, Palumbo RJ, Gamache ER, Curcio MJ. 2012. Host co-factors of the retrovirus-like transposon Ty1. Mob DNA 3:12.

22. Nishida Y, Pachulska-Wieczorek K, Blaszczyk L, Saha A, Gumna J, Garfinkel DJ, Purzycka KJ. 2015. Ty1 retrovirus-like element Gag contains overlapping restriction factor and nucleic acid chaperone functions. Nucleic Acids Res 43:7414-7431.

23. Gong M, Zhu H, Zhou J, Yang C, Feng J, Huang X, Ji G, Xu H, Zhu P. 2014. Cryo-electron microscopy study of insect cell-expressed enterovirus 71 and

coxsackievirus a16 virus-like particles provides a structural basis for vaccine development. J Virol 88:6444-6452.

24. Shabalina SA, Koonin EV. 2008. Origins and evolution of eukaryotic RNA interference. Trends Ecol Evol 23:578-587.

25. Cogoni C, Irelan JT, Schumacher M, Schmidhauser TJ, Selker EU, Macino G. 1996. Transgene silencing of the al-1 gene in vegetative cells of Neurospora is mediated by a cytoplasmic effector and does not depend on DNA-DNA interactions or DNA methylation. EMBO J 15:3153-3163.

26. Wang X, Hsueh YP, Li W, Floyd A, Skalsky R, Heitman J. 2010. Sex-induced silencing defends the genome of Cryptococcus neoformans via RNAi. Genes Dev 24:2566-2582.

27. Nicolas FE, Torres-Martinez S, Ruiz-Vazquez RM. 2003. Two classes of small antisense RNAs in fungal RNA silencing triggered by non-integrative transgenes. EMBO J 22:3983-3991.

28. Nicolas FE, Torres-Martinez S, Ruiz-Vazquez RM. 2013. Loss and retention of RNA interference in fungi and parasites. PLoS Pathog 9:e1003089.

29. Drinnenberg IA, Weinberg DE, Xie KT, Mower JP, Wolfe KH, Fink GR, Bartel DP. 2009. RNAi in budding yeast. Science 326:544-550.

30. Drinnenberg IA, Fink GR, Bartel DP. 2011. Compatibility with killer explains the rise of RNAi-deficient fungi. Science 333:1592.

Figures

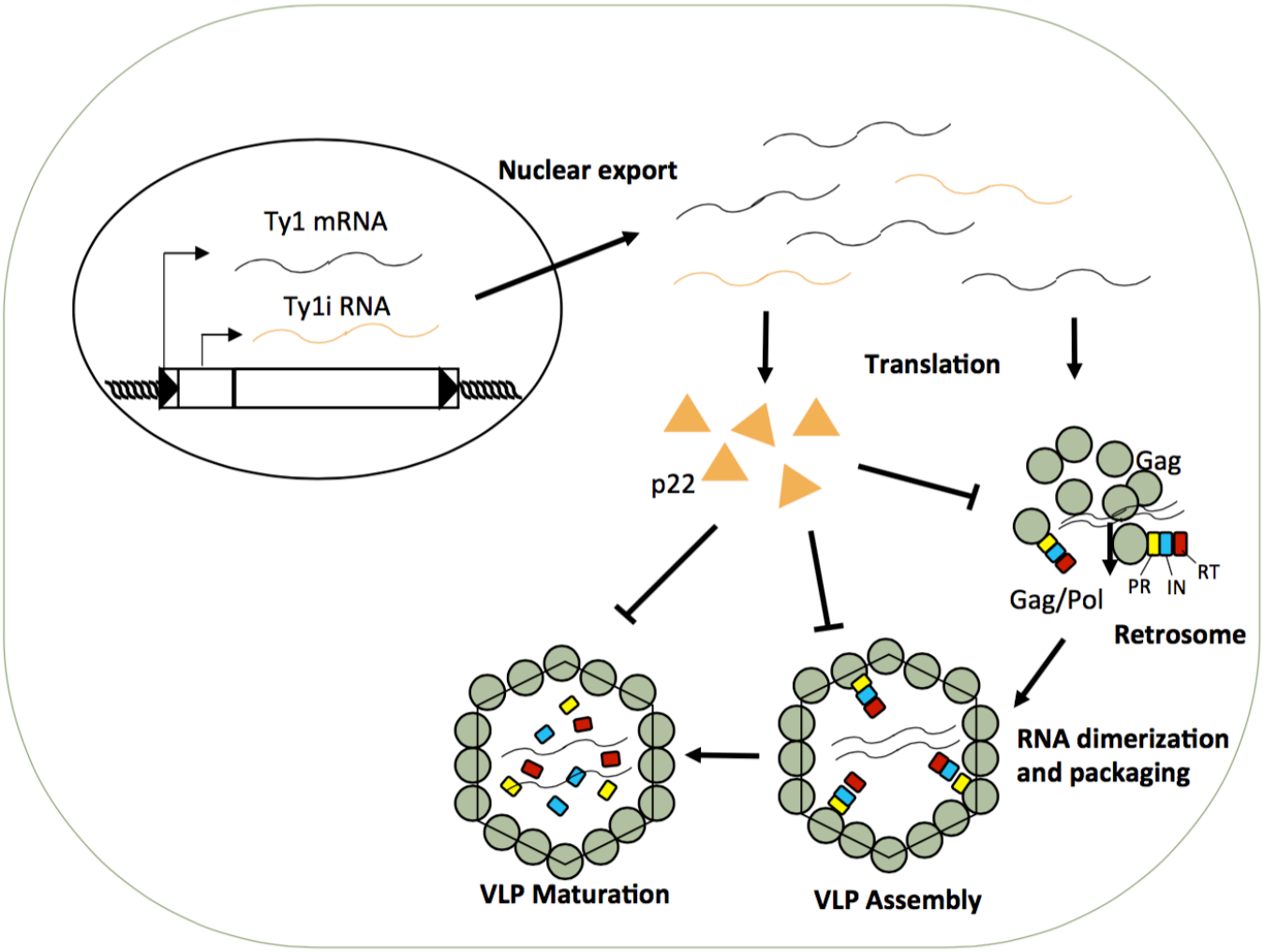

Figure 3-1. p22 affects multiple aspects of Ty1 replication.

APPENDICES

APPENDIX A

Purzycka KJ, Legiewicz M, Matsuda E, Eizentstat LD, Lusvarghi S, **Saha A**, Le Grice SF, Garfinkel DJ. 2013. Exploring Ty1 retrotransposon RNA structure within virus-like particles. Nucleic Acids Res 41:463-473.

Abstract

Ty1, a long terminal repeat retrotransposon of *Saccharomyces*, is structurally and functionally related to retroviruses. However, a differentiating aspect between these retroelements is the diversity of the replication strategies used by long terminal repeat retrotransposons. To understand the structural organization of cis-acting elements present on Ty1 genomic RNA from the GAG region that control reverse transcription, we applied chemoenzymatic probing to RNA/tRNA complexes assembled in vitro and to the RNA in virus-like particles. By comparing different RNA states, our analyses provide a comprehensive structure of the primer-binding site, a novel pseudoknot adjacent to the primer-binding sites, three regions containing palindromic sequences that may be involved in RNA dimerization or packaging and candidate protein interaction sites. In addition, we determined the impact of a novel form of transposon control based on Ty1 antisense transcripts that associate with virus-like particles. Our results support the idea that antisense RNAs inhibit retrotransposition by targeting Ty1 protein function rather than annealing with the RNA genome.

APPENDIX B

Nishida Y, Pachulska-Wieczorek K, Blaszczyk L, **Saha A**, Gumna J, Garfinkel DJ, Purzycka KJ. 2015. Ty1 retrovirus-like element Gag contains overlapping restriction factor and nucleic acid chaperone functions. Nucleic Acids Res 43:7414-7431.

Abstract

Ty1 Gag comprises the capsid of virus-like particles and provides nucleic acid chaperone (NAC) functions during retrotransposition in budding yeast. A subgenomic Ty1 mRNA encodes a truncated Gag protein (p22) that is cleaved by Ty1 protease to form p18. p22/p18 strongly inhibits transposition and can be considered an element-encoded restriction factor. Here, we show that only p22 and its short derivatives restrict Ty1 mobility whereas other regions of GAG inhibit mobility weakly if at all. Mutational analyses suggest that p22/p18 is synthesized from either of two closely spaced AUG codons. Interestingly, AUG1p18 and AUG2p18 proteins display different properties, even though both contain a region crucial for RNA binding and NAC activity. AUG1p18 shows highly reduced NAC activity but specific binding to Ty1 RNA, whereas AUG2p18 shows the converse behavior. p22/p18 affects RNA encapsidation and a mutant derivative defective for RNA binding inhibits the RNA chaperone activity of the C terminal region (CTR) of Gag-p45. Moreover, affinity pulldowns show that p18 and the CTR interact. These results support the idea that one aspect of Ty1 restriction involves inhibition of Gag-p45 NAC functions by p22/p18-Gag interactions.

APPENDIX C

Blaszczyk L, Biesiada M, **Saha A**, Garfinkel DJ, Purzycka KJ. 2017. Structure of Ty1 Internally Initiated RNA Influences Restriction Factor Expression. Viruses 9 (4):74.

Abstract

The long-terminal repeat retrotransposon Ty1 is the most abundant mobile genetic element in many *Saccharomyces cerevisiae* isolates. Ty1 retrotransposons contribute to the genetic diversity of host cells, but they can also act as an insertional mutagen and cause genetic instability. Interestingly, retrotransposition occurs at a low level despite a high level of Ty1 RNA, even though *S. cerevisiae* lacks the intrinsic defense mechanisms that other eukaryotes use to prevent transposon movement. p22 is a recently discovered Ty1 protein that inhibits retrotransposition in a dose-dependent manner. p22 is a truncated form of Gag encoded by internally initiated Ty1i RNA that contains two closely-spaced AUG codons. Mutations of either AUG codon compromise p22 translation. We found that both AUG codons were utilized and that translation efficiency depended on the Ty1i RNA structure. Structural features that stimulated p22 translation were context dependent and present only in Ty1i RNA. Destabilization of the 50 untranslated region (50 UTR) of Ty1i RNA decreased the p22 level, both in vitro and in vivo. Our data suggest that protein factors such as Gag could contribute to the stability and translational activity of Ty1i RNA through specific interactions with structural motifs in the RNA.

APPENDIX D (Review Article)

Garfinkel DJ, Tucker JM, **Saha A**, Nishida Y, Pachulska-Wieczorek K, Blaszczyk L, Purzycka KJ. 2016. A self-encoded capsid derivative restricts Ty1 retrotransposition in *Saccharomyces*. Curr Genet 62:321-329.

Abstract

Retrotransposons and retroviral insertions have molded the genomes of many eukaryotes. Since retroelements transpose via an RNA intermediate, the additive nature of the replication cycle can result in massive increases in copy number if left unchecked. Host organisms have countered with several defense systems, including domestication of retroelement genes that now act as restriction factors to minimize propagation. We discovered a novel truncated form of the *Saccharomyces* Ty1 retrotransposon capsid protein, dubbed p22 that inhibits virus-like particle (VLP) assembly and function. The p22 restriction factor expands the repertoire of defense proteins targeting the capsid and highlights a novel host–parasite strategy. Instead of inhibiting all transposition by domesticating the restriction gene as a distinct locus, Ty1 and budding yeast may have coevolved a relationship that allows high levels of transposition when Ty1 copy numbers are low and progressively less transposition as copy numbers rise. Here, we offer a perspective on p22 restriction, including its mode of expression, effect on VLP functions, interactions with its target, properties as a nucleic acid chaperone, similarities to other restriction factors, and future directions.

APPENDIX E (Review Article)

Pachulska-Wieczorek K, Blaszczyk L, Gumna J, Nishida Y, **Saha A**, Biesiada M, Garfinkel DJ, Purzycka KJ. 2016. Characterizing the functions of Ty1 Gag and the Gag-derived restriction factor p22/p18. Mob Genet Elements 6:e1154637.

Abstract

The long terminal repeat (LTR) and non-LTR retrotransposons comprise approximately half of the human genome, and we are only beginning to understand their influence on genome function and evolution. The LTR retrotransposon Ty1 is the most abundant mobile genetic element in the *S. cerevisiae* reference genome. Ty1 replicates via an RNA intermediate and shares several important structural and functional characteristics with retroviruses. However, unlike retroviruses Ty1 retrotransposition is not infectious. Retrotransposons integrations can cause mutations and genome instability. Despite the fact that *S. cerevisiae* lacks eukaryotic defense mechanisms such as RNAi, they maintain a relatively low copy number of the Ty1 retrotransposon in their genomes. A novel restriction factor derived from the C-terminal half of Gag (p22/p18) and encoded by internally initiated transcript inhibits retrotransposition in a dose-dependent manner. Therefore, Ty1 evolved a specific GAG organization and expression strategy to produce products both essential and antagonistic for retrotransposon movement. In this commentary we discuss our recent research aimed at defining steps of Ty1 replication influenced by p22/p18 with particular emphasis on the nucleic acid chaperone functions carried out by Gag and the restriction factor.