Ty3, a Position-specific Retrotransposon in Budding Yeast

SUZANNE SANDMEYER,¹,² KURT PATTERSON,¹ and VIRGINIA BILANCHONE¹

¹Department of Biological Chemistry; ²Department of Microbiology and Molecular Genetics, University of California, Irvine, CA

ABSTRACT Long terminal repeat (LTR) retrotransposons constitute significant fractions of many eukaryotic genomes. Two ancient families are Ty1/Copia (Pseudoviridae) and Ty3/Gypsy (Metaviridae). The Ty3/Gypsy family probably gave rise to retroviruses based on the domain order, similarity of sequences, and the envelopes encoded by some members. The Ty3 element of Saccharomyces cerevisiae is one of the most completely characterized elements at the molecular level. Ty3 is induced in mating cells by pheromone stimulation of the mitogen-activated protein kinase pathway as cells accumulate in G1. The two Ty3 open reading frames are translated into Gag3 and Gag3–Pol3 polyprotein precursors. In haploid mating cells Gag3 and Gag3–Pol3 are assembled together with Ty3 genomic RNA into immature virus-like particles in cellular foci containing RNA processing body proteins. Virus-like particle Gag3 is then processed by Ty3 protease into capsid, spacer, and nucleocapsid, and Gag3–Pol3 into those proteins and additionally, protease, reverse transcriptase, and integrase. After haploid cells mate and become diploid, genomic RNA is reverse transcribed into cDNA. Ty3 integration complexes interact with components of the RNA polymerase III transcription complex resulting in Ty3 integration precisely at the transcription start site. Ty3 activation during mating enables proliferation of Ty3 between genomes and has intriguing parallels with metazoan retrotransposon activation in germ cell lineages. Identification of nuclear pore, DNA replication, transcription, and repair host factors that affect retrotransposition has provided insights into how hosts and retrotransposons interact to balance genome stability and plasticity.

INTRODUCTION

Long terminal repeat (LTR) retrotransposons occur throughout eukaryotic phyla, but vary greatly in both the types of elements and the representation among species. The majority of some genomes are composed of these elements. The LTR retrotransposons are taxonomically divided into Pseudoviridae (or Ty1/Copia) and Metaviridae (or Ty3/Gypsy) elements based on genome organization and relatedness of proteins encoded (reviewed in references 1–3) (Gypsy Database 2.0, gydb.org). The Ty3/Gypsy family has shared ancestry with retroviruses and some members encode envelope enabling intercellular transmission. The eponymous founding elements of these LTR retrotransposon families, Ty1/Copia and Ty3/Gypsy, exist in Saccharomyces cerevisiae (Ty1 and Ty3) and in Drosophila (Copia and Gypsy).

This review is focused on Ty3, a representative of the Ty3/Gypsy retrotransposon group, which has been extensively characterized at the molecular level. Ty3 LTRs flank two overlapping open reading frames (ORFs), GAG3 and POL3, which encode structural and catalytic functions, respectively. Ty3 was discovered as a polymorphism associated with tRNA genes (tDNAs) and has been extensively studied as a model of targeted integration. A combination of approaches has implicated...
numerous host factors in control of Ty3 retrotransposition. Altogether, studies of Ty3 provide for a relatively comprehensive, albeit incomplete, picture of the ongoing remodeling of a eukaryotic genome by an LTR retrotransposon.

**TY3 DIVERSITY**

Ty3 is the sole representative of the Ty3/Gypsy family in *S. cerevisiae*. The reference *S. cerevisiae* strain BY4741 contains two closely related, full-length Ty3 elements one of which, YGRWTy3-1, is transpositionally active. In addition to the four LTRs associated with full-length elements, BY4741 contains 41 isolated LTRs (4–6) derived from recombination of the LTRs flanking full-length elements. These isolated LTRs were named saga elements (7, 8) before the discovery of full-length Ty3 (9). Ty1/Copia elements in *S. cerevisiae* are represented by Ty1, -2, -4, and -5. Ty1 (see Curcio, Lutz, and Lesage, Chapter 41) is the most abundant of these, with 32 full-length elements, two truncated elements with single LTRs, and 185 isolated LTRs (4–6).

Phylogenetic analysis of full-length elements and isolated LTRs has provided insight into the time of appearance and retrotransposition activity relative to host phylogeny based on occurrence of integrants shared among diverging host species and the divergence of flanking LTRs of individual full-length copies, as these are regenerated from the same template during retrotransposition (10). Comparison of Ty3 elements and LTRs in related *Saccharomyces* species showed that LTRs flanking full-length elements in the reference strain were identical and LTRs overall had >96% identity. In comparison, Ty1, Ty2, and Ty5 LTRs display 70 to 99% identity (5, 6). Although this suggests that Ty3 appeared recently relative to other elements, analysis of a large number of related *Saccharomyces* genomes identified ancient, shared Ty3 insertions pre-dating the appearance of *S. cerevisiae* (4). However, the presence of 24 highly similar, short-branch LTRs is consistent with recent retrotransposition of elements related to the full-length *S. cerevisiae* elements (4). Fifteen highly degenerate LTRs were identified in *S. cerevisiae* that were more similar to *Saccharomyces paradoxus* Ty3p than to *S. cerevisiae* Ty3 LTRs. Because common insertions preceding the divergence of *S. paradoxus* and *S. cerevisiae* were absent, these elements were concluded to have entered the *S. cerevisiae* genome via horizontal transmission from *S. paradoxus* or from some ancestral species shared between *S. paradoxus* and *S. cerevisiae* (4). Ty3 also has a greater proportion of solo LTRs relative to full-length elements than other Tys except Ty4. Although the reason for this is not clear, it is possible that this explains the absence of full-length copies of Ty3p from the modern *S. cerevisiae* genome (4).

**TY3 OVERVIEW**

Integrated, full-length copies of Ty3 are 5.4 kb in length, comprised of two LTRs of 340-bp each, flanking overlapping ORFs, GAG3 and POL3. Ty3 is transcribed into a 5.2-kb genomic (g)RNA that begins and ends in the LTRs (Figure 1). Translation of this RNA yields polyprotein precursors, a 34-kDa Gag3 and a 178-kDa Gag3–Pol3 product of frameshifting between GAG3 and POL3. Gag3 contains major structural protein domains CAPsid (CA), SPacer (SP), and NucleoCapsid (NC). Gag3–Pol3, in addition to those structural domains, contains catalytic domains PROtease (PR), Reverse Transcriptase (RT), and INtegrase (IN), and a spacer “J” between the PR and RT domains. Gag3 and Gag3–Pol3 assemble together with gRNA into immature 44 to 53 nm diameter virus-like particles (VLPs). Assembly of these occurs in cytoplasmic foci referred to as retrosomes (11) in which Ty3 proteins and RNA are concentrated with host assembly factors. Retrosomes form over the first several hours of Ty3 expression and become generally perinuclear. Within the VLP, precursor polyproteins undergo proteolytic processing by Ty3 PR into mature species. These populations of VLPs are much more diverse in size and morphology and range from 25 to 52 nm with electron dense cores. Subsequently Ty3 RT reverse transcribes gRNA into cDNA and remodeling of the VLP accompanies nuclear entry of a preintegration complex (PIC) of unknown composition. Integration into the transcription start site (TSS) of genes transcribed by RNA polymerase III (RNAP III) completes the cycle.

The reference *S. cerevisiae* strain BY4741 contains two full-length elements, YGRWTy3-1 and YILWTy3-1 (9, 12). YGRWTy3-1 is transpositionally active. YILWTy3-1 is transcribed and translated, but is inactivated by a frameshift mutation in the IN-coding region. The active element, YGRWTy3-1, is comprised of 340-bp LTRs flanking an internal coding domain of 4671 bp for a total length of 5351 bp. Ty3 null strains have been derived by deletion of endogenous elements (13).

**TY3 EXPRESSION**

**Mating control of Ty3 expression**

Ty3 transcription initiates in the upstream LTR and terminates downstream of that point in the downstream
LTR to yield the 5.2-kb RNA. This defines terminal RNA segments as U5 (unique to the 5′ end); R (repeated); and U3 (unique to the 3′ end) for a full-length element structure of U3-R-U5-inte

FIGURE 1 Ty3 retrotransposition. (A) Ty3 replication cycle. Pheromone binding to MATa or MATα pheromone receptors activates G protein-coupled mating signal transduction via mitogen-activated protein (MAP) kinase kinase kinase Ste7, MAP kinase kinase Ste11 and MAP kinase Fus3 (rose). Scaffold protein Ste5 (blue) supports specificity of their interaction preventing crosstalk with the filamentous growth pathway. Fus3 phosphorylates Dig1 and Dig2 negative regulators (gold) of Ste12 (dark blue), which then dissociate allowing Ste12 activation of RNA Pol II transcription of Ty3. Ty3 poly(A) RNA (maroon) is exported and translated into Gag3 and Gag3–Pol3 (tan), which then associate, together with the gRNA and RNA processing body (PB) factors, forming retrosomes within which Ty3 VLPs assemble. These foci become perinuclear over time. Assembly activates protease (PR) processing and maturation of the virus-like particles (VLPs). After cells mate (not shown) reverse transcription of the gRNA into cDNA occurs. Uncoating (dissociation of Gag3) presumably accompanies nuclear entry of the PIC. (A, B) Ty3 cDNA associates with RNAP III transcription initiation complexes composed of TFIIIB (yellow) and TFIIIC (green). In vitro TATA binding protein and Brf1 constitute the minimum target, but in vivo evidence suggests that TFIIIC can also be present. doi:10.1128/microbiolspec.MDNA3-0057-2014.f1

AMsience.org/MicrobiolSpectrum
FIGURE 2 Ty3 DNA, RNA, and protein. (A) DNA and RNA. The 5.4-kbp Ty3 element is transcribed into a 5.2-kb RNA. The major 5′ end of Ty3 RNA maps to 118 nucleotides (nt) inside the upstream long terminal repeat (LTR), and the 3′ end to heterogeneous positions between 243 and 273 nt inside the downstream LTR, as well as beyond the downstream LTR (14). The overlap resulting from termination downstream of the position of initiation results in a sequence that is repeated ("R") and defines 5′ U5 and 3′ U3 sequences. The initiator AUG of the GAG3 open reading frame (ORF) occurs at nucleotides 76 to 78 inside the Ty3 internal domain for a total 5′ untranslated region (UTR) of 193 nt; POL3 extends into the downstream LTR to define a 3′ UTR of ~227 nt (12, 14). Candidate upstream TATA and downstream polyadenylation sites are identified, but not experimentally verified. The RNA contains a bipartite primer binding site (PBS), which anneals to initiator tRNAMet in the upstream untranslated region and the downstream LTR (gold boxes). The GAG3 and POL3 ORFs encode Gag3 and Gag3–Pol3. (B) Protein. Gag3 is 290 amino acids (aa) and contains domains that mature via Ty3 PR processing into 207-aa capsid (CA), 27-aa spacer (SP), and 57-aa nucleocapsid (NC). Gag3–Pol3 contains those and additionally, protease (PR); reverse transcriptase (RT) starting at amino acids 536; and two forms of integrase (IN) domains (starting at amino acids 1012 and 1038) produced via a programmed frameshift. The POL3 ORF terminates within the downstream LTR so that the polypurine tract (PPT) plus strand primer is actually within the IN-coding region. (C) Reverse transcription of Ty3 genomic RNA. The tRNA primes synthesis of a minus-strand strong stop containing U5 and R segments, which then transfers to the 3′ end and primes extension of the minus strand. The plus-strand strong stop intermediate is initiated with cleavage by RNaseH at the downstream end of the PPT just outside the downstream LTR and is extended through U3, R, and U5 and likely copied into the 3′ end of the RNA then transferred to the 5′ end of the DNA and extended to form the plus strand of the cDNA. Although as described in the text minus- and plus-strand strong-stop intermediates have been identified, the overall flow described is based on the retrovirus model. An additional possibility (not shown) is that the 5′ and 3′ ends are transiently joined in a lariat RNA (see text). Bottom, the full-length cDNA has two extra base pairs on each end derived from a 2-nt offset between the priming sites and the LTR ends of the integrated element. Integrate (IN) processes 2 nt from each 3′ end and mediates the nucleophilic attacks of the resulting hydroxyls at 5-bp staggered positions flanking the RNA polymerase III (RNAPIII) transcription initiation sites. The integration site is repaired, resulting in 5-bp direct repeats flanking the ends of the Ty3 element.


Figure 2 continues on next page
characteristic of LTR retrotransposons and to retroviruses. The Ty3 GAG3 ORF lies downstream of R-U5, but the POL3 ORF overlaps U3.

*Saccharomyces cerevisiae* exists in haploid MATa and MATα mating types that secrete mating-type-specific pheromones that bind to receptors triggering mitogen-activated protein (MAP) kinase signaling. Ty3 transcription and retrotransposition are subject to mating type control (Figure 1A). In haploid cells, Ty3 is expressed at low levels. However, the full-length Ty3 elements (13, 14) and some LTRs are induced up to 80-fold in mating cells (15). This activates the MAP

**FIGURE 2 (continued)**

ASMscience.org/MicrobiolSpectrum
kinase Fus3 to relieve Dig1,2 repression of positive trans-
cription factor Ste12 bound to pheromone response el-
ements. These pheromone response elements occur in
 promoters of genes mediating physiological changes in
preparation for mating including agglutinin production,
enhanced vesicular trafficking, polarized cell growth
(shmooing), and arrest in G1 of the cell cycle (reviewed
in reference 16). Deletion analysis identified two adja-
cent PRE sequences within the Ty3 U3.

Pheromone induction of Ty3 transcription in G1-
 arrested cells is accompanied by translation of Ty3 pro-
teins, and formation and maturation of VLPs. However,
reverse transcription is delayed until cells resume cycling
(17). Experiments in which transposing elements were
activated in one mating type, and integrations of those
elements were detected into plasmids carried in the op-
posite mating type demonstrated that this control of
cDNA production promotes invasion of Ty3 into new
host genomes (18). Once cells have mated, Ty3 is re-
pressed by the action of the MAT a1/a2 heterodimer
formed from proteins expressed from the MAT loci of
the two haploid mating cell types. Deletion analysis
identified MAT a1/a2 repressor binding sites in the U3
region (13).

Ty3 is also regulated in response to environmental
conditions. Similar to Ty1, it is induced in bas1Δ cells
during adenylate nucleotide deprivation. BAS1 encodes a
Myb-related factor that regulates transcription of genes
in purine and histidine biosynthetic pathways and mei-
otic recombination at specific genes (19-21). Although
the basis of induction has not been explored for Ty3, Ty1
induction is inversely correlated with anti-sense tran-
scription and depends on both the absence of Bas1 and
the presence of the bHLH transcription factor Tye7 (22).

Interaction of Ty3 RNA polymerase II and
flanking tDNA RNA polymerase III promoters

The unique Ty3 specificity for integration at the RNAP
III TSS provides a context in which to examine inter-
actions of RNAP II and RNAP III promoters. Investi-
gation of the interactions between the SUP2 tDNA
RNAP III promoter and the adjacent Ty3 LTR RNAP II
promoter showed that mutations that inactivated the
tDNA promoter greatly enhanced pheromone induc-
tion of the LTR promoter (23). The interference by
RNAP III promoters with RNAP II repression was fur-
ther defined and shown to be generalizable (24) and
dependent upon localization of tDNA and flanking
sequences to the nucleolus (25, 26). Hence, similar
to expression of chromodomain retrotransposons that

TY3 RETROTRANSPOSITION ASSAYS

Because retrotransposition is a relatively rare event,
occurring in less than 1% of induced cell cultures, the
sensitivity of assays used in genetic screens affects the
spectrum of Ty or host mutants recovered. Genetic
screens require patching naive cells, replica plating
patches onto one or more types of selective medium and
screening for papillation frequency. Given the rarity of

Ty1/Copia and Ty3/Gypsy coexistence
in budding yeast

Ty1 and Ty3 share a subset of host factors and are both
induced by MAP kinase signaling in response to envi-
ronmental stimuli. As described above, Ty3 is induced
in haploid cells by pheromone signaling through the mat-
ing MAP kinase pathway, which activates transcription
factor Ste12 (13, 18). Ty1 is induced under MAP kinase
activation of Ste12/Tec1 in diploid cells, in response to
nutritional deprivation (32). Despite common upstream
components of these pathways, cross talk between them
is controlled by mating MAP kinase Fus3-mediated
degradation of Tec1 in haploid cells and mating type
suppression of Ty3 in diploid cells (33). Ty1 constitutes
a high fraction of RNA in haploid cells (34), but Ty1
proteins are actively degraded during the mating pher-
romone response (35-37). Hence, despite overlapping
induction pathways and shared host factors, Ty1 and
Ty3 occupy discrete retrotransposition niches.
events, identification of mutants deficient in transposition becomes nontrivial. Identification of the step in Ty3 retrotransposition affected by mutations is further complicated because expression at the RNA, protein, and even cDNA level is relatively robust, but cells in which transposition occurs are infrequent. One possible explanation of this apparent disparity is that Ty3 VLPs assemble in cytoplasmic clusters. As clusters increase in size, the ability of individual VLPs to access nuclear pores may become limiting.

Ty3 can be induced under native or synthetic regulation. Induction under the native pheromone promoter allows evaluation of host factors in the physiologic context of retrotransposition, but transcription increases as cells accumulate in G1, and so is relatively asynchronous. Alternatively Ty3 has been engineered to be synchronously inducible in the presence of galactose by substitution of the GAL1–10 upstream activating sequence for the upstream Ty3 U3 region (9, 13, 14). Galactose induces a synchronous response, but has the disadvantage that galactose metabolism is under catabolite repression, so that rapid induction requires pregrowth in a nonglucose carbon source in which the growth rate slows.

Ty3 targeting offers the possibility of monitoring retrotransposition by selecting for insertions into target traps. One such target is composed of a pair of divergent tDNAs, one of which is a suppressor tDNA (18). Expression from the target suppressor is blocked by mutual interference of the tDNA transcription complexes so that cells containing a suppressible allele in a nutritional background, retrotransposed cells can be selected on medium lacking histidine and containing 5-fluoro-uracil, and required for polyprotein processing (43). Based on the retrovirus precedent, it is likely that dimerization of Ty3 Retrotransposition

Ty3 can be induced under native or synthetic regulation. Induction under the native pheromone promoter allows evaluation of host factors in the physiologic context of retrotransposition, but transcription increases as cells accumulate in G1, and so is relatively asynchronous. Alternatively Ty3 has been engineered to be synchronously inducible in the presence of galactose by substitution of the GAL1–10 upstream activating sequence for the upstream Ty3 U3 region (9, 13, 14). Galactose induces a synchronous response, but has the disadvantage that galactose metabolism is under catabolite repression, so that rapid induction requires pregrowth in a nonglucose carbon source in which the growth rate slows.

Ty3 targeting offers the possibility of monitoring retrotransposition by selecting for insertions into target traps. One such target is composed of a pair of divergent tDNAs, one of which is a suppressor tDNA (18). Expression from the target suppressor is blocked by mutual interference of the tDNA transcription complexes so that cells containing a suppressible allele in a nutritional background, retrotransposed cells can be selected on medium lacking histidine and containing 5-fluoro-uracil, and required for polyprotein processing (43). Based on the retrovirus precedent, it is likely that dimerization of Ty3 Retrotransposition

Ty3 can be induced under native or synthetic regulation. Induction under the native pheromone promoter allows evaluation of host factors in the physiologic context of retrotransposition, but transcription increases as cells accumulate in G1, and so is relatively asynchronous. Alternatively Ty3 has been engineered to be synchronously inducible in the presence of galactose by substitution of the GAL1–10 upstream activating sequence for the upstream Ty3 U3 region (9, 13, 14). Galactose induces a synchronous response, but has the disadvantage that galactose metabolism is under catabolite repression, so that rapid induction requires pregrowth in a nonglucose carbon source in which the growth rate slows.

Ty3 targeting offers the possibility of monitoring retrotransposition by selecting for insertions into target traps. One such target is composed of a pair of divergent tDNAs, one of which is a suppressor tDNA (18). Expression from the target suppressor is blocked by mutual interference of the tDNA transcription complexes so that cells containing a suppressible allele in a nutritional background, retrotransposed cells can be selected on medium lacking histidine and containing 5-fluoro-uracil, and required for polyprotein processing (43). Based on the retrovirus precedent, it is likely that dimerization of Ty3 Retrotransposition

Ty3 can be induced under native or synthetic regulation. Induction under the native pheromone promoter allows evaluation of host factors in the physiologic context of retrotransposition, but transcription increases as cells accumulate in G1, and so is relatively asynchronous. Alternatively Ty3 has been engineered to be synchronously inducible in the presence of galactose by substitution of the GAL1–10 upstream activating sequence for the upstream Ty3 U3 region (9, 13, 14). Galactose induces a synchronous response, but has the disadvantage that galactose metabolism is under catabolite repression, so that rapid induction requires pregrowth in a nonglucose carbon source in which the growth rate slows.

Ty3 targeting offers the possibility of monitoring retrotransposition by selecting for insertions into target traps. One such target is composed of a pair of divergent tDNAs, one of which is a suppressor tDNA (18). Expression from the target suppressor is blocked by mutual interference of the tDNA transcription complexes so that cells containing a suppressible allele in a nutritional background, retrotransposed cells can be selected on medium lacking histidine and containing 5-fluoro-uracil, and required for polyprotein processing (43). Based on the retrovirus precedent, it is likely that dimerization of Ty3 Retrotransposition

Ty3 can be induced under native or synthetic regulation. Induction under the native pheromone promoter allows evaluation of host factors in the physiologic context of retrotransposition, but transcription increases as cells accumulate in G1, and so is relatively asynchronous. Alternatively Ty3 has been engineered to be synchronously inducible in the presence of galactose by substitution of the GAL1–10 upstream activating sequence for the upstream Ty3 U3 region (9, 13, 14). Galactose induces a synchronous response, but has the disadvantage that galactose metabolism is under catabolite repression, so that rapid induction requires pregrowth in a nonglucose carbon source in which the growth rate slows.

Ty3 targeting offers the possibility of monitoring retrotransposition by selecting for insertions into target traps. One such target is composed of a pair of divergent tDNAs, one of which is a suppressor tDNA (18). Expression from the target suppressor is blocked by mutual interference of the tDNA transcription complexes so that cells containing a suppressible allele in a nutritional background, retrotransposed cells can be selected on medium lacking histidine and containing 5-fluoro-uracil, and required for polyprotein processing (43). Based on the retrovirus precedent, it is likely that dimerization of
PR within the VLP promotes its activation. In addition to PR itself, Gag3–Pol3 processing by PR generates 55-kDa Ty3 RT and 61-kDa and 58-kDa IN proteins (46). The significance of the difference in N-terminal processing between these IN forms is unknown. However, the 61-kDa IN species expressed as a recombinant protein is capable of in vitro strand-transfer activity (47). Within cells expressing wild-type (WT) Ty3, VLPs are heterogeneous, representing a mixture of mature and immature forms (43).

In addition to catalytic domains with known functions encoded in \( POL3 \), a Junction (J) domain inferred to be 10 kDa is encoded between PR- and RT-coding regions (43, 48). This includes a 26-aa motif, which occurs twice in YGRWTy3-1 and three times in YILWTy3-1. Deletion of J within boundaries chosen to reconstruct the processing site decreased Ty3 cDNA levels dramatically, but did not abrogate retrotransposition (48). Chimeras constructed between the YGRWTy3-1 and YILWTy3-1 elements showed that the J domain does not account for the inactivity of the YILWTy3-1 (12).

Fluorescence microscopy provides a powerful tool for examining the appearance and subcellular localization of retrotransposon proteins and RNA. In cells undergoing pheromone induction, a monomeric cherry-tagged version of the Gag3 precursor polyprotein appears within about 10 min as a diffuse cytoplasmic signal and a green fluorescent protein (GFP) tagged version of the full-length Gag3–Pol3 precursor polyprotein appears within about 30 min (our unpublished results) (49). Immunoblotting shows that processed products, likely reflecting VLP formation, become more abundant than Gag3 between 2 to 4 hours (our unpublished results), and clusters of VLPs can also be detected during this time by transmission electron microscopy (our unpublished results). Anti-GFP immuno-electron microscopy shows that Ty3 proteins tagged with GFP localize to these clusters (49). Southern blotting detects cDNA within 6 hours.

**Frameshifting**

Greater amounts of structural than catalytic proteins are required to assemble VLPs (28). The ratio of these components is controlled by a programmed plus-one frameshift near the downstream end of \( GAG3 \). The relative abundance of Gag3 and the presence of this domain in Gag3–Pol3 not only drives assembly, but...
insures inclusion of catalytic proteins (28, 45, 50). The key frameshifting event is mediated by a 7-nucleotide (nt) sequence (GCG AGU U) read as Ala-Ser within the GAG3 frame. Within this sequence, the Ser codon (AGU) is decoded by a rare tRNA, stimulating plus-one frameshifting into the GCG GUU frame. This sequence is translated as Ala-Val, resulting in formation of full-length Gag3–Pol3 (42). Shifting is further stimulated by a downstream enhancer structural element of 12 nt. The Ty3 recoding shift is essentially executed by the incoming amino-acyl tRNA. Hence, this programmed frameshifting differs from both the retrovirus “slippery” minus-one frameshift and the plus-one Ty1 frameshift, which depend upon shifting of ribosome-bound tRNAs. The inefficiency of this process causes the Gag3–Pol3 precursor to be produced at 5% to 10% of the level of structural proteins (28, 42). Given this mechanism, it is not surprising that growth conditions affect the frequency of frameshifting and modulate retrotransposition. For example, frameshifting increases in cells grown in galactose-containing medium (51).

**Virus-like particle structure**

Immature VLPs contain Gag3, Gag3–Pol3, dimeric Ty3 gRNA (52), tRNAs including Ty3 primer initiator tRNA\textsuperscript{Met} (tRNA\textsuperscript{Met}) (53) and probably other host RNAs and proteins (52, 54). Gag3 is the major structural protein of immature Ty3 VLPs, and its expression in *Escherichia coli* is sufficient to achieve formation of aberrant particles (55). Formation of VLPs correlates strongly with ability of mutants to undergo proteolytic processing and reverse transcription. The simplest interpretation is that the VLP shell not only triggers processing to activate protein domain functions, but also selectively sequesters, concentrates and protects enzymes and substrates. Ty3 PR and RT active site mutants, and WT VLP populations have been characterized by transmission electron microscopy and atomic force microscopy (44, 56–58). Immature VLPs generated from Ty3 PR active site mutants are relatively uniform, and appear roughly spherical with diameters measured by atomic force microscopy of 44 to 53 nm. Striking physical features of the VLPs, particularly evident in atomic force microscopy preparations of immature forms, are knob-like capsomeres with six-fold symmetry interspersed with knobs of five-fold symmetry suggesting a pseudo-icosahedral structure. Transmission electron microscopy of immature VLPs shows an outer layer that is presumably the Gag3 and associated RNA. Populations of VLPs in cells expressing WT Ty3 are more heterogeneous, with diameters ranging from 25 to 52 nm. Many of these have electron-dense centers, presumably reflecting NC and complexed RNA released from the shell. Smaller particles could be cores that have generated cDNA and have remodeled in some way.

Ty3 VLPs are extremely stable, and Gag3 spontaneously aggregates or forms particles (unpublished data). Ty1 particles have also been isolated and studied extensively, and must be similarly stable (59). In contrast, retroviral cores undergo maturation coincident with budding and so are enveloped in the mature form. These core particles can be stripped of envelope and readily dissociated (60). VLP stability may be a general feature of retrotransposons necessitated by intracellular maturation in the absence of envelope.

**Capsid**

The 290-aa Gag3 is processed into 207-aa CA (processed to remove the N-terminal Met), 26-aa SP and 57-aa NC species (43, 44). A CA-SP species is also prominent among Gag3 products of processing (56). CA is acetylated on the N-terminal Ser (unpublished data). Although retroviral Gag CA domains are not generally conserved in sequence, they are remarkably similar in structure. They share N-terminal domain (NTD) and C-terminal domain (CTD) bundles of alpha helices (61, 62) with an intervening linker. Modeling suggests that, similar to retroviral CA domains, the CA domain of Ty3 Gag3 forms alpha helical bundles with the NTD contributing the outer surface of the VLP (58). The N-terminal 20-aa of Ty3 Gag3 contains a late domain motif (YPXL). Late motifs are named after their function “late” in the retrovirus lifecycle when structural proteins interact with “Endosomal Sorting Complexes Required for Transport” (ESCRT) proteins to sequester maturing VLPs into plasma membrane budding structures (63). Intriguingly although Ty3 has not been observed budding into intracellular membranes, density equilibrium centrifugation suggests that it is transiently associated with membranes (our unpublished results), and ESCRT proteins are required for Ty3 retrotransposition (64). Although retroviral CA domains are mainly conserved at the structural level, a “Major Homology Region” (MHR) is found in Gag CA domains and also occurs in Ty3 CA from amino acids 86 to 100 (QGX2EXFX3X3H) (65). The Ty3 MHR occurs N-terminal in Ty3 Gag3 compared with the MHR in retroviral Gag, but mutations in the MHR cause similar defects in Ty3 and retroviral particle morphogenesis.

Analysis of Gag3 mutants generated by Ala scanning identified residues that affect assembly as evaluated by transmission electron microscopy visualization of VLPs,
RNA packaging, and concentration of Ty3 protein into foci. Assembly was particularly sensitive to disruption of the N-terminal 100 aa (58). Only a few individual mutations completely abrogated particle formation and in one case this was accompanied by appearance of extensive cytoplasmic filaments. Visualization with RFP and GFP fusions showed that these filaments contained Gag3 and at least two components of processing bodies (PB) (58). Two-hybrid analysis showed that interactions occur between the NTD and NTD, as well as between the NTD and CTD and that introduction of mutations with effects on assembly disrupted these interactions (45, 58).

**Nucleocapsid**

The 57-aa Ty3 NC domain of Gag3 is critical for capture of the gRNA. The NC NTD contains 17 basic aa and is followed by one copy of the conserved zinc knuckle CX2CX4HX4C present in one or two copies in retroviral NC domains (11, 50). Mutations in conserved residues within and around the zinc knuckle eliminated concentration of Ty3 RNA into retrosome foci, and significantly reduced gRNA incorporation into VLPs (packaging) and retrotransposition (57). Unexpectedly, substitution of the first and second, but not the third and fourth, residues of the CCHC motif caused accumulation of Gag3 in nuclear, partially condensed particles. Because substitution of any one of the CCHC residues should disrupt zinc binding and gRNA packaging, such disparate phenotypes suggest that this subdomain has more than one role in assembly. In addition, the finding that Gag3 accumulates efficiently in the nucleus of these mutants suggests that, similar to the scenario proposed for Rous sarcoma virus gRNA packaging (66, 67), Ty3 Gag3 unbound to gRNA might at some stage invade the nucleus to scavenge for gRNA. Potentially related scenarios for Gag perinuclear capture of newly exported gRNA are proposed for human immunodeficiency virus 1 (HIV-1) (68, 69) and Ty1 (70).

**Spacer**

A SP domain separates Gag3 CA and NC domains, and a similar situation exists in a subset of retrovirus Gag proteins, including those of HIV-1 and Rous sarcoma virus (71, 72). These SP domains are not conserved, but are thought to participate in assembly, possibly substituting for functions performed by the CTD of CA for retroviruses lacking SP (73) (reviewed in reference 11). Ty3 SP is distinguished from retrovirus SP domains in that it is longer and more acidic. As is the case for retroviruses, removal of SP is important for morphogenesis. Mutations at the Ty3 SP–NC junction that disrupt PR processing also block processing of CA–SP, but the converse is not the case. This situation suggests that Gag3 processing is ordered (56), similar to what occurs for HIV-1 (74). Because failure to produce free NC does not block gRNA packaging but blocks appearance of cDNA, these observations argue that the NC domain of immature Gag3 functions in packaging, but mature NC must chaperone nucleic acid interactions during reverse transcription. Disruption of CA–SP processing also blocks retrotransposition but subsequent to cDNA production. Surprisingly, deletion of the SP domain by reconstructing a hybrid processed CA–NC mutant, allowed a low level of retrotransposition, indicating that SP is not absolutely essential. In contrast, conversion of acidic SP residues to Ala disrupted VLP formation and eliminated retrotransposition. Together these observations are consistent with a model in which the Ty3 acidic SP domain could interact with the basic NTD of NC in a nonessential condensation function that promotes assembly. However, if SP is present, it is essential that it be released from CA to allow progression through some step subsequent to reverse transcription, such as VLP uncoating.

**Genomic RNA**

The Ty3 gRNA contains cis-acting sequences, enabling Gag3-mediated packaging (psi), priming of reverse transcription, and integration. Ty3 gRNA is dimeric in VLPs (52) and by analogy with retroviruses has roles in packaging and reverse transcription. The primer for Ty3 minus-strand reverse transcription is tRNA\textsuperscript{Met} (53), which anneals to complementary primer binding site (PBS) split between just downstream of U5 and within U3 (75). The primer for plus-strand reverse transcription is a PPT just upstream of U3. cis-acting sequences were collectively first delimited by determining that Ty3 nucleotides 429 to 4979 of the 5.2-kb RNA were dispensable for retrotransposition in the presence of a helper Ty3 to provide proteins (28). Subsequently, an assay in which packaged RNA was protected from exonuclease digestion further delimited the Ty3 packaging sequence (psi) (76). These experiments showed that an abbreviated RNA with the Ty3 5′ untranslated region (UTR) lacking both the upstream and downstream PBS and having a heterologous 3′ UTR is sufficient to mediate packaging of a short RNA encoding Gag3 (76) (Figure 2C). At least a subset of retroviral psi sequences, including that of murine leukemia virus, are upstream sequences containing hairpin structures. Within these hairpin loops, short inverted repeats seed
gRNA dimer formation, and trigger annealing of the RNA sequences within the opposing gRNA hairpins. The resulting gRNA dimers expose binding sites for the NC domain of Gag (77). Although such features are not identified for Ty3, the 5′ end of the gRNA does contain the region required in cis for protection by Gag3.

To better understand the subcellular localization of VLP assembly, Ty3 RNA was fluorescently tagged. RNA tagged with tandem hairpin binding sites for RNA MS2 bacteriophage CAPsid (MS2-CA) can be visualized if expressed in cells together with MS2-CA fused to a fluorescent protein reporter (78). Ty3 gRNA tagged with the MS2-CA binding site is active for retrotransposition and was visualized with MS2 CA fused to red fluorescent protein (RFP) or GFP (49, 57, 76). Examination of cells expressing Ty3 RNA visualized with MS2-CA-GFP showed that gRNA colocalizes in foci with Gag3 identified using anti-CA antibodies (49).

**Retrosome localization of virus-like particle assembly**

One of the striking features of Drosophila (79) and S. cerevisiae (54) expressing Ty3/Gypsy and Ty1/Copia retrotransposons (80) is occurrence of VLP clusters or retrosomes. Several observations support the interpretation that, at least in the case of Ty3, retrosomes are associated with assembly of Ty3 protein and RNA into VLPs (Figure 1A). First, Ty3 RNA, proteins, and VLPs accumulate in clusters and individual VLPs are rarely observed outside clusters; second, mutations that disrupt assembly either abrogate formation of clusters or cause less dense clusters of poorly formed particles; third, deletion of genes encoding host proteins required for assembly also disrupts cluster formation (49, 56–58) (our unpublished results). Studies of Ty3 retrosome formation have used antibodies against individual Ty3 proteins as well as GFP, RFP, and monomeric cherry (mCherry) as in vivo reporters for Ty3 RNA and protein (49). Ty3 Gag3 and Gag3–Pol3 have been fused alternatively to GFP, RFP or mCherry. Within 1 hour of Ty3 induction, fluorescently tagged Ty3 gRNA and proteins progress from diffuse fluorescence to small foci, which aggregate over the next several hours into larger foci containing VLPs and representing retrosomes (49) (our unpublished results). Quantitation of the ability of Ty3 RNAs lacking various Ty3 gRNA attributes (e.g., frameshift, POL3 ORF, PBS) or having UTRs from heterologous RNAs showed that in the presence of Gag3, RNAs localize to foci if they contain either the 5′ Ty3 UTR or a second long, poorly translated ORF (76). However, GAG3–POL3 RNA in which the Ty3 UTR are substituted with heterologous UTRs is not efficiently packaged, consistent with the 5′ UTR containing psi. Hence, localization to foci does not insure assembly of Ty3 RNA into particles.

**Retrosome association with RNA processing body components**

Screens for Ty1 and Ty3 host factors identified conserved components of RNA PB as factors required for WT levels of retrotransposition (64, 81). PB are RNA granules containing concentrations of RNA deadenylation-dependent degradation components including deadenylation enzymes, translation suppressors, decapping activators and enzymes, helicases, and exonuclease. PB were initially identified as sites of mRNA degradation (82), but this is unlikely to be their exclusive function. A subset of RNAs in PB is competent to resume polysome-associated translation (83). In addition, at least some poly(A)-dependent degradation occurs cotranslationally (84). Indeed further investigation indicated that PB and stress granule (SG) ribonucleoprotein particles, which reversibly sequester nontranslating RNAs during stress (85), share a subset of components (86). A revised interpretation is that these granules represent different ends of a spectrum of ribonucleoprotein complexes (87). PB proteins colocalizing with Ty3 retrosomes (49) include translational repressors Pat1 and Dhh1 (88–90); decapping activator Lsm1 (91); decapping protein Dcp2 (92); and 5′ exonuclease Xrn1 (87, 93–95).

Galactose induction of Ty3 expression causes formation and/or enlargement of PB, suggesting that PB proteins interact directly with Ty3 proteins and/or RNA. Examination of galactose-induced Ty3 protein and RNA fluorescent reporters in cells expressing PB-GFP reporters showed that Ty3 RNA and protein and PB components colocalize (11, 49). PB formation is also observed under conditions independent of galactose induction of Ty3. These conditions include nutritional stress (86), stationary phase (96), and mating (97). Ty3 protein and PB factors also colocalize in mating cells (49).

On the surface, it is surprising that retroelement assembly is associated with factors engaged in RNA degradation. However, as discussed above, not all PB targeted RNAs are degraded and the natural sequestration of RNAs in PB away from translation could provide a mechanism for Ty3 RNA to transition from translation into assembly. The POL3 ORF, downstream of the frameshift would have a lower density of ribosomes and could be subject to binding by PB translational repressors. Translational repression might enhance access of Gag3 to the 5′ packaging site, thereby protecting
the 5’ end from decapping and degradation. Greater interaction with PB or trafficking factors would then promote assembly by concentrating VLP components (76). In the absence of such enhanced concentration, Ty3 proteins and RNA might fail to assemble. PB factors are also required for Ty1 assembly (98, 99). However, Ty1 retrosomes are not completely coincident with PB or SG foci (98, 99) (see Curcio, Lutz and Lesage, Chapter 41).

**REVERSE TRANSCRIPTASE AND REVERSE TRANSCRIPTION**

Reverse transcription is the defining feature of the retroelement replication. Non-LTR retrotransposon RT containing solely the polymerase domain (Pol) likely preceded LTR retrotransposon RTs that contain Pol and RNaseH domains (3). The Pol catalytic activity is responsible for cDNA production from both plus-strand RNA and minus-strand cDNA templates. The RNaseH domain is responsible for degradation of template RNA in DNA–RNA heteroduplexes as well as the cleavage at PPT to create the plus-strand primer. Pol and RNaseH domains, each have conserved carboxylate triad active sites competent to chelate two Mg2+ or Mn2+ ions. However, these reactions differ in both their structural context and in the reactions they mediate: in Pol, a hydroxyl attacks a dNTP; whereas in RNaseH, H2O is activated to hydrolyze a polynucleotide chain (see Hughes, Chapter 46) (100, 101).

**Relationship of Ty3 RT to other RT enzymes**

Despite conservation of RT Pol and RNaseH domains among retroviral RT, there is diversity in quaternary structure: gamma-retroviruses, such as murine leukemia virus, have monomeric RT (Pol-RNaseH); alpharetroviruses, such as Rous sarcoma virus, use an RT/RT-IN heterodimer; and lenti-retroviruses such as HIV-1, use a Pol-RNaseH/Pol (p61/p56) to highlight a few variations (see Hughes, Chapter 46; Skalka, Chapter 48). Retroviruses likely evolved from progenitor LTR retrotransposons related to Ty3/Gypsy. Ty3 RT has about 25% identity in the Pol domain with monomeric murine leukemia virus RT (9). Ty3/Gypsy retrotransposon RNaseH domains occur immediately after the Pol domain. Retroviral RNaseH domains are not only separated from the Pol domain by a RNaseH fold, that lacks the key carboxylate triad, but are actually more closely related to RNaseH domains of cellular enzymes than RNaseH domains of LTR retrotransposons. An attractive explanation is that retroviruses acquired a cellular RNaseH domain, but retained the progenitor retroelement RNaseH domain as a catalytically inactive connector between Pol and RNaseH (102).

**Ty3 reverse transcriptase structure and active sites**

*In vivo*, Ty3 Gag3–Pol3 is processed into 55-kDa RT and 116-kDa RT-IN fusion species (46). *In vitro* recombinant 55-kDa RT displays an ability to extend synthetic primers on an RNA template comparable to that of recombinant retroviral RT (103, 104). Inspection shows Ty3 RT amino acids D151–X59–Y211–L212–D213–D214 consistent with the conserved retroviral Pol active-site motif, D-Xn-Y-L-D-D. The Ty3 protein sequence D358–E401–D426–H427–Y459–D469 is consistent with the RNaseH carboxylate triad in other members of the Ty3/Gypsy family, D-E-DH-Y-D (104). The conserved D-E-D triad is required for metal binding and activity, as are the His and Tyr residues (105). However, consistent with different origins of the RNaseH domain, the retroviral RNaseH motif, D-E-D-H-D, differs from the Ty3/Gypsy motif, in that Tyr is lacking and the His residue is moved away from the carboxylate triad (e.g. HIV-1 D443-E478-D498-D549).

The crystal structure of the Ty3 55-kDa RT bound to a 16-bp RNA/DNA heteroduplex PPT was the first retrotransposon RT crystal structure. It revealed a novel, substrate-induced dimer structure (106). The bound nucleic acid is intermediate between the A and B forms, a polynucleotide duplex with little structural deformation. In the Ty3 RT, subunits “A” and “B” have similar conserved subdomain folds, but these are arranged asymmetrically. Subunit A displays the characteristic Pol right-handed topology with palm subdomain containing the active site carboxylates, fingers stabilizing the RNA primer and thumb interacting with template DNA. In contrast, subunit B thumb and RNaseH subdomains are rotated by approximately 90°. This arrangement obstructs interaction between the subunit A RNaseH active site and scissile phosphate, but leaves an unobstructed path for subunit B RNaseH. Altogether the structure suggests that catalytic activities are shared between the two subunits. In support of that interpretation, RT proteins mutated in the half-dimer interface or in the RNaseH active site lack *in vitro* activity, but complement when combined *in vitro*. Hence, Ty3 and HIV-1 RT have similar conserved subdomains, but the overall structures show that subunits play different roles. In the case of Ty3 RT, substrate binding induces RT dimerization, whereas in the case of HIV-1, RT forms a heterodimer in the absence of substrate. Ty3 RT
enzymatic activity is split between Pol, contributed by subunit A, and RNaseH contributed by subunit B. In contrast, HIV-1 Pol and RNaseH activities are contained in p66 and the p51 subunit functions as a scaffold. Nonetheless, p51 of HIV RT and subunit B of Ty3 RT have similar subdomain arrangements, including the positions of the vestigial RNaseH p51 connector domain and active Ty3 subunit B RNaseH.

Mechanism of reverse transcription

The sites for initiation of reverse transcription of Ty3 minus-strand and plus-strand cDNA synthesis intermediates and full-length cDNA are identified. Overall, observations are consistent with the sequence of events proposed for replication of retrovirus gRNA (Figure 2C, see legend for detailed description) (see Skalka, Chapter 48; Hughes, Chapter 46). Ty3 minus-strand reverse transcription is primed from tRNA Meti (53) annealed to a bipartite PBS (75) and the plus-strand is primed from a downstream PPT. Overall, Ty3 RT performs an amazingly diverse set of activities: RNA- and DNA-dependent DNA synthesis, strand transfers, degradation of the RNA component of the RNA/DNA heteroduplex, and RNA endonucleolytic cleavages to create the plus-strand primer.

Ty3 cDNA primers

Ty3 minus-strand synthesis is primed by tRNA Meti. However, this tRNA rather than annealing to a 17-nt PBS as is the case for retrovirus tRNA primers, binds to a bipartite PBS, an 8-nt segment 2 nt downstream of U5 and two adjacent segments thousands of bases downstream within U3, for an additional 23 nt (75). Ty1 also uses a bipartite tRNA Meti PBS with a short upstream segment and two longer downstream segments. However, in the case of Ty1, the bipartite PBS is complete in the 5′ end (107). Ty retroelements differ from retroviruses, which typically have 17–18 bases of unpaired 5′ ends of the two bridging tDNAMeti molecules form an intermediate lariat with

duplex agree on mostly base-paired and A-form duplex lacking major structural distortions. However, nuclear magnetic resonance revealed sugar puckers at position +1 (G) of the RNA strand (where in vitro cleavage is between −1 and +1) from C3′ end to mixed C3′/C2′ end (111). Perturbations using locked nucleic acid analogs to constrain sugar rings and abasic tetrahydrofuran linkages, which lack base pairing, indicated that the the 5′ and 3′ ends of the PPT are more sensitive than internal positions to substitution with novel constituents. This suggests that PPT ends contribute to RNaseH recognition (112, 113). Underscoring these apparent distinctions from retroviral PPT, Ty3 RNaseH fails to recognize at least one example of a retrovirus PPT, that of HIV (AAAAGAAAAGGGGGGA) (104, 114).

RNA–RNA interactions supporting strand transfer

Packaging into core particles and participation in strong-stop cDNA intermediates strand transfers during reverse transcription are two requirements of LTR retrotransposon and retrovirus gRNAs. Dimer formation by gRNA and direct interactions between the 5′ and 3′ ends of the gRNAs, respectively, have been proposed to participate in these activities. It is noteworthy that tRNA Meti is shared as the minus-strand primer by distantly related elements Ty3 (53), Ty1 (53) and Ty5 (115). In the case of Ty3, which has a bipartite PBS so that annealed tRNA Meti bridges the 5′ and 3′ gRNA ends, the primer is proposed to facilitate strand transfer (75). An alternative proposal is that a transient covalent structure connects 5′ and 3′ ends, thereby supporting transfer of intermediates. In the case of Ty1 and Ty3, the lariat debranching enzyme, Dbr1, is required for WT levels of cDNA production (116). Based on this requirement, and consistent with in vitro observations, it was proposed that a 5′: 3′ lariat transiently links the Ty1 gRNA ends (117). Although the evidence for Ty1 is incomplete (118), experiments with HIV-1, which also requires Dbr1, appear supportive of an intermediate lariat in an alternative reverse transcription pathway (119).

Ty3 gRNA dimerization

Ty3 VLPs contain gRNA dimers (52). In the case of retroviruses, the dimerization initiation sequence maps within the psi packaging sequence, and this duplex structure is proposed to be a feature that identifies gRNA for packaging (reviewed in references 77 and 120). Gabus et al. (75) suggested that palindromic 5′ ends of the two bridging tDNAMeti molecules form an interface between the two Ty3 gRNAs. In addition to
linking 5′ and 3′ ends, this would mediate dimerization of the gRNA. In an in vitro system with mini-Ty3 element substrate, minus-strand strong-stop transfer and gRNA dimer formation were dependent upon the bipartite PBS and tRNA\textsuperscript{Met} primer (75, 103). However, the bipartite Ty3 PBS is not required for Ty3 gRNA packaging, suggesting either that dimer formation is not required as it is for retroviruses, or that tRNA\textsuperscript{Met} does not provide this interface (76). Ty1, which also uses tRNA\textsuperscript{Met} to prime reverse transcription, was similarly proposed to dimerize through the tRNA interface. However, recent analysis of Ty1 VLP RNA using Selective 2′ Hydroxyl Acylation Analyzed by Primer Extension (SHAPE) did not identify the proposed tRNA\textsuperscript{Met} interface, but rather a palindromic sequence within the 5′ end of the gRNA similar to the retroviral dimerization initiation sequence (121).

**NUCLEAR ENTRY**

Ty3 accesses integration targets via the nuclear pore complex

Ty3 VLPs assemble in the cytoplasm, but must access nuclear targets. In fungi the nuclear envelope does not break down during mitosis so that NPC are the gateway for multimolecular complexes. The yeast NPC is a conserved, eight-fold symmetric outer- and inner-ring and channel structure with a nuclear basket and a central meshwork of filamentous proteins rich in Phe-Gly (FG) repeats (122–124). NPC are dynamic and participate in multiple activities including gene gating, replication, and dsDNA break repair (29, 125). Because Ty3 VLPs are generally above the ~39-nm upper limit for complexes accommodated by the pore channel (126), it is likely that cDNA transit through the NPC is accompanied by significant remodeling also known as uncoating. This requirement is shared with fungal retroelements Ty1 (127, 128) and Tf1 (129). In nondividing animal cells, those retroviruses that are infectious, including HIV-1 and avian sarcoma leukemia virus, also access chromosomal DNA through the NPC (130, 131) (see Bushman and Craigie, Chapter 45). One unifying scenario is that the NPC plays a dynamic role in promoting entry by transforming both retrotransposon and VLPs and retrovirus cores into PIC.

**Nuclear pore complex components as host factors**

Consistent with the idea that nuclear access of retroelements involves cooperation between retroelements and host factors, multiple factors that facilitate (cofactors) or antagonize (restriction factors) retrotransposition have been identified. Genomewide screens for Ty3 host factors identified members of the FG filaments (Nup159 and Nup100); and outer ring Nup84 complex (Nup120, Nup133, and Nup84) as Ty3 restriction factors. Gtr1, a negative regulator of Ran; importin Kap120; inner ring adapters Nup59 and Nup157; and FG filament Nup116 (combined with Nup100 constitutes hNup98) were identified as cofactors (64, 132, 133).

Deletions of genes encoding components of the Nup84 complex increased retrotransposition frequency. The NPC outer ring Nup84-Nup120-Nup133 (hNup107-hNup160-hNup107, respectively) subcomplex participates in NPC organogenesis, and yeast mutants lacking one of these components cluster pores on the nuclear envelope (134, 135). In nup133Δ and nup120Δ mutants, the Ty3 retrosome co-localizes with the clustered NPC. Physical co-localization between Ty3 VLP clusters and NPC is consistent with cumulative association of fluorescent Ty3 protein clusters on the periphery of DAPI-stained nuclei and with localization observed in some transmission electron microscopy images (133). Although it is attractive to think that mutations in the Nup84 complex enhance access of VLPs to NPC across a greater surface of the nuclear envelope, it is equally possible that distortion caused by loss of NPC outer ring coat proteins reduces the stringency of the NPC size filter.

Nups identified as Ty3 cofactors might mediate the VLP docking required for nuclear entry, or play a role in uncoating. Members of the GLFG subset (Nup110 and Nup116) of FG Nups were identified as Ty3 retrotransposition cofactors (64, 132). A complicating factor in examination of the role of FG Nups is the fact that individual proteins are essential and the different classes of FG repeats are redundant (122, 136). The roles of FG repeat families were dissected using mutants in which combinations of deletions of specific FG repeats have been achieved so that strains are viable, but lack certain types of FG repeats (136). These experiments highlighted a role for GLFG repeat Nups, Nup100 and Nup116 (133). In vitro investigation showed not only that assembly-competent immature and mature VLPs interact with GLFG repeat Nups, but that GLFG repeats specifically interact with recombinant Gag3 (133).

An attractive model is that FG hydrophobic filaments interact with the VLP surface and promote dissociation of CA. Although specific contacts for Gag3 subdomains in uncoating have not been identified, mutations in CA
NTD and SP block retrotransposition subsequent to cDNA synthesis \(56, 58\). This phenotype is consistent with an uncoating defect.

**Redundancy of mechanisms of nuclear entry**

A complicating factor in understanding nuclear entry of all retroelements is the multiplicity of potential VLP mediators of nuclear import. In the case of HIV-1, MA, Vpr, IN and CA have been implicated in nuclear entry \(130, 137–139\). In addition to Ty3 Gag \(49\) implication in nuclear entry, a bipartite basic nuclear localization sequence occurs in IN (amino acids 401–436 of IN) \(140\). Mutations in this motif block integration at a post-cDNA synthesis stage and also block nuclear localization of IN expressed as an independent protein. Hence, there could be redundant routes of PIC nuclear entry, or Gag3 and IN could cooperate in sequential NPC association and translocation to insure efficient and directional nuclear entry.

**INTEGRASE AND INTEGRATION**

**Functions of integrase**

Mutations in Ty3 IN block retrotransposition at multiple steps: before or during cDNA synthesis, 3′ end processing, and strand transfer into the genome. The best understood of these steps is its role in targeted integration. The substrate for this activity is the ends of the Ty3 cDNA. As occurs for some retroelements including retroviruses, the Ty3 extra-chromosomal cDNA has two extra base pairs on each end and these are removed from the 3′ ends by IN before strand transfer \(46\). Integrated Ty3 has terminal 8-bp inverted repeats, the terminal dinucleotides of which are conserved among transposable elements \(5′-\mathrm{TGTTGTTAATACAAACA-3′}\). The cDNA is joined to host DNA by IN with strand transfers staggered by 5 bp, which are filled in on each strand so that the integrated Ty3 ends are flanked by 5-bp direct repeats of host sequence (Figure 2C).

Comparison of retrotransposons and retroviruses suggests that the former benefit from integration that avoids disruption of host functions by avoiding ORFs or targeting heterochromatin or both, whereas the latter benefit from targeting expressed regions. Yeast Ty1–4 and *Dictyostelium* TRE5 and TRE3 are targeted to tDNAs \(141–145\); Ty5 integrates into heterochromatin \(146\); Maggy chromodomain element targets H3K9Me \(147\), and Tf1 targets intergenic regions (see Levin, Chapter 43). Of the characterized retroelements, Ty3 has the most specific pattern of integration.

**Structural features of Ty3 integrase**

Long terminal repeat retrotransposon and retrovirus integrases have related three-domain structures. Alignment of the Ty3 61-kDa IN with other retroelement IN proteins shows conservation of a catalytic core domain flanked by less-conserved NTD and CTD domains. The first crystal structure of a retrovirus IN, that of prototypic foamy virus IN complexed oligonucleotides representing cDNA ends and target DNA provided insight into this class of proteins \(148, 149\). Overall retroviral IN proteins function as a tetramer composed of dimers bound to the ends of the cDNA with the inner subunits participating in strand transfer. The NTD of each inner subunit functions in *trans* with the catalytic core domain of the other subunit. Outer subunits of each dimer are thought to have similar folds, but whether they have distinct functions in target association is not yet known \(150\). The crystal structure showed that prototypic foamy virus IN has an N-terminal extended domain of about 100 residues. This is not highly conserved among retroviral IN proteins but is thought to make contacts with the viral DNA backbone in the case of prototypic foamy virus IN. The IN NTD is typically defined by a conserved HX_{3–5}H_{23–32}CX_{2}C motif near its N terminus. The position of the Ty3 zinc-binding motif, H93-X-H95-X101-X92-C131-X2-C134, indicates that Ty3 IN also has an N-terminal extended domain. The Ty3 catalytic core domain contains approximately a dozen residues highly conserved among Ty3/Gypsy and retrovirus IN proteins. These include Ty3 residues D164-X60-D225-X35-E261 representing the D-Xn-D-X35-E Mg^{2+} binding motif common to polynucleotidyl esterases, including DNA cut-and-paste transposable elements \(151, 152\). Mutations in this motif do not affect assembly or reverse transcription, but abrogate 3′-end processing and integration \(46\). Similarity between Ty3 IN and retroviral IN proteins is sufficient to support *in vitro* strand-transfer activity of a small subset of 27 chimeras of Ty3, prototypic foamy virus, and HIV IN proteins \(153\).

Just C-terminal to the catalytic core domain, is the Ty3 GPF/Y motif, G398-P399-F400, which is found in a subset of retroelement IN proteins \(154\). In Tf1 this motif is located in a subdomain that participates in multimerization \(155\). Overall, the CTD of IN is poorly conserved among retroviruses and LTR retrotransposons \(154\).

**Functions of nonconserved integrase domains**

Ala-scanning mutagenesis of the Ty3 IN NTD and CTD produced mutations in each domain that underwent
processing, but lacked cDNA (156). This pattern of IN mutant phenotypes is similar to that of class II mutants of HIV-1 IN (157). These form immature cores. Retrovirus immature cores disassemble upon proteolytic processing and reassemble as mature cores inside the envelope (158). However, class II mutants fail to correctly re-assemble and fail to make cDNA (159). The similarities of the mutant phenotypes suggest that Ty3 IN NTD and CTD have a role in post-processing rearrangement of domains required for cDNA synthesis.

The Ty3 IN CTD contains a bipartite basic nuclear localization sequences (R401-RVVKKINDAYELDL NSHKKBHRVINQFLKRFVYR-436) sufficient to mediate nuclear import of ectopic IN (140). As indicated above, the Ty3 nuclear localization sequence is required at some post-cDNA synthesis step, consistent with a role in nuclear import (140).

A large fraction of Ty3/Gypsy elements have IN proteins with chromodomains in the CTD and these have been taxonomically classified as Chromoviruses. The chromodomain is ~50 aa and mediates interactions with dimethyl and trimethyl H3K9, a signature of heterochromatin (142). This domain occurs in Ty3/Gypsy Maggy and Tf1 IN (147), but has been replaced by an alternative domain in Ty3 (154). A plausible explanation for the replacement is that the targeted dimethyl and trimethyl H3K9 modification is lacking in S. cerevisiae (160).

**Mechanism of Ty3 targeting**

Motivation for studies of Ty3 stemmed in large part from its remarkable integration specificity for the TSS of genomic RNAP III-transcribed genes including tDNAs (8, 141, 161), SNR6 (U6 RNA), and RDN5 genes (5S RNA) (162) (Figure 1B). The RNAPIII TSS occurs in a favored context rather than an absolute consensus, and is flanked by transcription factors. This suggested early on that transcription factors might similarly position Ty3 integration. In vivo S. cerevisiae RNAPIII transcription requires general factors TFIIIC and TFIIIB. TFIIIC associates with promoter elements boxA and boxB of about 10 bp each located 20 bp and ~20 to ~100 bp downstream of the TSS, respectively. TFIIIC directs sequence independent binding of TFIIIB upstream of the TSS (163, 164). TFIIIB is comprised of subunits TATA binding protein (TBP), Brf1 and Bdp1. It is considered the initiation factor because once bound, it is sufficient in vitro to mediate multiple rounds of RNAP III transcription (165, 166). SNR6 and a subset of tDNAs, have an upstream TATA element that contributes to transcription and TSS selection (167, 168).

In vitro, in a purified system with a TATA-containing template, TBP mediates binding of Brf1, and the two are sufficient for transcription (169). Genomewide chromatin immunoprecipitation–DNA microarray experiments showed that in growing cultures, yeast RNAPIII genes are mostly occupied by TFIIIB, TFIIIC, and RNAPIII (170–172).

The target requirements of Ty3 for integration are similar to requirements for transcription. Mutagenesis of targets showed that promoter elements are required for Ty3 integration in vivo, and underscored roles for the TATA box and boxA in copositioning the TSS and integration site (162, 173). In vitro Ty3 integration can be detected by PCR using primers binding within Ty3 and downstream of the target. Early experiments used VLPs as a source of cDNA and IN protein. Experiments showed that TFIIIC and TFIIIB are required to target integration to TATA-less tDNAs (174), but that TFIIIC is dispensable at genes where the upstream TATA element directly mediates binding of TBP and Brf1 (175, 176). In vitro at least, RNAPIII is not required and competes with the PIC for access to targets (177). A refined system uses recombinant IN and a fusion of Brf1 and TBP constructed as a structural mimetic (Brf1NTD-TBP-Brf1CTD) (178). In this reaction, in the presence of Mg2+, IN targets strand transfer of the 3’ end of a DNA duplex representing the terminal 20 bp of Ty3 cDNA into the SNR6 TSS (47). Surprisingly, in the presence of Mn2+, strand transfer independent of Brf1-TBP-Brf1 occurs within the plasmid target proximal to sequences resembling the Ty3 LTR ends, rather than the RNAPIII TSS (47).

**Genomewide identification of Ty3 targets and RNAPIII factor binding sites**

Next generation sequencing enables genomewide assessment of retroelement targeting and use of Ty3 as a probe for RNAPIII-transcribed genes, as well as an entrée to additional features of RNAPIII activation (38). Next generation sequencing analysis of genomic DNA from cells representing ~10,000 independent Ty3-HIS3 retrotransposition events identified ~300 significant insertion sites. These included 275 tDNAs, and the eight additional types of RNAPIII-transcribed genes (iYGR033C, RNA170, RPR1, SCR1, SNR52, SNR6, ZOD1, and RDN5) and 18 integrations associated with LTRs. Because of the frequent occurrence of boxA and boxB sequences in the S. cerevisiae genome, it was possible that integrations would associate with boxA–boxB pairs resembling RNAPIII promoters. Five insertion sites were identified that were not known RNAP
III-transcribed genes. However, insertions were flanked by boxA- and boxB-like sequences. Two of the five sites were subcloned and shown to act as plasmid-borne targets. Targeting was sensitive to mutations in conserved positions of the boxB promoter element. There are boxB sites in the *S. cerevisiae* genome which are bound by “Extra TFIIC” (ETC), but little or no TFIIB (170–172).

These are not targeted by Ty3, further underscoring the requirement for TFIIB of Ty3 integration.

One of the striking observations enabled by next generation sequencing of Ty3 insertion sites was that tDNA families, within which sequences are similar or identical, generated widely different numbers of sequencing reads (38). Ty3 could use identical genes at widely differing frequencies because of different chromatin contexts. RNAP III genes undergo specific types of epigenetic modifications and bind chromatin remodelers and condensins (179–190). In addition to possible differences in chromatin context, tDNAs also act as replication fork barriers (191), and subpopulations localize to the nucleolus (31, 192) and nuclear pores (30). Overall, it seems that we are just beginning to learn about the dynamic activities of RNAP III-transcribed genes. Ty3 offers a potentially exquisitely sensitive, indelible, in vivo probe for transient interactions between these genes and their chromatin caretakers.

Integration host factors

Although host RNAP III machinery plays a clear role in targeting, most genes encoding subunits of key complexes are essential and were not identified in host factor screens (64, 132). Exceptions were the La protein, a small-RNA chaperone (41) and Tfc1, a subunit of TFIIC (193). Despite the association of La with RNAP III-transcribed precursor RNAs (194), the *lp14* mutant had reduced Ty3 cDNA, so that it seems likely that the role of La precedes the integration step. A mutagenized strain expressing a C-terminal truncation of Tfc1 was identified in a screen for genes which affected recovery of integrations that activated expression of the suppressor tDNA in the divergent tDNA target described above. Although Ty3 cDNA integrates into this target in both orientations, Ty3 integration occurred in only one orientation in this mutant. This result indicates that the Ty3 PIC itself is asymmetric. A combination of genetic experiments, mutagenesis and in vitro pull-down assays indicated that the N-terminal extended domain of IN interacts directly with Tfc1 (193).

The example of fungal retrotransposons and availability of RNAi and high throughput sequencing spurred productive investigation into host factors involved in targeting retrovirus integration. Although there are no known examples of the highly specific targeting or targeting of RNAP III genes, as observed for a subset of retrotransposons, retroviral integration is also nonrandom. Retrotransposons generally avoid coding regions and a subset target inactive chromatin via interactions between IN chromodomains and H3K9Me. In contrast, retroviruses integrate into chromatin associated with epigenetic marks of transcriptionally active chromatin (195). HIV-1 associates with chromatin by interaction between the IN catalytic core domain and LEDGF/P75. This factor interacts with an epigenetic mark of active chromatin, H3K36Me3, via the LEDGF PWWP domain (reviewed in reference 196) (see Bushman and Craigie, Chapter 45). The murine leukemia virus IN associates with targets via interaction between the IN CTD and Bromo-domain and ExtraTerminal domains proteins, which bind acetylated H3 and H4 tails (197) (see Bushman and Craigie, Chapter 45).

**TY3–HOST INTERACTIONS**

Genetic screens and cytological analyses have identified large numbers of candidate Ty host factor genes. Importantly, many of these cluster by function (64, 81, 132, 198–200). Overall there was about 15% overlap between the genes identified in Ty3 and Ty1 screens despite the relatively different assays used. Genes identified in Ty3 genomewide screens were grouped according to the gene ontology process (Figure 4). Cofactors assisting retrotransposition included those involved in catabolism, PB components, nucleoporins, checkpoint control, and vesicular trafficking proteins including ESCRT factors. The most striking class of restriction factors antagonizing retrotransposition were those related to DNA replication, particularly helicases, and certain RNA transporters, including Npl3 andSac3. A small, but significant, percentage of cofactors and restriction factors that overlapped with those identified in Ty1 screens included replication factors (64, 199, 201), PB proteins (81, 198), and several RTT genes of various functions including chromatin remodelers and histone modification factors (64, 199).

Genes involved in replication and responses to DNA damage were the major class of restriction factors. Deletions of genes encoding replication factors Csm3, Rrm3, Mrc1, Rad6, Bre1, Sgs1 and Hpr5 enhanced Ty3 retrotransposition. Work in the Curcio laboratory on Ty1 showed that replication stress or DNA damage caused by the absence of a subset of DNA maintenance factors triggers increased Ty1 cDNA synthesis. This was traced
to activation of checkpoint proteins Rad24 and Rad9, as the effect was significantly reduced in backgrounds lacking one of the checkpoint genes (201). Although epistasis analysis was not conducted for Ty3, retrotransposition was decreased in the \( \text{rad24}\Delta \) background (64) consistent with checkpoint control promoting Ty3 retrotransposition. However, tDNA targeting might also be affected directly by mutations in replication factors. Transfer RNA genes block replication forks coming from downstream (191). If incoming replication forks disrupt targeting, fork stalling could increase integration. Alternatively, fork breakage caused by loss of key helicases could trigger localization to NPC (125), enhancing accessibility of tDNA targets to incoming integrases.

**FIGURE 4** Hierarchical clustering of Ty3 cofactors and restriction factors by gene ontology groups. Gene ontology (GO) analysis was performed using the GO SLIM Biological Process mapping tool available through Saccharomyces Genome Database (http://www.yeastgenome.org). Knockout mutants identified as having either increased "Up" or decreased "Down" Ty3 retrotransposition phenotypes were analyzed for GO: Biological Process terms. Enriched categories were determined using chi-squared test. GO categories were considered enriched if two criteria were met: (i) the \( P \)-value was <0.05 and (ii) the number of genes in the enriched category exceeded 10% of the total number of genes in the Up or Down list. Enriched categories were converted to a heat map with hierarchical clustering using R; values represent the \(-\log(P\text{-value})\) scaled from 0 (no significance) to 1, blue coloring reflects the intensity of significance. doi:10.1128/MicrobiolSpectrum.MDNA3-0057-2014.f4
As discussed above, Ty3 retrotransposition increased in mutants with deletions of genes encoding a subset of Nup84 complex components, *nup84Δ, nup120Δ*, and *nup133Δ* (64). We speculate that clustering of mutant pores and retromes could enhance nuclear access. Alternatively, these mutations could simply make the pore more permissive for PIC translocation (133).

Among the Ty3 cofactor genes were vesicular trafficking proteins. In budding yeast, ESCRT proteins were identified as essential for the late stages of trafficking of ubiquitinated surface proteins from the plasma membrane into multivesicular bodies (202). Subsequently, these factors were shown to be critical for retrovirus budding, which is the topological equivalent of multivesicular body formation (62). ESCRT proteins identified as Ty3 cofactors include ESCRTI (Vps28), ESCRTII, (Vps25, Vps22, Vps36), and ESCRTIII (Vps20, Snf7), as well as ESCRT-related proteins Vps4 and Bro1 (64) and Vps27. Although there is no evidence of Ty3 budding, it could associate with membranes in the course of morphogenesis. An interesting alternative possibility is raised by the recent implication of ESCRT components in NPC maintenance (203).

Retrovirus genomewide host screens, with HIV-1 being a major focus, have identified large numbers of host factors, among them, many nuclear proteins (see Bushman and Craigie, Chapter 45). Metadata analysis of nine such screens concluded that ∼5% of all human genes encode factors related to HIV-1 and showed low, but statistically significant, overlap (204). As in the case of the Ty elements, overlap is limited by differences among screening strategies. Ty3 host factors (64, 132) were compared with factors identified in a subset of retrovirus siRNA host factor screens (205–208) using the Overlapper tool (http://hivsystemsbiology.org/GenesListOverlapper/app).

These results are shown in Table 1. Functional overlaps include overlaps involved in vesicular trafficking, nuclear access (porins and translocation factors), DNA repair, ubiquitin signaling, mediator complex/transcription, RNA binding, and DNA and RNA unwinding (reviewed in references 62, 209–211).

Nucleoporins and NPC translocation factors were identified as host cofactors for Ty3 and for HIV-1, which infects nondividing cells. Prominent among these are the hydrophobic Nups enriched in FG repeats (122, 123). Human HsNup98, an FG Nup (212, 213), interacts with HIV-1 Gag and is required for HIV-1 nuclear entry. Budding yeast homologs of HsNup98, ScNup116, and ScNup100 are required for WT levels of Ty3 nuclear entry and recombinant ScNup116 interacts with Ty3 VLPs and recombinant Gag3 in vitro. A role in for this class of Nups in moving the uncoating VLP through the pore was proposed (133). *Schizosaccharomyces pombe* basket NUP, SpNup124, was isolated as a factor required for nuclear entry of Ty3/Gypsy element TFI Gag cores (129). The *S. cerevisiae* homolog, ScNup1, is essential, and so was underrepresented in Ty screens. However, HsNup153, the human homolog, is required for HIV-1 nuclear entry and mutations shift the integration bias, suggesting that basket Nups could affect NPC-proximal integration site selection (214). These observations suggest that interactions between Gag and FG Nups are a common feature of retroelement nuclear pore translocation and possibly uncoating.

Ribonucleoprotein granule components are common themes in retrotransposon and retrovirus assembly, but in seemingly distinct roles. During assembly, RNA and protein from LTR retrotransposon Tys and non-LTR retrotransposon LI LINEs, interact productively with PB and SG components (133, 215, 216). Mouse mammary tumor virus assembles intracellularly in association with YB-1-enriched ribonucleoprotein granules (217). YB-1 is a translational regulatory protein which accumulates in SG during stress, but is also found in PB. Expression of mouse mammary tumor virus Gag causes granules to increase in size, and these are distinct from SG and PB. HIV-1, which assembles into cores during extrusion into budding plasma membrane, actively antagonizes formation of SG, but associates with Staufen in cytoplasmic ribonucleoprotein particles (218). Host HIV-1 restriction factors, including apolipoprotein B mRNA editing enzyme catalytic polypeptide-like (APOBEC) cytidine deaminase family members, overexpressed Mov10 helicase (219–221), and RNAi components, associate with PB (210, 222). The PB factors also restrict endogenous retroviruses (223). However, studies disagree about the positive contributions of PB factors to retrovirus infection. DEAD box helicase DDX6 (*ScDhh1*) is ascribed a positive role in the intracellular packaging of prototypic foamy virus RNA, but studies differ as to roles in HIV-1 core assembly (224, 225). Overall these studies suggest that individual components of ribonucleoprotein granules, including PB and SG, interact with assembling retrovirus RNA and proteins, but assembly is not associated with these granules *per se*. Although the examples are limited, they are consistent with ribonucleoprotein granules sequestering proteins to protect host RNAs from inappropriate degradation or translational suppression, while providing a source of soluble proteins to interact with specific components of viruses. These associate with specialized granule components intracellularly, or complete assembly during budding on the plasma membrane.
### TABLE 1 Host factors in common between Ty3 and HIV-1

<table>
<thead>
<tr>
<th>ORF</th>
<th>Gene name</th>
<th>Full name</th>
<th>Ty3 ΔTPN[^1]</th>
<th>Human homolog</th>
<th>Retrovirus host factor[^2]</th>
</tr>
</thead>
<tbody>
<tr>
<td>YBR284W</td>
<td>YBR284W</td>
<td></td>
<td>up, down</td>
<td>AMPD1, AMPD2, AMPD3</td>
<td></td>
</tr>
<tr>
<td>YAR0035W</td>
<td>SWD1</td>
<td>Set1c, WD40 repeat protein</td>
<td>up, down</td>
<td>RBBP5, WDR72</td>
<td></td>
</tr>
<tr>
<td>YPR119W</td>
<td>CLB2</td>
<td>CycLin B</td>
<td>up</td>
<td>CCNA1, CCNA2, CCNB1, CCNB2, CCNB3, CCNE1, CCNE2, CNTD2</td>
<td></td>
</tr>
<tr>
<td>YPL256C</td>
<td>CLN2</td>
<td>CycLinN</td>
<td>up</td>
<td>CCNB1, CCNB2, CNTD2</td>
<td></td>
</tr>
<tr>
<td>YPL024W</td>
<td>RM1</td>
<td>RecQ Mediated genome Instability</td>
<td>up</td>
<td></td>
<td></td>
</tr>
<tr>
<td>YOR144C</td>
<td>ELG1</td>
<td>Enhanced Level of Genomic instability</td>
<td>up</td>
<td></td>
<td></td>
</tr>
<tr>
<td>YOR39W</td>
<td>CKB2</td>
<td>Casein Kinase Beta' subunit</td>
<td>up</td>
<td>CSNK2B</td>
<td></td>
</tr>
<tr>
<td>YMR052C</td>
<td>POP2</td>
<td>PKG promoter directed OverProduction</td>
<td>up</td>
<td>CNOT7, CNOT8</td>
<td></td>
</tr>
<tr>
<td>YNL242W</td>
<td>ATG2</td>
<td>AuTophaGy related</td>
<td>up</td>
<td>ATG2A, ATG2B</td>
<td></td>
</tr>
<tr>
<td>YNL138W</td>
<td>SRV2</td>
<td>Suppressor of RasVal19</td>
<td>up</td>
<td>CAP1, CAP2</td>
<td></td>
</tr>
<tr>
<td>YMR190C</td>
<td>SG1</td>
<td>Slow Growth Suppressor</td>
<td>up</td>
<td>BLM, RECQL, RECQL5, WRN</td>
<td></td>
</tr>
<tr>
<td>YMR048W</td>
<td>CSM3</td>
<td>Chromosome Segregation in Meiosis</td>
<td>up</td>
<td></td>
<td></td>
</tr>
<tr>
<td>YML051W</td>
<td>GAL80</td>
<td>GALactose metabolism</td>
<td>up</td>
<td>GFOD1, GFOD2</td>
<td></td>
</tr>
<tr>
<td>YLR399C</td>
<td>BDF1</td>
<td>BromoDomain Factor</td>
<td>up</td>
<td>BRD2, BRD3, BRD4, BRDT, CREBBP, EP300</td>
<td></td>
</tr>
<tr>
<td>YLR268W</td>
<td>SEC22</td>
<td>SECretory</td>
<td>up</td>
<td>SEC22A, SEC22C</td>
<td></td>
</tr>
<tr>
<td>YLR235C</td>
<td>YLR235C</td>
<td></td>
<td>up</td>
<td></td>
<td></td>
</tr>
<tr>
<td>YLR015W</td>
<td>BRE2</td>
<td>BREfelin A sensitivity</td>
<td>up</td>
<td>ASH2L</td>
<td></td>
</tr>
<tr>
<td>YLL044W</td>
<td>YLL044W</td>
<td></td>
<td>up</td>
<td></td>
<td></td>
</tr>
<tr>
<td>YLL030C</td>
<td>RRT7</td>
<td>Regulator of rDNA Transcription</td>
<td>up</td>
<td></td>
<td></td>
</tr>
<tr>
<td>YLL002W</td>
<td>RTT109</td>
<td>Regulator of Ty1 Transposition</td>
<td>up</td>
<td></td>
<td></td>
</tr>
<tr>
<td>YKR082W</td>
<td>NUP133</td>
<td>Nuclidean Pore</td>
<td>up</td>
<td>NUP133</td>
<td></td>
</tr>
<tr>
<td>YKL135C</td>
<td>APL2</td>
<td>clathrin Adaptor Protein complex Large chain</td>
<td>up</td>
<td>AP1B1, AP2B1, AP4B1</td>
<td></td>
</tr>
<tr>
<td>YKL068W</td>
<td>NUP100</td>
<td>Nuclidean Pore</td>
<td>up</td>
<td>NUP98, NUP1L</td>
<td></td>
</tr>
<tr>
<td>YKL057C</td>
<td>NUP120</td>
<td>Nuclidean Pore</td>
<td>up</td>
<td></td>
<td></td>
</tr>
<tr>
<td>YKL041W</td>
<td>VPS24</td>
<td>Vacular Protein Sorting</td>
<td>up</td>
<td>CHMP3</td>
<td></td>
</tr>
<tr>
<td>YLO92W</td>
<td>SR52</td>
<td>Suppressor of Rad Six</td>
<td>up</td>
<td></td>
<td></td>
</tr>
<tr>
<td>YLL047C</td>
<td>RRT101</td>
<td></td>
<td>up</td>
<td>CUL1</td>
<td></td>
</tr>
<tr>
<td>YL021W</td>
<td>MRS1</td>
<td>Mitochondrial RNA Splicing</td>
<td>up</td>
<td></td>
<td></td>
</tr>
<tr>
<td>YL090W</td>
<td>ICE2</td>
<td>Inheritance of Cortical ER</td>
<td>up</td>
<td></td>
<td></td>
</tr>
<tr>
<td>YL069C</td>
<td>RPS24B</td>
<td>Ribosomal Protein of the Small subunit</td>
<td>up</td>
<td>RPS24</td>
<td></td>
</tr>
<tr>
<td>YL015W</td>
<td>BAR1</td>
<td>BARrner to the alpha factor response</td>
<td>up</td>
<td>CTS5, CTSE, NAPSA, PGA3, PGA4, PGA5, PGC, REN</td>
<td></td>
</tr>
<tr>
<td>YL011W</td>
<td>TIR3</td>
<td>Tlp1-Related</td>
<td>up</td>
<td></td>
<td></td>
</tr>
<tr>
<td>YL009C-A</td>
<td>EST3</td>
<td>Ever Shorter Telomerases</td>
<td>up</td>
<td></td>
<td></td>
</tr>
<tr>
<td>YHR177W</td>
<td>YHR177W</td>
<td></td>
<td>up</td>
<td></td>
<td></td>
</tr>
<tr>
<td>YHR031C</td>
<td>RRM3</td>
<td>rDNA Recombination Mutation</td>
<td>up</td>
<td>PIF1</td>
<td>KIAA0368</td>
</tr>
<tr>
<td>YHL030W</td>
<td>ECM29</td>
<td>ExtraCellular Mutant</td>
<td>up</td>
<td></td>
<td></td>
</tr>
<tr>
<td>YHL003C</td>
<td>LAG1</td>
<td>Longevity Assurance Gene</td>
<td>up</td>
<td>CERS1, CERS2, CERS3, CERS4, CERS5, CERS6, TRAM1, TRAM1L1, TRAM2</td>
<td></td>
</tr>
<tr>
<td>YGR270W</td>
<td>YTA7</td>
<td>Yeast Tat-binding Analog</td>
<td>up</td>
<td></td>
<td></td>
</tr>
<tr>
<td>YGRO56W</td>
<td>RSC1</td>
<td>Remodel the Structure of Chromatin</td>
<td>up</td>
<td>PBRM1</td>
<td></td>
</tr>
<tr>
<td>YGL056W</td>
<td>BUD13</td>
<td>BUD site selection</td>
<td>up</td>
<td>BUD13</td>
<td></td>
</tr>
<tr>
<td>YGL066W</td>
<td>SGF73</td>
<td>SaGa associated Factor, 73 kDa</td>
<td>up</td>
<td>ATXN7, ATXN7L1, ATXN7L2, ATXN7L3, ATXN7L3B</td>
<td></td>
</tr>
<tr>
<td>YGL058W</td>
<td>RAD6</td>
<td>RADiaction sensitive</td>
<td>up</td>
<td>UBE2A, UBE2B, UBE25, UBE2W</td>
<td></td>
</tr>
<tr>
<td>YER155C</td>
<td>BEM2</td>
<td>Bud EMergence</td>
<td>up</td>
<td>ARHGAP21, ARHGAP23, ARHGAP35, CHN1, GMIP, HMHA1, RACGAP1, SYDE1, SYDE2</td>
<td></td>
</tr>
<tr>
<td>YER060W-A</td>
<td>FCY22</td>
<td>FluoroCYtosine resistance</td>
<td>up</td>
<td></td>
<td></td>
</tr>
<tr>
<td>YER057C</td>
<td>HMF1</td>
<td>Homologous Mmf1p Factor</td>
<td>up</td>
<td>HRSP12</td>
<td></td>
</tr>
<tr>
<td>YEL013W</td>
<td>VAC8</td>
<td>VACuole related</td>
<td>up</td>
<td>ANKAR, ARMC3, ARMC4, CTNNB1, JUP</td>
<td></td>
</tr>
<tr>
<td>YEL004W</td>
<td>YEA4</td>
<td></td>
<td>up</td>
<td>SLC35B4</td>
<td></td>
</tr>
<tr>
<td>YDR448W</td>
<td>ADA2</td>
<td>transcriptional ADAptor</td>
<td>up</td>
<td>TADA2A, TADA2B</td>
<td></td>
</tr>
</tbody>
</table>

(continued)
### TABLE 1 Host factors in common between Ty3 and HIV-1 (continued)

<table>
<thead>
<tr>
<th>Gene name</th>
<th>Full name</th>
<th>Ty3 ΔTPN</th>
<th>Human homolog</th>
<th>Retrovirus host factor</th>
</tr>
</thead>
<tbody>
<tr>
<td>YDR432W</td>
<td>NPL3</td>
<td>up</td>
<td>CIRBP, DMKN, HNRNPA1, HNRNPA1L2, HNRNPA2B1, HNRNPA3, LOR, RBMX, RBMXL1, RBMXL2, RBMY1A1, RBMY1B, RBMY1D, RBMY1E, RBMY1F, RBMY1J, SRSF1, SRSF9</td>
<td></td>
</tr>
<tr>
<td>YDR290W</td>
<td>YDR290W</td>
<td>up</td>
<td></td>
<td></td>
</tr>
<tr>
<td>YDR289C</td>
<td>RTT103</td>
<td>up</td>
<td>RPRD1A, RPRD1B, RPRD2</td>
<td></td>
</tr>
<tr>
<td>YDR176W</td>
<td>NGG1</td>
<td>up</td>
<td>TADA3</td>
<td></td>
</tr>
<tr>
<td>YDR159W</td>
<td>SAC3</td>
<td>up</td>
<td>MCM3AP, SAC3D1</td>
<td></td>
</tr>
<tr>
<td>YDR074W</td>
<td>TPS2</td>
<td>up</td>
<td></td>
<td></td>
</tr>
<tr>
<td>YDL16W</td>
<td>NUP84</td>
<td>up</td>
<td>NUP107</td>
<td>(206)</td>
</tr>
<tr>
<td>YDL074C</td>
<td>BRE1</td>
<td>up</td>
<td>RNF20, RNF40</td>
<td></td>
</tr>
<tr>
<td>YDL056W</td>
<td>MBP1</td>
<td>up</td>
<td></td>
<td></td>
</tr>
<tr>
<td>YCR020W</td>
<td>HTL1</td>
<td>up</td>
<td></td>
<td></td>
</tr>
<tr>
<td>YCL061C</td>
<td>MRC1</td>
<td>up</td>
<td></td>
<td></td>
</tr>
<tr>
<td>YCL037C</td>
<td>SRO9</td>
<td>up</td>
<td></td>
<td></td>
</tr>
<tr>
<td>YBR277C</td>
<td>YBR277C</td>
<td>up</td>
<td></td>
<td></td>
</tr>
<tr>
<td>YBR269C</td>
<td>FMP21</td>
<td>up</td>
<td>C6orf57</td>
<td></td>
</tr>
<tr>
<td>YBL094C</td>
<td>YBL094C</td>
<td>up</td>
<td></td>
<td></td>
</tr>
<tr>
<td>YBL065C</td>
<td>LDB7</td>
<td>up</td>
<td></td>
<td></td>
</tr>
<tr>
<td>YPL084W</td>
<td>BRO1</td>
<td>BCK1-like resistance to Osmotic shock</td>
<td>PDCD6IP, PTPN23, RHPN1, RHPN2</td>
<td></td>
</tr>
<tr>
<td>YNL236W</td>
<td>SIN4</td>
<td>Switch Independent</td>
<td>down</td>
<td></td>
</tr>
<tr>
<td>YMR154C</td>
<td>RIM13</td>
<td>Regulator of IME2</td>
<td>down</td>
<td></td>
</tr>
<tr>
<td>YPR173C</td>
<td>VPS4</td>
<td>Vacuolar Protein Sorting</td>
<td>down</td>
<td>KATNA1, KATNAL1, KATNAL2, VPS4A, VPS4B</td>
</tr>
<tr>
<td>YPL225W</td>
<td>YPL225W</td>
<td>down</td>
<td>PBDC1</td>
<td></td>
</tr>
<tr>
<td>YPL150W</td>
<td>YPL150W</td>
<td>down</td>
<td>MELK, IPO11</td>
<td></td>
</tr>
<tr>
<td>YPL125W</td>
<td>KAP120</td>
<td>down</td>
<td>IPO11</td>
<td></td>
</tr>
<tr>
<td>YPL065W</td>
<td>VPS28</td>
<td>Vacuolar Protein Sorting</td>
<td>down</td>
<td>VPS28</td>
</tr>
<tr>
<td>YPL042C</td>
<td>SNSN</td>
<td>down</td>
<td>CDK19, CDK8</td>
<td></td>
</tr>
<tr>
<td>YPL002C</td>
<td>SNF8</td>
<td>Snf8 NonFermenting</td>
<td>down</td>
<td>SNF8</td>
</tr>
<tr>
<td>YOR275C</td>
<td>RIM20</td>
<td>Regulator of IME2</td>
<td>down</td>
<td>PDCD6IP, PTPN23</td>
</tr>
<tr>
<td>YOR270C</td>
<td>VPH1</td>
<td>Vacuolar pH</td>
<td>down</td>
<td>ATP6VOA1, ATP6VOA2, ATP6VOA4, TCIRG1</td>
</tr>
<tr>
<td>YOR055W</td>
<td>YOR055W</td>
<td>down</td>
<td>PPCDC</td>
<td></td>
</tr>
<tr>
<td>YOR054C</td>
<td>VHS3</td>
<td>Viable in a Hal3 St4 background</td>
<td>down</td>
<td></td>
</tr>
<tr>
<td>YOL083W</td>
<td>ATG34</td>
<td>AuTophaGy related</td>
<td>down</td>
<td></td>
</tr>
<tr>
<td>YNR059W</td>
<td>MNT4</td>
<td>MaNnosylTransferase</td>
<td>down</td>
<td></td>
</tr>
<tr>
<td>YNR006W</td>
<td>VPS27</td>
<td>Vacuolar Protein Sorting</td>
<td>down</td>
<td>GGA1, GGA2, GGA3, HGS, TOM1, TOM1L1, TOM1L2, WDFY1, WDFY2, ZFYVE21</td>
</tr>
<tr>
<td>YNL224C</td>
<td>SQS1</td>
<td>SQuelch of Splicing suppression</td>
<td>down</td>
<td>SON</td>
</tr>
<tr>
<td>YNL136W</td>
<td>EAF7</td>
<td>Esa1-Associated Factor</td>
<td>down</td>
<td>MRGBP</td>
</tr>
<tr>
<td>YNL059C</td>
<td>ARP5</td>
<td>Actin-Related Protein</td>
<td>down</td>
<td>ACTR5</td>
</tr>
<tr>
<td>YMR158W</td>
<td>YMR158W</td>
<td>down</td>
<td></td>
<td></td>
</tr>
<tr>
<td>YMR123W</td>
<td>PKR1</td>
<td>Pichia farinosa Killer toxin Resistance</td>
<td>down</td>
<td></td>
</tr>
<tr>
<td>YMR116C</td>
<td>ASC1</td>
<td>Absence of growth Suppressor of Cyp1</td>
<td>down</td>
<td>GNB2L1, WDR31</td>
</tr>
<tr>
<td>YMR077C</td>
<td>VPS20</td>
<td>Vacuolar Protein Sorting</td>
<td>down</td>
<td>CHMP6, CHMP7</td>
</tr>
<tr>
<td>YMR058W</td>
<td>FET3</td>
<td>FErrous Transport</td>
<td>down</td>
<td></td>
</tr>
<tr>
<td>YML121W</td>
<td>GTR1</td>
<td>GTP binding protein Resemblance</td>
<td>down</td>
<td>RAGA, RRAGB</td>
</tr>
<tr>
<td>YML097C</td>
<td>VPS9</td>
<td>Vacuolar Protein Sorting</td>
<td>down</td>
<td>GAPVD1, RABGEF1, RIN1, RIN2, RIN3, RINL, VPS9D1</td>
</tr>
<tr>
<td>YML081C</td>
<td>ATP18</td>
<td>ATP synthase</td>
<td>down</td>
<td></td>
</tr>
<tr>
<td>YML075C</td>
<td>RPL6A</td>
<td>Ribosomal Protein of the Large subunit</td>
<td>down</td>
<td>RPL6</td>
</tr>
<tr>
<td>YLR423C</td>
<td>ATG17</td>
<td>AuTophaGy related</td>
<td>down</td>
<td></td>
</tr>
<tr>
<td>YLR417W</td>
<td>VPS36</td>
<td>Vacuolar Protein Sorting</td>
<td>down</td>
<td>VPS36</td>
</tr>
</tbody>
</table>

(continued)
### TABLE 1 Host factors in common between Ty3 and HIV-1 (continued)

<table>
<thead>
<tr>
<th>ORF</th>
<th>Gene name</th>
<th>Full name</th>
<th>Ty3 ΔTPN1</th>
<th>Human homolog</th>
<th>Retrovirus host factor</th>
</tr>
</thead>
<tbody>
<tr>
<td>YLR242C</td>
<td>ARV1</td>
<td>ARE2 Required for Viability</td>
<td>down</td>
<td>ARV1</td>
<td></td>
</tr>
<tr>
<td>YLR025W</td>
<td>SNF7</td>
<td>Sucrose NonFermenting</td>
<td>down</td>
<td>CHMP4A, CHMP4B, CHMP4C</td>
<td></td>
</tr>
<tr>
<td>YKR099W</td>
<td>BA51</td>
<td>BASal</td>
<td>down</td>
<td>MYB, MYBL1, MYBL2, SNAPC4</td>
<td></td>
</tr>
<tr>
<td>YKR020W</td>
<td>VPS51</td>
<td>Vacular Protein Sorting</td>
<td>down</td>
<td>DBR1</td>
<td></td>
</tr>
<tr>
<td>YKL149C</td>
<td>DBR1</td>
<td>DeBRanching</td>
<td>down</td>
<td>DBR1</td>
<td></td>
</tr>
<tr>
<td>YJR137C</td>
<td>MET5</td>
<td>METHionine requiring</td>
<td>down</td>
<td>DNR1</td>
<td></td>
</tr>
<tr>
<td>YJR121W</td>
<td>ATP2</td>
<td>ATP synthase</td>
<td>down</td>
<td>ATP5A1, ATP5B</td>
<td>(209)</td>
</tr>
<tr>
<td>YJR120W</td>
<td>YJR120W</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>YJR102C</td>
<td>VPS25</td>
<td>Vacular Protein Sorting</td>
<td>down</td>
<td>VPS25</td>
<td></td>
</tr>
<tr>
<td>YJR032W</td>
<td>CPR7</td>
<td>Cyclisoprotein-sensitive Proline Rotamase</td>
<td>down</td>
<td>NKTR, PPID, PPIG</td>
<td></td>
</tr>
<tr>
<td>YJL006C</td>
<td>CTX2</td>
<td>Carboxy-Terminal domain Kinase</td>
<td>down</td>
<td>CCNT1, CCNT2</td>
<td>(209) (206)</td>
</tr>
<tr>
<td>YIL134W</td>
<td>FLX1</td>
<td>Flavin eXchange</td>
<td>down</td>
<td>SLC25A32</td>
<td></td>
</tr>
<tr>
<td>YIL115C</td>
<td>NUP159</td>
<td>NUclear Pore</td>
<td>down</td>
<td>MAEA</td>
<td></td>
</tr>
<tr>
<td>YIL097W</td>
<td>FYV10</td>
<td>Function required for Yeast Viability</td>
<td>down</td>
<td>DDB1</td>
<td></td>
</tr>
<tr>
<td>YIL073C</td>
<td>SPO22</td>
<td>SPOrulation</td>
<td>down</td>
<td></td>
<td></td>
</tr>
<tr>
<td>YHR082C</td>
<td>KSP1</td>
<td></td>
<td>down</td>
<td>CHEK2</td>
<td></td>
</tr>
<tr>
<td>YHR026W</td>
<td>VMA16</td>
<td>Vacular Membrane Atpase</td>
<td>down</td>
<td>ATP6V0B</td>
<td>(208)</td>
</tr>
<tr>
<td>YHR008C</td>
<td>SOD2</td>
<td>SuperOxide Dismutase</td>
<td>down</td>
<td>SOD2</td>
<td></td>
</tr>
<tr>
<td>YGR246C</td>
<td>BRF1</td>
<td>B-Related Factor</td>
<td>down</td>
<td>BRF1</td>
<td>(209)</td>
</tr>
<tr>
<td>YGR188C</td>
<td>BUB1</td>
<td>Budding Uninhibited by Benzimidazole</td>
<td>down</td>
<td>BUB1, BUB1B</td>
<td></td>
</tr>
<tr>
<td>YGR167W</td>
<td>CTC2</td>
<td>Clathrin Light Chain</td>
<td>down</td>
<td>CLTA, CLTB</td>
<td>(206)</td>
</tr>
<tr>
<td>YGL212W</td>
<td>VAM7</td>
<td>Vacular Morphogenesis</td>
<td>down</td>
<td>STX8</td>
<td></td>
</tr>
<tr>
<td>YGL173C</td>
<td>XRN1</td>
<td>eXoribonuclease</td>
<td>down</td>
<td>XRN1</td>
<td></td>
</tr>
<tr>
<td>YGL156W</td>
<td>AMS1</td>
<td>Alpha-MannoNsidase</td>
<td>down</td>
<td>MAN2C1</td>
<td></td>
</tr>
<tr>
<td>YGL078C</td>
<td>DBP3</td>
<td>Dead Box Protein</td>
<td>down</td>
<td>DDX21, DDX50</td>
<td>(208)</td>
</tr>
<tr>
<td>YGL057C</td>
<td>GEP7</td>
<td>GEnetic interactors of Prohibitins</td>
<td>down</td>
<td>DDX21</td>
<td></td>
</tr>
<tr>
<td>YFR019W</td>
<td>FAB1</td>
<td>Forms Aploid and Binucleate cells</td>
<td>down</td>
<td>PIKFYVE, ZFYVE21, ZFYVE28</td>
<td></td>
</tr>
<tr>
<td>YFR012W</td>
<td>DCV1</td>
<td>Demands Cdc28 kinase activity for Viability</td>
<td>down</td>
<td></td>
<td></td>
</tr>
<tr>
<td>YFR009W</td>
<td>GCN20</td>
<td>General Control Nonderepressible</td>
<td>down</td>
<td>ABCF1, ABCF3</td>
<td></td>
</tr>
<tr>
<td>YFL007W</td>
<td>BLM10</td>
<td>Bleomyacin resistance</td>
<td>down</td>
<td>PSM4</td>
<td></td>
</tr>
<tr>
<td>YFL001W</td>
<td>DEG1</td>
<td>DEpressed Growth rate</td>
<td>down</td>
<td>PUS3</td>
<td></td>
</tr>
<tr>
<td>YER173W</td>
<td>RAD24</td>
<td>RADiation sensitive</td>
<td>down</td>
<td>RAD17</td>
<td></td>
</tr>
<tr>
<td>YER148W</td>
<td>SPT15</td>
<td>SuPpressor of Ty</td>
<td>down</td>
<td>TBP, TBPL1, TBPL2</td>
<td>(209)</td>
</tr>
<tr>
<td>YER145C</td>
<td>FTR1</td>
<td>Fe TTransporter</td>
<td>down</td>
<td></td>
<td></td>
</tr>
<tr>
<td>YER105C</td>
<td>NUP157</td>
<td>NUclear Pore</td>
<td>down</td>
<td>NUP155</td>
<td>(206)</td>
</tr>
<tr>
<td>YEL044W</td>
<td>IES6</td>
<td>Into Eighty Subunit</td>
<td>down</td>
<td>INO80C</td>
<td></td>
</tr>
<tr>
<td>YDR493W</td>
<td>M2M1</td>
<td>Mitochondrial Zinc Maintenance</td>
<td>down</td>
<td></td>
<td></td>
</tr>
<tr>
<td>YDR323C</td>
<td>PEP7</td>
<td>carboxyPEPtidase Y-deficient</td>
<td>down</td>
<td>ZFYVE20</td>
<td></td>
</tr>
<tr>
<td>YDR139C</td>
<td>RUB1</td>
<td>Related to ubiquitin</td>
<td>down</td>
<td>NEDD8</td>
<td></td>
</tr>
<tr>
<td>YDR067C</td>
<td>OCA6</td>
<td>Oxidant-induced Cell-cycle Arrest</td>
<td>down</td>
<td>DDX6</td>
<td>(206)</td>
</tr>
<tr>
<td>YDL160C</td>
<td>DHH1</td>
<td>DEAD box Helicase Homolog</td>
<td>down</td>
<td>MAP2K1, MAP2K2, MAP2K3, SBK2, SLK, STK10, TAOK1, TAOK2, TAOK3</td>
<td>(206)</td>
</tr>
<tr>
<td>YDL159W</td>
<td>STE7</td>
<td>STErile</td>
<td>down</td>
<td></td>
<td></td>
</tr>
<tr>
<td>YDL118W</td>
<td>YDL118W</td>
<td></td>
<td>down</td>
<td></td>
<td></td>
</tr>
<tr>
<td>YDL088C</td>
<td>ASM4</td>
<td></td>
<td>down</td>
<td></td>
<td></td>
</tr>
<tr>
<td>YBR293W</td>
<td>VBA2</td>
<td>Vacular Basic Amino acid trans- porter</td>
<td>down</td>
<td>SLC18A1, SLC18A2, SLC18A3, SLC18B1</td>
<td>(206)</td>
</tr>
<tr>
<td>YBR263W</td>
<td>SHM1</td>
<td>Serine HydroxylMethyltransferase</td>
<td>down</td>
<td>SHMT1, SHMT2</td>
<td></td>
</tr>
<tr>
<td>YBR262C</td>
<td>AIM5</td>
<td>Altered Inheritance rate of Mitochondria</td>
<td>down</td>
<td></td>
<td></td>
</tr>
<tr>
<td>YBR103W</td>
<td>SIF2</td>
<td>Sir4p-Interacting Factor</td>
<td>down</td>
<td>TBL1X, TBL1XR1, TBL1Y, THOC3, WDR17</td>
<td></td>
</tr>
<tr>
<td>YBR044C</td>
<td>TCM62</td>
<td>TriCarboxylic acid cycle Mutant</td>
<td>down</td>
<td>HSPD1</td>
<td></td>
</tr>
<tr>
<td>YBR021W</td>
<td>FUR4</td>
<td>5-Fluorouridine sensitivity</td>
<td>down</td>
<td>NSUN2</td>
<td></td>
</tr>
<tr>
<td>YBL024W</td>
<td>NCL1</td>
<td>NuCLEAR protein</td>
<td>down</td>
<td>EPS15, EPS15L1, GRAP, GRAP2, GRB2, ITSN1, ITSN2, REPS1, REPS2, SLA, SLA2</td>
<td></td>
</tr>
<tr>
<td>YBL007C</td>
<td>SLA1</td>
<td>Synthetic Lethal with ABP1</td>
<td>down</td>
<td></td>
<td></td>
</tr>
<tr>
<td>YAL027W</td>
<td>SAW1</td>
<td>Single-strand Annealing Weakened</td>
<td>down</td>
<td>SMARCAD1</td>
<td></td>
</tr>
<tr>
<td>YAL019W</td>
<td>FUN30</td>
<td>Function Unknown Now</td>
<td>down</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

(continued)
The presence of a mixture of components in RNA granules that variously restrict and promote retroelement assembly might extend to the retroelement expression in germ cell-related lineages. In these cells, cycles of genome demethylation are accompanied by activation of retrotransposon transcription and post-transcriptional control in perinuclear granules containing PB components. Mouse, fly, nematode, and zebrafish germ cell-related lineages accumulate in these granules. When RNAi is suppressed, these cells undergo retrotransposition (95, 226–228). Based on the role of components of these granules in budding yeast retrotransposition, where RNAi is naturally absent, we speculate that in retrotransposing germ cell systems, at least a subset of these components interact with assembling retrotransposon proteins and RNAs.

**PERSPECTIVES**

The dawn of genomewide analysis enabled by next generation sequencing shows host genes embedded in a sea of retroelements. Together with other retrotransposons, LTR retrotransposons account for ongoing expansion, adaptation, and remodeling of eukaryotic genomes. Many consequences, noncoding RNAs and RNAi among them, were unexpected and raise new questions.

In the rush to apply these findings, retrotransposition itself remains largely unexplored. In animal cells, despite RNAi, LTR element retrotransposition is ongoing in germ cells, early-stage embryos, the brain, and some cancers. We know little of the mechanics of retrotransposition in these cells, and less about real-time effects of retrotransposition on those and descendant cells.

Ty3/Gypsy elements represent a major class of LTR retrotransposons, and dominate many plant and some animal genomes. Ty3, the sole representative in budding yeast, is one of the more completely understood of these elements at the molecular level. However, it is virtually a total anomaly. Classified as a Chromovirus, Ty3 lacks the defining IN chromodomain, and stands out for the specificity of its targeting. Most retrotransposons are expressed and retrotranspose in single cells, but Ty3 in the natural context of mating, retrotransposes only after diploids are formed. LTR retrotransposons are suppressed by RNAi, but due to the evolutionary loss of Dicer and Argonaute functions, this system is absent in budding yeast. Perhaps most surprisingly, Ty3 VLP assembly requires components of the RNA processing system that supposedly suppresses and degrades LTR retrotransposon RNAs in metazoans. What can we expect to learn from such an exceptional element?

Work in the Ty3 system has provided considerable information about the biochemistry of retrotransposition, much of which is proving generalizable. Three areas especially illustrate this point. First, Ty3 targeting provided the first clear evidence for interaction with host chromatin/transcription factors, and still offers the most precise tool for understanding the interaction between IN and target protein complexes and DNA. Targets in this system can be precisely predicted and manipulated to understand requirements of strand transfer. Second, Ty3/Tf1 Gypsy element retrotransposons show clear interactions with nuclear pores, and combined with targeting, can be used to explore the links between entry and integration. Known and essentially identical targets of Ty3 used at widely different frequencies, offer potential insights into unknown areas of nuclear dynamics.

### TABLE 1  Host factors in common between Ty3 and HIV-1 (continued)

<table>
<thead>
<tr>
<th>ORF</th>
<th>Gene name</th>
<th>Full name</th>
<th>Ty3 ΔTPN¹</th>
<th>Human homolog</th>
<th>Retrovirus host factor²</th>
</tr>
</thead>
<tbody>
<tr>
<td>YAL013W</td>
<td>DEP1</td>
<td>Disability in regulation of Expression of genes involved in Phospholipid biosynthesis</td>
<td>down</td>
<td>BFP2, DES, GFAP, IFLT1, INA, KRT1, KRT10, KRT12, KRT13, KRT14, KRT15, KRT16, KRT17, KRT18, KRT19, KRT2, KRT20, KRT23, KRT24, KRT25, KRT26, KRT27, KRT28, KRT3, KRT31, KRT32, KRT33A, KRT33B, KRT34, KRT35, KRT36, KRT37, KRT38, KRT39, KRT4, KRT40, KRT5, KRT6A, KRT6B, KRT6C, KRT7, KRT71, KRT72, KRT73, KRT74, KRT75, KRT76, KRT77, KRT78, KRT79, KRT8, KRT80, KRT81, KRT82, KRT83, KRT84, KRT85, KRT9, LMNA, LMNB1, LMNB2, NEFM, NES, SYNC, SYN1, VIM (208)</td>
<td></td>
</tr>
</tbody>
</table>

¹Genes identified in genomewide screens for Ty3 host factors (64, 133).
²Retrovirus host genes identified in Ty3 screens and in retrovirus host screens (206–209).
Third, because Ty3 is not regulated by post transcriptional RNA interference, it offers a rare opportunity to study retrotransposition itself. It cannot be a coincidence that Ty3 retrotransposes in the yeast analog of germ cells during cell fusion, and animal retrotransposons are activated in germ cells. We know quite a bit about how retrotransposition is suppressed. Now we should think about what happens when it occurs.

At the center of all this, we have little understanding of how retrotransposition affects cells in real time. In part this is because retrotransposition occurs in a few cells per population. The recent advent of single cell biology will surely produce a wave of insights into those rare cells that are undergoing retrotransposition. Are there clusters of retrotransposition in those cells? Does retrotransposition target broken replication forks as suggested by mutant screens? One DNA break induces checkpoint arrest; do retrotransposon cDNAs have the same effect? What determines the nuclear winner between competing processes of integration and destruction of invading cDNA? Does retrotransposition encrypt an individual cell record, as theorized for neuronal tissues? Is retrotransposition a trigger in the cancers with which it is associated? We know many mechanisms through which retrotransposons are tolerated by hosts. Is there also a kind of retrotransposon race that determines how retrotransposon composition differs so dramatically among eukaryotic species? Answers to these and other questions will lead us to a better understanding of our genomes.

ACKNOWLEDGMENTS
Ty3 research from the Sandmeyer laboratory was supported by PHS grant GM33281. We thank Ivan Chang and Parth Sitlani (UCI) for helpful discussions and other past and present members of the Sandmeyer laboratory who made this work possible. We thank Stuart Le Grice, NCI Frederick, MD, for helpful discussions, and Mercy Chang for artistic illustrations.

REFERENCES

Sandmeyer et al.


