



Conservation of mRNA quality control factor Ski7 and its diversification through changes in alternative splicing and gene duplication

Alexandra N. Marshall^a, Jaeil Han^a, Minseon Kim^a, and Ambro van Hoof^{a,1}

^aDepartment of Microbiology and Molecular Genetics, University of Texas Health Science Center at Houston and University of Texas Graduate School of Biomedical Sciences, Houston, TX 77030

Edited by Lynne E. Maquat, University of Rochester School of Medicine and Dentistry, Rochester, NY, and approved May 7, 2018 (received for review February 2, 2018)

Eukaryotes maintain fidelity of gene expression by preferential degradation of aberrant mRNAs that arise by errors in RNA processing reactions. In *Saccharomyces cerevisiae*, Ski7 plays an important role in this mRNA quality control by mediating mRNA degradation by the RNA exosome. Ski7 was initially thought to be restricted to *Saccharomyces cerevisiae* and close relatives because the *SKI7* gene and its paralog *HBS1* arose by whole genome duplication (WGD) in a recent ancestor. We have recently shown that the preduplication gene was alternatively spliced and that Ski7 function predates WGD. Here, we use transcriptome analysis of diverse eukaryotes to show that diverse eukaryotes use alternative splicing of *SKI7/HBS1* to encode two proteins. Although alternative splicing affects the same intrinsically disordered region of the protein, the pattern of splice site usage varies. This alternative splicing event arose in an early eukaryote that is a common ancestor of plants, animals, and fungi. Remarkably, through changes in alternative splicing and gene duplication, the Ski7 protein has diversified such that different species express one of four distinct Ski7-like proteins. We also show experimentally that the *Saccharomyces cerevisiae* *SKI7* gene has undergone multiple changes that are incompatible with the Hbs1 function and may also have undergone additional changes to optimize mRNA quality control. The combination of transcriptome analysis in diverse eukaryotes and genetic analysis in yeast clarifies the mechanism by which a Ski7-like protein is expressed across eukaryotes and provides a unique view of changes in alternative splicing patterns of one gene over long evolutionary time.

gene duplication | mRNA quality control | HBS1L | alternative splicing

Both gene duplication and alternative splicing have long been thought to contribute to diversification of gene function over evolutionary time, although the details of their contribution are still poorly understood. The yeast genes *SKI7* and *HBS1* provide an example where the current function of the yeast genes is well understood, but their evolutionary diversification has not been fully explored. At the molecular level, Ski7 has been shown to directly bind to both the Ski complex and the RNA exosome and is thought to mediate mRNA delivery from the Ski complex to the RNA exosome, allowing for mRNA degradation by the RNA exosome (1–3). This general function of Ski7 requires only the N-terminal domain of Ski7. The C-terminal domain has a more poorly understood function in the specific recognition of mRNAs that lack a stop codon and mediates their degradation by the RNA exosome (4). The paralog Hbs1 has distinct and nonoverlapping functions and does not bind to either the Ski complex or the RNA exosome. Instead Hbs1 delivers a partner protein, Dom34, to a ribosome when it has stalled during translation. Thereby Hbs1 allows Dom34 to disassemble a stalled ribosome into its 60S and 40S subunits (5–7).

The *SKI7* and *HBS1* genes arose as part of a whole genome duplication (WGD) about 100 Mya, after the *Saccharomyces* lineage diverged from the *Lachancea* lineage (8–11). Although

Ski7 function was initially thought to have arisen after the duplication, we have previously shown that the single *Lachancea kluyveri* ortholog can carry out both Ski7 and Hbs1 functions and thus, that Ski7 function was already present before whole genome duplication (8, 10, 11). In addition, we have shown that the *L. kluyveri* gene encodes two different proteins through alternative splicing with a shorter splice isoform performing the function of *HBS1* and a longer splice isoform performing the function of *SKI7* (10). Furthermore, based on sequence analysis of other fungal species, we concluded that alternative splicing of a single *SKI7/HBS1* gene is conserved in ascomycetes, basidiomycetes, and possibly animals, but a key change in splicing pattern occurred in an early ascomycete (10). More recently, studies have confirmed that the human ortholog (*HBS1L*) is alternatively spliced and suggested that the minor splice isoform resembles Ski7 in its ability to bind to the RNA exosome (1, 12). However, it was suggested that alternative splicing of *HBS1L* was restricted to vertebrates and not found in invertebrate animals (12), implying that alternative splicing may have arisen independently in fungi and vertebrates.

To get a more complete overview of the evolutionary history of alternative splicing of this gene, we have now analyzed transcriptome data from more diverse eukaryotes. Our general approach was to identify *SKI7/HBS1* homologs and analyze the splicing pattern by

Significance

The rapid degradation of mRNAs that lack a stop codon is critical to fidelity of gene expression and in yeast, it requires Ski7. Ski7 function is not fully understood and *SKI7*-like genes are not apparent in other organisms. We show that in most eukaryotes Ski7 is expressed as an alternative splice isoform from the *HBS1* gene. This most conserved example of alternative splicing probably arose in the common ancestor of animals, fungi, and plants. However, in six taxa alternative splicing was replaced by duplicated genes. After each duplication the *SKI7*-like gene has undergone several changes that we analyzed experimentally. The results clarify how duplicated genes diversify, identify novel *SKI7*-like genes, and reveal changes in nonstop mRNA decay.

Author contributions: A.N.M., J.H., M.K., and A.v.H. designed research, performed research, analyzed data, and wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

Published under the PNAS license.

Data deposition: The data reported in this paper have been deposited in the Gene Expression Omnibus (GEO) database, <https://www.ncbi.nlm.nih.gov/geo> (accession no. SRP129038).

¹To whom correspondence should be addressed. Email: ambro.van.hoof@uth.tmc.edu.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1801997115/-DCSupplemental.

Published online July 2, 2018.

RNA sequencing analysis. We detected alternative splicing of *SKI7/HBS1* in many animals and fungi, and thus extended our analysis to species that diverged earlier. We thereby identified alternative splicing in a slime mold and in several plant species but not in any eukaryote that diverged even earlier. Overall, the results suggest that alternative splicing of *SKI7/HBS1* arose in a common ancestor of animals, fungi, and plants but that the alternative splicing pattern changed several times, instead of once at the base of the ascomycetes. Furthermore, we have found that duplication of this gene in multiple lineages was followed by loss of alternative splicing, loss of an N-terminal domain of unknown function, and loss of a catalytic residue in the GTPase domain. We experimentally verify in *Saccharomyces cerevisiae* that each of these changes contributes to the functional divergence of Ski7 from Hbs1. These changes in alternative splicing and gene duplications provide additional insight in the expression of the *SKI7/HBS1* gene in eukaryotes and into how they have affected the Ski7 function.

Results and Discussion

A Single *SKI7/HBS1* Gene Is Alternatively Spliced in Diverse Animals, Fungi, and Plants. Our approach to characterize the splicing pattern of *SKI7/HBS1* orthologs in diverse eukaryotes was to reanalyze transcriptome sequencing data in diverse eukaryotes with available high-quality genome sequences. Preliminary investigations indicated that existing annotation of splice isoforms was often incorrect or incomplete. We therefore narrowed our analysis to species with available raw transcriptome sequencing data and an available high-quality genome sequence. We reanalyzed the primary data by mapping transcriptome sequencing reads to the genome and carefully manually curating splice isoforms of the *SKI7/HBS1* gene. The best transcriptome sequencing-based evidence for (alternative) splicing is the presence of junction reads, which are reads that partially map to one exon and partly to the next exon. To be conservative, we only concluded that a gene was alternatively spliced if both splice isoforms were supported by multiple junction reads.

As a proof of principle, we analyzed RNAseq data for *L. kluyveri* and humans, the two species where alternative splicing of a single *SKI7/HBS1* gene has previously been investigated in detail (10, 12). We generated RNAseq data from *L. kluyveri* and reanalyzed publicly available raw reads from humans. For both species, junction reads clearly confirmed the use of alternative 3' splice sites to generate two different transcripts (Fig. 1 *A* and *B*), suggesting that this method could reliably detect alternative splicing of *SKI7/HBS1*.

In many other diverse eukaryotes, alternative splicing of *SKI7/HBS1* was not previously annotated and/or the annotated splicing patterns were inconsistent with our RNAseq analysis. For example, the *Cryptococcus neoformans* gene was annotated in GenBank with a single splice isoform that contains nine exons (Fig. 1 *C*). We had previously noted, based on EST data, that this was likely incorrect but could not definitively determine the correct splicing pattern based on sequence and EST analysis alone (10). The reanalysis of RNAseq data clearly showed that one of the annotated introns was not supported by any junction reads and seemed to have been annotated to eliminate an in-frame stop codon that interfered with annotation of a single protein being produced from this gene. Transcriptome analysis indicated that this stop codon is an authentic feature and that the exon containing it is an alternatively spliced exon. Specifically, multiple junction reads show that the third exon can be spliced to either the fourth or fifth (previously annotated as the sixth) exon. In contrast, no junction reads support the previously annotated fourth intron. Thus, instead of this gene producing one mRNA and one protein, it encodes two mRNAs, with exon 4 either included or skipped (Fig. 1 *C*). Exon 4 includes a stop codon, which may trigger nonsense-mediated mRNA decay. Consistent with this, we noted that the exon 4-included splice isoform is more

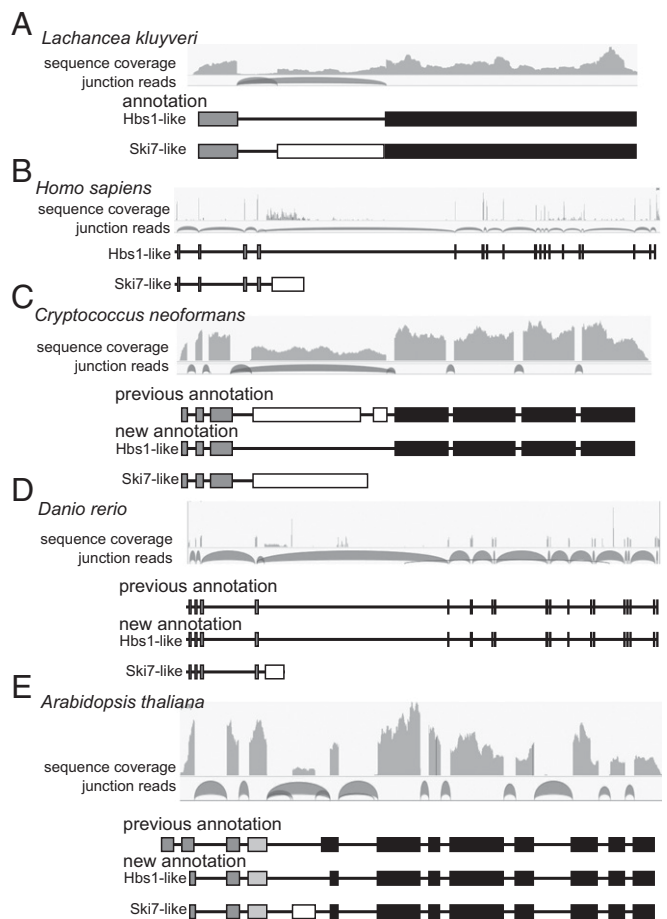


Fig. 1. RNA sequencing analysis reveals alternative splicing patterns of a single *SKI7/HBS1* gene. RNA sequencing results from *L. kluyveri* (*A*), humans (*B*), *C. neoformans* (*C*), zebrafish (*D*), and *A. thaliana* (*E*) are depicted. Shown for each species from *Top* to *Bottom* are the coverage of RNA sequencing reads, identifying exons; junction reads, identifying which exons are spliced to each other; any previous annotation of splicing pattern; and annotation of Hbs1 and Ski7 splicing pattern derived from the RNA sequencing results. Black box shows exon encoding a GTPase domain; dark gray box, exon encoding an Hbs1 N-terminal domain; white box, exon encoding Ski7-specific domain; and light gray box, exon encoding a plant-specific Zn-finger domain.

abundant in RNAseq data from a strain that lacks nonsense-mediated mRNA decay (*upf1Δ*) (13). This suggests that Ski7-mediated mRNA decay may be regulated by nonsense-mediated mRNA decay in *Cryptococcus*. The analyses of *C. neoformans* convinced us that RNAseq reanalysis could readily identify splicing patterns that were not obvious from the gene sequence alone.

Similarly, Fig. 1 *D* and *E* show that the *Danio rerio* (zebrafish) and *Arabidopsis thaliana* orthologs include an alternatively spliced exon that was previously unannotated. The splicing pattern in zebrafish closely resembles the pattern previously shown for humans (compare Fig. 1 *B* and *D*), and our transcriptome analyses showed that this splicing pattern is also found in a diverse set of other animals, including vertebrates, insects, mollusks, sea anemones, and sponges. The same splicing pattern was also detected in *Capsaspora owczarzewski*, a lineage that is related to the animals. We conclude that the human splicing pattern is not restricted to vertebrates as previously suggested (12), but instead is widely conserved in animals.

We also detected strong evidence for alternative splicing of *SKI7/HBS1* in diverse fungal species, including ascomycetes, basidiomycetes, zygomycetes, and chytrids (Fig. 2 and *SI Appendix*,

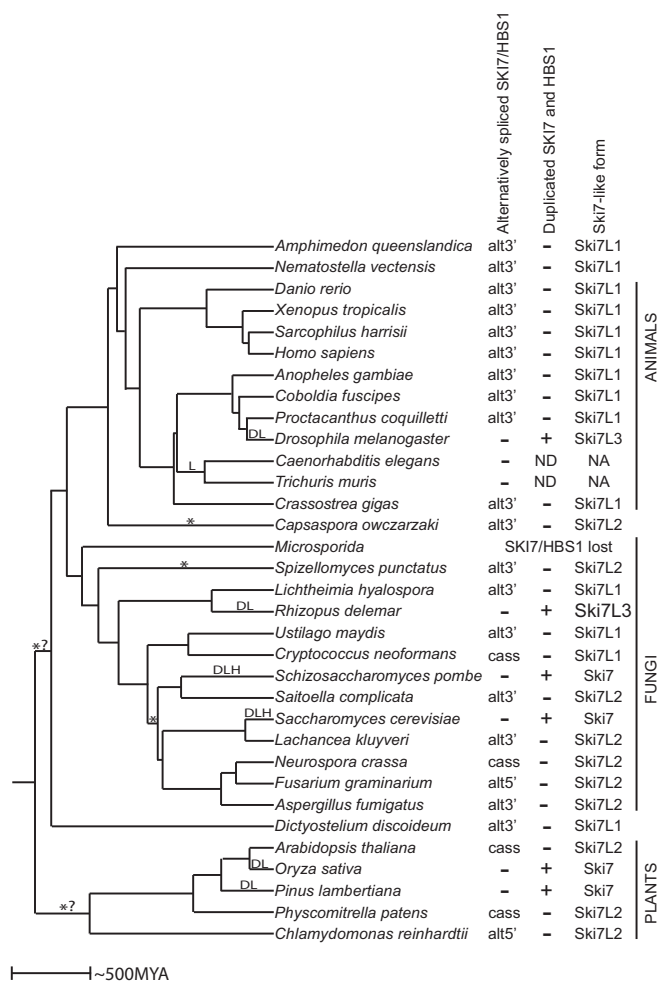


Fig. 2. Alternative splicing of a single *SKI7/HBS1* is conserved in diverse eukaryotes. The tree reflects the species tree, with branch lengths reflecting time of divergence as estimated by www.timetree.org. D indicates a gene duplication event; L, a loss of alternative splicing; *, a switch between a Ski7 that contains or lacks the GTPase domain; H, a loss of the catalytic histidine; and *? indicates that it is not clear whether one change of alternative splicing occurred in the common ancestor of plants or in the common ancestor of animals and fungi. Alternative splicing patterns include the use of alternative 3' splice sites (alt3'), the use of alternative 5' splice sites (alt5'), and the use of a cassette exon (cass). N.A., not applicable; N.D., not detected. Ski7-like protein forms Ski7L1, Ski7L2, and Ski7L3 are defined in Fig. 3E.

Table S1). Most fungi use alternative 3' splice sites, but some ascomycetes use a cassette exon or alternative 5' splice sites (e.g., *Neurospora crassa* and *Fusarium graminearum*, respectively). Thus, alternative splicing of *SKI7/HBS1* occurs throughout the fungi, but fungi use more diverse alternative splicing patterns than animals.

To determine when alternative splicing arose, we further expanded our analysis to include eukaryotes that diverged earlier and discovered alternative splicing in plants and the slime mold *Dictyostelium discoideum* (Fig. 2). As in the ascomycetes, plants used diverse mechanisms of alternative splicing, including alternative 5' splice sites and cassette exons. Even though the splicing pattern varied between all of these animals, fungi, and plants, the alternative splicing always affected an exon that encodes an intrinsically disordered region (white rectangles in figures) between the N-terminal (gray rectangles) and GTPase domains (black rectangles). Even earlier diverging eukaryotes (including *Giardia intestinalis*, *Leishmania donovani*, and *Trichomonas vaginalis* from the kingdom Excavata) encoded a clearly recognizable

SKI7/HBS1 ortholog, but RNAseq analyses indicate these Excavata genes do not contain introns and thus are not alternatively spliced. Similarly, although the *HBS1* ortholog from the Alveolata *Tetrahymena thermophila* and the Stramenopile *Phytophthora infestans* contained introns, RNAseq analysis showed no evidence of alternative splicing.

The splicing pattern in some eukaryotes was less obvious, especially in cases where the genome is large and incompletely sequenced, or the transcriptome data are of poor quality. However, in many cases, a different transcriptome dataset in the same species or a closely related species provided insight into that lineage. This approach provided conclusive evidence for alternative splicing of *SKI7/HBS1* in diverse eukaryotes (Figs. 1 and 2). The most parsimonious explanation of all these observations is that alternative splicing arose relatively early during eukaryotic divergence (at least a billion years ago), before animals, fungi, and plants diverged from each other but possibly after their divergence from other eukaryotes. Although some alternative splicing events have previously been suggested to be conserved from a common animal/fungal ancestor (14), this is unique evidence for alternative splicing of a gene being conserved from an even earlier common ancestor also shared with plants. Thus, the alternative splicing of *SKI7/HBS1* is the most ancient alternative splicing of a specific gene we are aware of.

Multiple Independent Duplications of the *SKI7/HBS1* Gene Were Followed by Loss of Alternative Splicing. Although we identified alternative splicing of a single *SKI7/HBS1* gene in many eukaryotes, eight clades formed exceptions to this pattern. In six of these eight clades, the alternatively spliced gene is replaced by a pair of duplicated genes that have lost the capacity for alternative splicing. We have previously shown that in the *Saccharomyces* and *Schizosaccharomyces* lineages a single alternative spliced *SKI7/HBS1* gene was duplicated, and alternative splicing was subsequently lost in both copies, such that one gene now encodes the Hbs1 protein, while the other gene encodes the Ski7 protein (10). These events occurred independently of each other after their divergence from *Lachancea* and *Saitoella*, respectively (Fig. 2 and ref. 10).

The third and fourth clade that lost alternative splicing of a single *SKI7/HBS1* gene are the *Oryza sativa* (rice) and *Pinus lambertiana* (pine) lineages. Similar to the *Saccharomyces* and *Schizosaccharomyces* cases, we found that in rice and pine the alternatively spliced gene was replaced by duplicate genes that show no evidence of alternative splicing (Fig. 3A and B). Other plants, including *A. thaliana* and the moss *Physcomitrella patens*, contain only one *SKI7/HBS1* gene that according to our RNAseq analyses is alternatively spliced. In both rice and pine, one of the duplicated genes includes an exon with clear sequence similarity to the alternative exon in *Arabidopsis*, while the other gene lacks this exon (Fig. 3 and *SI Appendix*, Fig. S1). We conclude that in the rice and pine lineages a single alternatively spliced *SKI7/HBS1* gene was replaced with pairs of duplicated genes. Because rice and *Arabidopsis* are more closely related to each other than to pine, the simplest explanation is that the duplications in the pine and rice lineages occurred independently.

The fifth clade that lost alternative splicing of a single *SKI7/HBS1* gene contains *Drosophila melanogaster* (Fig. 3C). The *D. melanogaster HBS1* gene lacks introns and is not alternatively spliced. Despite this, no readily recognizable duplicate gene is present. To explore when this loss of alternative splicing occurred, we analyzed related species. Preliminary sequence analysis suggested that flies closely related to *Drosophila* also contained an *HBS1* gene that did not appear to be alternatively spliced (including *Musca*, and *Megaselia* species). However, several mosquito species (including *Aedes*, *Anopheles*, and *Culex* species) seemed to have an additional alternative exon, which was confirmed by transcriptome analysis of *Anopheles gambiae* (Fig. 3C). Specifically,

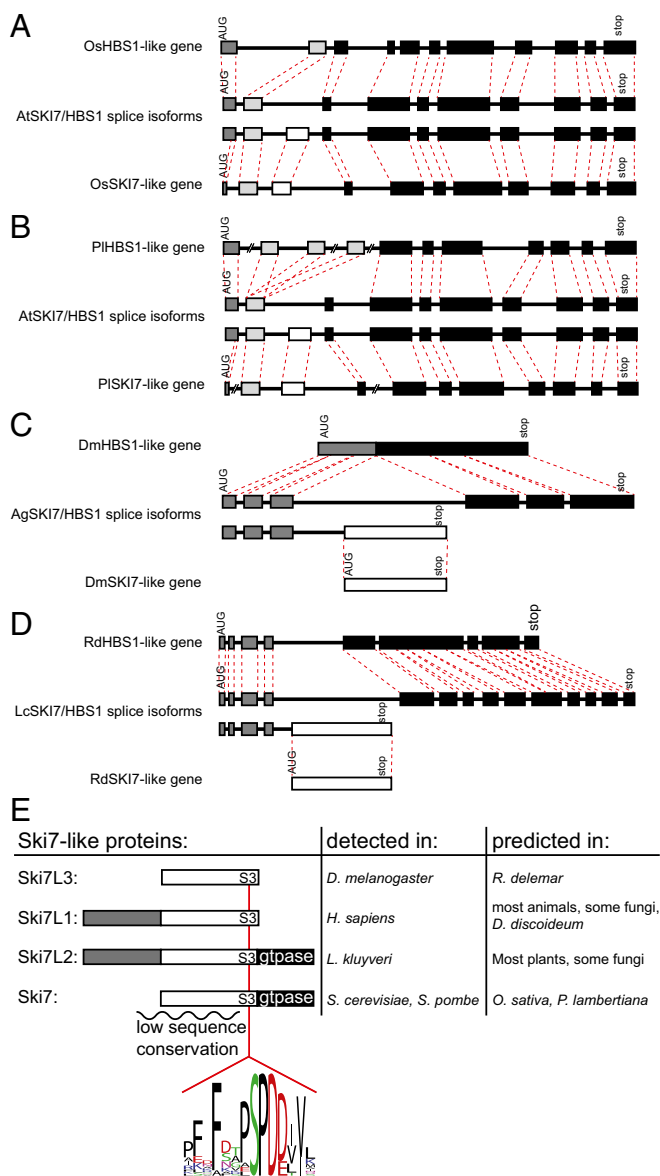


Fig. 3. Animals, fungi, and plant genomes encode one of four different Ski7-like proteins. (A–D) Duplication of SKI7/HBS1 is often followed by loss of alternative splicing and loss of the N-terminal domain. The duplicated genes of rice (Os), pine (PI), fruit flies (Dm), and *Rhizopus* (Rd) are depicted and compared with the alternatively spliced genes of *Arabidopsis* (At), mosquitos (Ag), and *Lichtheimia* (Lc). Coding exons are depicted as in Fig. 1. Red dotted lines connect homologous exon boundaries. The 5' UTR introns are not depicted. Intron and exon sizes are generally drawn to scale within each gene, with the exception of the very large introns of pine. (E) Depiction of the four different Ski7-like proteins in diverse eukaryotes. An intrinsically disordered region of the protein with low sequence conservation is indicated with a squiggly line. S3 indicates a conserved motif shared by all Ski7-like proteins. Black, white, and dark gray rectangles are used as in Fig. 1. The sequence logo depicts the S3 motif in all of the species analyzed. The Ski7-like proteins have been experimentally detected by Western blot and/or mass spectrometry in some organisms (Middle column; refs. 10, 12, 15, 16, and 37–41) and in other organisms they are predicted by our RNAseq analysis (Right column).

multiple junction reads supported splicing of exon 3 to either exon 4 or exon 5. Subsequent analysis also found alternative splicing of a single SKI7/HBS1 gene in the fly species *Proctacanthus coquilletti* and *Coboldia fuscipes*. Therefore, the loss of alternative splicing in the fruit fly lineage occurred within the Muscomorpha infraorder. Comparison of multiple fly and mosquito *HBS1* genes revealed

that the sequence encoded by the *An. gambiae* alternative exon 4 is not present in *Drosophila HBS1* genes. Position-specific iterated (PSI)-BLAST analysis starting with the exon 4-encoded sequence revealed a very rapidly evolving and completely uncharacterized gene in *D. melanogaster* (CG17982) and related fly species (Fig. 3C). Although the encoded protein is uncharacterized, it has been shown to copurify with the RNA exosome (15), like the Ski7 protein from yeast and the minor splice isoform of humans (1–3, 12, 16). Although CG17982 was not previously recognized as a potential Ski7 homolog, it was the only uncharacterized protein in the RNA exosome purification of Forler et al. (15). We conclude that CG17982 is a strong candidate for a Ski7 homolog, although it is highly diverged. This suggests that the alternatively spliced SKI7/HBS1 gene was replaced by two genes that lack alternative splicing in the *Drosophila* lineage.

The sixth clade where a single alternatively spliced SKI7/HBS1 gene is replaced by a pair of duplicated genes contains the mucoromycete fungus *Rhizopus delemar* (Fig. 3D). This duplication and loss of alternative splicing appear to have occurred after the divergence of mucoromycetes because *Lichtheimia corymbifera* does contain a single alternatively spliced SKI7/HBS1 gene.

Unlike the six clades described above that contain duplicate SKI7 and HBS1-like genes, the last two clades that lacked evidence of a single alternatively spliced SKI7/HBS1 gene may completely lack one or both proteins. The seventh clade where we failed to find an alternatively spliced SKI7/HBS1 gene is the Microsporidia. Microsporidia are pathogenic fungi that replicate inside an infected animal host cell and have unusually small genomes that have lost many genes (17). Microsporidia lack any recognizable SKI7 or HBS1 genes, and thus the SKI7/HBS1 gene may have been one of the many genes lost in this lineage. We speculate that the microsporidial Ski2 and RNA exosome retain functionality in the absence of a recognizable Ski7 but conclusively showing this will be challenging in these experimentally intractable organisms.

The eighth clade that appeared to lack an alternatively spliced SKI7/HBS1 gene contains *Caenorhabditis elegans*. We unsuccessfully pursued the same strategy as detailed above for *D. melanogaster* to find a candidate very rapidly diverging duplicate gene. Specifically, we analyzed the transcriptome of the early diverging nematode *Trichuria muris*. This indicated that its *HBS1* gene also encodes one splice isoform. It therefore appears that the loss of SKI7/HBS1 alternative splicing occurred in an ancient ancestor of the nematodes.

The apparent loss of Ski7 in both nematodes and Microsporidia is more ancient than the duplications in the *Drosophila*, *Saccharomyces*, *Schizosaccharomyces*, *Rhizopus*, *Oryza*, and *Pinus* lineages. It is not clear whether this large time span has resulted in a highly diverged and unrecognizable Ski7-like protein in nematodes and Microsporidia or whether Ski7 function has been lost in these taxa. Interestingly, microsporidia and nematode genomes do contain readily recognizable genes for the Ski7 binding partner Ski2 (the most conserved subunit of the Ski complex) and the RNA exosome.

Loss of Alternative Splicing Is Not an Inevitable Consequence of SKI7/HBS1 Duplication. The results above suggest that duplication of an alternatively spliced SKI7/HBS1 gene was followed by loss of alternative splicing in six different lineages (fruit flies, budding yeast, fission yeast, *Rhizopus*, rice, and pine). In contrast, a single alternatively spliced SKI7/HBS1 ortholog was retained through two or three rounds of WGD in the human and zebrafish lineages, respectively. This raises the question of how often a duplicated pair of alternatively spliced genes is resolved into two genes with the complimentary splice patterns. WGD events are especially informative, because they provide evidence of gene duplication that is independent of the presence of the duplicated gene in extant species. We therefore examined the SKI7/HBS1 orthologs in other animal, fungal, and plant lineages that underwent well-documented WGD. In many cases either one or both alternatively spliced genes were retained after WGD.

Xenopus laevis has a duplicated genome relative to *Xenopus tropicalis* (18). Our transcriptome analyses indicate that the *X. tropicalis* *SKI7/HBS1* gene is alternatively spliced, using the same pattern as all other animals studied (Fig. 2). *X. laevis* contains two recognizable *SKI7/HBS1* genes on chromosomes 5S and 5L (the paralogous chromosomes resulting from WGD). Although the *X. laevis* genome sequence contains too many gaps to conclusively identify splicing patterns and be included in Fig. 2, the chromosome 5L *SKI7/HBS1* gene appears to be intact and capable of alternative splicing. In contrast, the chromosome 5S gene appears to have acquired frame shift mutations in eight different exons, and three exons appear to be deleted. Thus, the gene on 5S is likely to be a pseudogene that is in the process of being lost from the genome. Therefore, after WGD, *X. laevis* maintained one alternatively spliced *SKI7/HBS1* gene, while the other gene degenerated to a pseudogene.

WGD is much more common in plants and has been followed by various fates for the *SKI7/HBS1* gene. For example, the soybean genome (*Glycine max*) underwent WGD after its divergence from pigeon pea (*Cajanus cajan*) (19) and appears to contain two *SKI7/HBS1* genes that are alternatively spliced. Although WGD is rare in fungi, the ascomycete *Hortaea werneckii* lineage underwent a recent WGD (20) and has also retained two *SKI7/HBS1* genes that both appear to be alternatively spliced. This suggests that even though gene duplication has led to loss of alternative splicing of *SKI7/HBS1* in six different lineages, other duplications have not (yet) affected alternative splicing.

To complement the above study of *SKI7/HBS1* fate after multiple independent duplications, we analyzed our *L. kluyveri* transcriptome data for evidence of alternative splicing in other genes and compared those to the post-WGD genes in *S. cerevisiae*. Consistent with alternative splicing being rare in the ascomycetes, we detected only two convincing examples of the use of alternative splice sites in addition to *SKI7/HBS1*. The *L. kluyveri* *SAKL0A08096* gene uses alternative 3' splice sites (SI Appendix, Fig. S2). In *S. cerevisiae*, both copies of this gene (*YSH1* and *SYC1*) have been retained and have lost their intron (and thus the capacity for alternative splicing). The *S. cerevisiae* Ysh1 protein closely resembles the *L. kluyveri* long splice isoform containing an N-terminal endoribonuclease domain and a C-terminal protein interaction domain (SI Appendix, Fig. S2 A and D). The *S. cerevisiae* Syc1 protein on the other hand closely resembles the short splice isoform with only the protein interaction domain (SI Appendix, Fig. S2D). The third *L. kluyveri* gene with evidence for alternative splice sites is *SAKL0G06116*. This gene uses alternative 5' splice sites that are separated from each other by four nucleotides. In *S. cerevisiae*, one ortholog of this gene (*SRC1*) has retained this alternative splicing event, while the other ortholog (*HEH2*) has lost the intron (21). Thus, the common ancestor of *Saccharomyces* and *Lachancea* likely used alternative splice sites for three of its genes. WGD replaced all three of these genes with pairs of duplicated genes and five of the six duplicated genes lost alternative splicing.

Overall, these analyses suggest that duplication of the *SKI7/HBS1* gene and other alternatively spliced genes often results in loss of alternative splicing, although this outcome is not obligate.

Changes in Splicing Pattern Repeatedly Diversify the Ski7-Like Proteins. To get a more detailed understanding of the evolution of the *SKI7/HBS1* function, we also analyzed the encoded protein sequences. The *S. cerevisiae* Hbs1 protein consists of a small N-terminal domain and a C-terminal GTPase domain (6). In all animals and fungi analyzed, one of the splice isoforms (or duplicated genes) had this same domain organization. Although sequence conservation of the N-terminal domain is low, it is rich in acidic and tyrosine residues (SI Appendix, Fig. S3). This splice isoform is also similar to the orthologous sequences from the Alveolata, Stramenopile, and Excavata clades (SI Appendix, Fig.

S3). The land plants gained a novel exon after they diverged from *Chlamydomonas reinhardtii*. This exon is included in both their Ski7s and Hbs1s and encodes a putative Zn finger (Fig. 1E, light gray boxes; SI Appendix, Fig. S1). Thus, the Hbs1-like protein from diverse eukaryotes is well conserved from the ancient common ancestor of plants, animals, fungi, Alveolata, Stramenopile, and Excavata.

In contrast, the Ski7-like splice isoform was much more diverse and not readily aligned by multiple sequence alignment. The Ski7 proteins from the *Saccharomyces* genus are distinguishable from their Hbs1 paralogs by three short conserved motifs: S1, S2, and S3 (10). While the S1 and S2 motifs are not obviously conserved beyond the Saccharomycetacea, the S3 motif is recognizable in the human exon that is specific for the Ski7-like splice isoform (10). We also detected a similar motif in all other plants, animals, and fungi analyzed (with the exception of the nematodes and Microsporidia as discussed above; Fig. 3E). A recent crystal structure indicates that the S3 motif forms part of the RNA exosome interaction (1), suggesting that this interaction is conserved across eukaryotes and distinguishes Ski7 function from Hbs1 function.

Although the S3 motif is conserved across eukaryotes, the domain organization of the Ski7-like protein is poorly conserved (Fig. 3E). Specifically, in the human Ski7-like splice isoform, the S3 motif is encoded in the alternative exon. In other animals, the S3 motif is also encoded in the alternative exon and followed from 9 to 35 codons later by the stop codon. We refer to these Ski7-like proteins as Ski7L1 (Fig. 3E and SI Appendix, Fig. S4). Some fungi also contain a Ski7L1 protein, with the conserved S3 motif encoded in the alternative exon and closely followed by a stop codon (e.g., *Cryptococcus* and *Lichtheimia*; Figs. 1C and 3E and SI Appendix, Fig. S4). In contrast, in the Ski7-like splice isoform of *L. kluyveri* the S3 motif is followed by a GTPase domain shared with the Hbs1-like isoform. We refer to this form as Ski7L2 (Fig. 3E and SI Appendix, Fig. S5). Among the species examined, four clades have the Ski7L2 form (plants, ascomycetes, *Spizellomyces punctatus* and *C. owczarzaki*). This requires at least four switches between a Ski7L1 and a Ski7L2 splicing pattern gene (Fig. 2, indicated with *). Because the GTPase domain of Ski7 is known to be important for nonstop mRNA decay in *S. cerevisiae*, species that contain a Ski7L1 may use a different mechanism to degrade nonstop mRNAs or lack a specific mechanism to recognize these mRNAs.

Duplication of *SKI7/HBS1* Further Diversifies the Ski7-Like Proteins.

Duplication and loss of alternative splicing resulted in two different domain organizations not found in the nonduplicated *SKI7/HBS1* genes. In *S. cerevisiae*, *S. pombe*, *O. sativa*, *D. melanogaster*, and *R. delemar*, the N-terminal domain was lost from the Ski7-like isoform. In the first three species, the preduplication gene encoded a Ski7L2 form, and thus the postduplication form closely resembles Ski7 in that it has the Ski7 specific domain and a C-terminal GTPase domain (Fig. 3E). Since this domain organization most closely resembles the *S. cerevisiae* protein, we refer to it simply as Ski7. In contrast, the *D. melanogaster* and *R. delemar* Ski7-like genes evolved from a preduplication Ski7L1 that lacks a GTPase domain, and therefore, the postduplication genes also lack the GTPase domain and consist solely of a domain with low sequence conservation (Fig. 3E). This low sequence conservation explains why the fly protein was not previously annotated as having any similarity to Ski7. The structure of the *Drosophila* and *Rhizopus* proteins is thus distinct in that it lacks both the N-terminal Hbs1 domain and the GTPase domain, and we therefore called it Ski7L3. Thus, for five of the six duplications the Ski7-like gene lacks the N-terminal domain of the ancestral protein. In the sixth case, pine, the N-terminal domain is present in the Ski7-like gene but has diverged in sequence to a point where it may no longer be functional (SI Appendix, Fig. S1). Therefore, duplication

of *SKI7/HBS1* is often followed by loss (or divergence) of the N-terminal domain, which further diversifies Ski7-like isoforms.

Although the postduplication *Saccharomyces* and *Schizosaccharomyces* Ski7 proteins retain a recognizable GTPase domain, they both diverged from the consensus GTPase sequence. Most notably, the histidine that catalyzes GTP hydrolysis is replaced by a serine in *Saccharomyces* species and unalignable in the *Schizosaccharomyces* species (10, 22). These proteins are thus likely pseudo-GTPases. In contrast, the duplicated rice and pine Ski7s retain all conserved GTPase motifs. The analyses above suggest that while *S. cerevisiae* Hbs1 closely resembles the preduplication protein, *S. cerevisiae* Ski7 has undergone three major changes since its duplication: loss of alternative splicing, loss of the N-terminal domain, and divergence of the GTPase sequence to a pseudo-GTPase.

Three Major Postduplication Changes of Ski7 Are Incompatible with Hbs1 Function and Explain Its Specialization. We next used experimental genetics in *S. cerevisiae* to understand the functional consequences of the three major changes identified by RNAseq and protein sequence comparisons. We have previously shown experimentally that loss of alternative splicing may have contributed to gene specialization (10); however, the consequences of loss of the N-terminal domain and GTPase activity have not been experimentally addressed. To investigate the functional consequences of these three changes, we introduced them individually into the *L. kluyveri* gene.

We began by deleting the N-terminal domain from *L. kluyveri* *SKI7/HBS1*. Specifically, the N-terminal domain in *L. kluyveri* is encoded within the first exon, and we therefore deleted the sequences between the start codon and the Ski7-specific exon (*L. kluyveri*- Δ exon1). As previously reported, the short *L. kluyveri* Hbs1-like splice isoform fully complements the growth defects of an *rps30a* Δ *hbs1* Δ strain. We previously reported that the Ski7-like long splice isoform did not fully complement this growth defect (10), but on closer examination, we noticed partial complementation. Importantly, the *L. kluyveri*- Δ exon1 construct was completely nonfunctional in complementing *rps30a* Δ *hbs1* Δ (Fig. 4A). In contrast, this construct is fully capable of complementing a *dcp1-2 ski7* Δ strain, and thus encodes a protein that can function as Ski7 but cannot function as Hbs1. This suggests that loss of alternative splicing in the *SKI7* gene after duplication was not sufficient for complete loss of Hbs1 function and that other changes such as loss of the N-terminal domain contributed.

Translation factor GTPases, including eIF1A, eRF3, and Hbs1 contain five conserved motifs (G1–G5), and each of these motifs is readily recognizable in the preduplication *L. kluyveri* *SKI7/HBS1* gene (Fig. 4B). After duplication in the *S. cerevisiae* lineage, each of these motifs was strictly conserved in Hbs1, but poorly conserved in Ski7 (Fig. 4B). Strikingly, the catalytic histidine of motif G3 is conserved in ScHbs1 but changed to a serine in ScSki7. Consistent with this change, ScSki7 has very low if any GTPase activity (22). We therefore replaced the catalytic histidine in LkSki7/Hbs1 with either a serine to mimic the ScSki7 or with an alanine to completely eliminate reactivity of the side chain (mutants G3b and G3a, respectively). GTPase activity of translation factor GTPases is tightly controlled by conformational changes; the ribosome promotes the rotation of the histidine into the active site and a hydrophobic gate reduces this movement (23). A highly conserved valine in the G1 motif forms the hydrophobic gate, but this valine is not conserved in Ski7 in the *Saccharomyces* genus. We changed this valine to glycine to impair the hydrophobic gate (mutant G1a). One of the few conserved residues in the GTPase motifs that are conserved in *Saccharomyces* Ski7 is a lysine in motif G1. We changed this lysine to alanine in LkSki7/Hbs1 (mutant G1b).

The G3a and G3b point mutants that change the catalytic histidine were unable to complement the *rps30a* Δ , *hbs1* Δ growth phenotype (Fig. 4B). In contrast, the G1a mutant that targeted

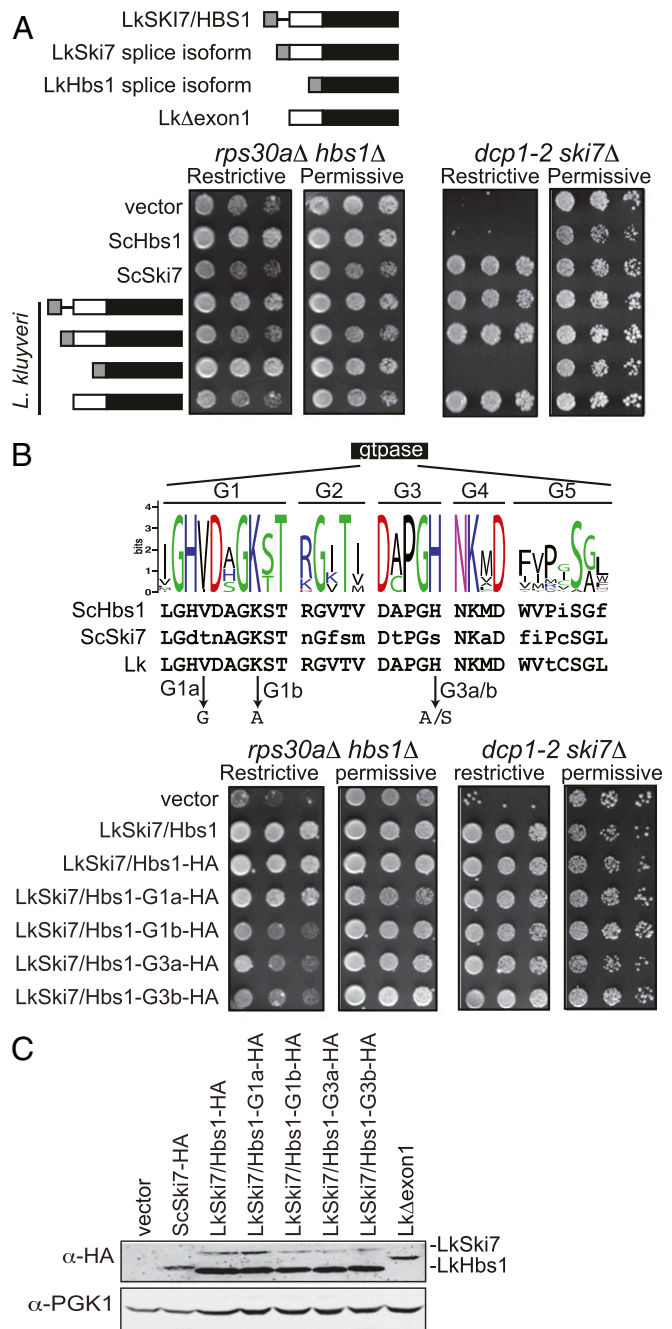


Fig. 4. The GTPase activity and N-terminal domain encoded by *LkSKI7/HBS1* are required for its Hbs1 function, but dispensable for its Ski7 function. (A, Top) Schematic showing the constructs used to express different versions of *LkSKI7/HBS1*. (Bottom) Serially diluted samples of the indicated strains were spotted onto SC -Ura solid media and grown at permissive and restrictive temperatures. The *Bottommost* construct that is missing its N-terminal domain (gray rectangle) fails to function as Hbs1, but is fully functional for Ski7 function. (B, Top) Sequence logo showing the consensus GTPase motifs for translation factor GTPases along with the sequences found in the LkSki7/Hbs1 proteins, ScHbs1 and ScSki7 (lowercase: mutations since the common ancestor of *Lachancea* and *Saccharomyces*) and the sequences for four mutations in the G1 and G3 motifs. (Bottom) Growth assay demonstrating that three of four GTPase mutants fail in Hbs1 function, but all are fully functional for Ski7 function. The growth assay was performed as in A. (C) The mutations have no effect on the expression level. Ski7-like and Hbs1-like proteins were detected using an anti-HA antibody. Pgk1 is shown as a loading control.

the hydrophobic gate was able to complement. Importantly, all three of these mutants fully complemented the growth phenotype of the *dcp1-2*, *ski7Δ* strain. Finally, although the lysine of the G1 motif is conserved in both ScSki7 and ScHbs1, mutating it in G1b impaired the Hbs1 function of LkSki7/HBS1 but had no effect on its Ski7 function. This suggests that many amino acid changes in *LkSKI7HBS1* are compatible with Ski7 function but not with Hbs1 function. Together these results suggest that the loss of alternative splicing, loss of the N-terminal domain, and nonsynonymous mutations in the GTPase domain likely played important roles in the specialization of Ski7 following duplication.

The Role of the GTPase Domain of Ski7 in Nonstop mRNA Degradation.

Ski7 has both a general role in exosome-mediated cytoplasmic RNA decay in *S. cerevisiae*, and a specific, less-understood role in the degradation of mRNAs that lack an in-frame stop codon (nonstop mRNA decay) (3, 16). The growth phenotype we assayed above reflects the role in general mRNA decay and does not address a role in nonstop decay. Previous work demonstrated that the nonstop decay pathway requires the GTPase domain of Ski7 in *S. cerevisiae* (16). However, this GTPase domain is not present in Ski7L1 forms of many other eukaryotes, and as the mechanism of nonstop decay in other eukaryotes is poorly understood, we wanted to determine if and how this function changed following duplication.

We used two assays to assess nonstop mRNA decay. In one assay, we used a *his3-nonstop* reporter. The *his3-nonstop* mRNA is rapidly degraded by the nonstop decay pathway, resulting in a low level of the His3 enzyme that is required for histidine biosynthesis. Thus, a *ski7Δ his3-nonstop* strain grows in the absence of added histidine and this effect can be complemented by an *ScSKI7* plasmid. The *LkSKI7HBS1* gene complemented *ski7Δ* and reduced the growth in the absence of histidine (Fig. 5A). To determine whether this activity was affected by any of the three major postduplication changes described above (loss of alternative splicing, loss of the N-terminal domain, and nonsynonymous mutations in the GTPase domain), we examined the effect of these changes on this complementation. The mutant *LkSKI7/HBS1* genes functioned as well as the wild-type gene in this assay. However, on careful examination, we noted that each of the *LkSKI7/HBS1* complemented strains grew slightly better than *SKI7* (or *ski7Δ* complemented with *ScSKI7*), suggesting that *LkSKI7/HBS1* may not fully complement the nonstop mRNA decay phenotype.

Because the *his3-nonstop* growth assay is hard to quantitate and only indirectly reflects mRNA decay activity, we next directly examined the effect on *pgk1-nonstop* mRNA. The *pgk1-nonstop* reporter mRNA is a variant of the *S. cerevisiae* *PGK1* gene from which all in-frame stop codons have been removed (4). Previous work has shown that the *pgk1-nonstop* mRNA is rapidly degraded by the RNA exosome and that the mRNA is stabilized in a *ski7Δ* mutant (4, 16). Consistent with previous results, our experiment showed that this accumulation of the *pgk1-nonstop* mRNA is rescued by the addition of the *ScSKI7* gene (Fig. 5B). In contrast, the *LkSKI7/HBS1* gene consistently resulted in a lower *pgk1-nonstop* mRNA level than the empty vector control but the effect was small and not statistically significant (average level = 0.82 of the vector control, $n = 6$, $P = 0.1$ by the Student's *t* test; Fig. 5B and C). As shown in Fig. 5C, all of the mutant *LkSKI7/HBS1* genes functioned similarly to the starting *LkSKI7HBS1* gene in this assay. Thus, the *pgk1-nonstop* assay indicates that the *LkSKI7HBS1* gene functions suboptimally in nonstop decay. Interestingly, the major changes that affected loss of Hbs1 activity are distinct from the changes that affect the nonstop decay function. This indicates that selection for optimal nonstop mRNA decay did not drive the major postduplication changes in Ski7, but that instead other so far unknown sequence changes affect nonstop mRNA decay.

The above results indicated that Ski7 affects nonstop mRNAs without a need for an intact GTPase active site. Interestingly, the biochemical role of Hbs1 is to use GTP hydrolysis for the

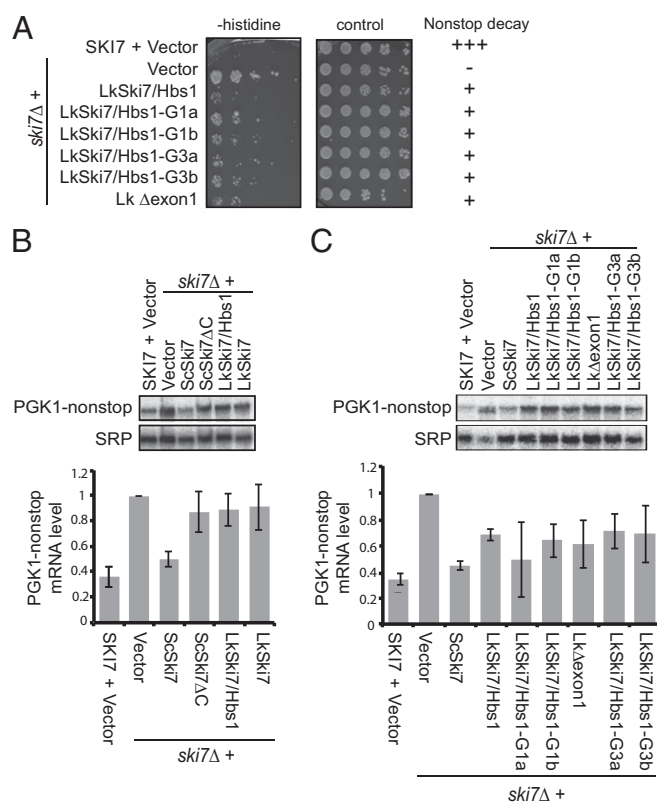


Fig. 5. LkSki7 performs nonstop mRNA decay suboptimally, and sequence changes that disrupt Hbs1 function do not affect nonstop mRNA decay. (A) *His3-nonstop* growth assay demonstrating partial complementation by the *L. kluyveri* gene. Strains containing a *his3-nonstop* reporter were serially diluted and spotted onto media lacking (Left) or containing (Right) histidine. In cells with a wild-type *SKI7* gene (Top row) the *his3-nonstop* mRNA is rapidly degraded, limiting the synthesis of histidine and therefore growth in the absence of histidine. In the *ski7Δ* strain (second row), a defect in nonstop mRNA decay results in growth. (B and C) Northern blot analyses of the degradation of a *pgk1-nonstop* mRNA. In both panels a representative blot is shown at Top, with quantitation and normalization to SRP RNA for three biological replicates shown at the Bottom. The bar graphs indicate average and SD.

disassembly of stalled ribosomes into 40S and 60S subunits for reuse in a variety of biological contexts (5, 6, 12). For example, if a ribosome is stalled within the coding region, Hbs1, together with Dom34 and Rli1, disassembles the ribosome, resulting in a 60S subunit that still contains the nascent peptide (7, 24). The nascent peptide is subsequently ubiquitinated by Ltn1 and degraded by the proteasome (25). Therefore, Hbs1 is required to generate the Ltn1 substrate. Proteins encoded by nonstop mRNAs are also targeted by Ltn1 and the proteasome (26). To test whether Ski7, in analogy to Hbs1, is required to generate the substrate for Ltn1, we analyzed the effect of double mutants. As we have previously reported, both *ski7Δ* and *ltn1Δ* result in increased expression of the *his3-nonstop* reporter (26) (Fig. 6A). Similarly, both *ski7Δ* and *ltn1Δ* result in increased expression of Pkg1-nonstop protein (Fig. 6B). Importantly, the *ski7Δ*, *ltn1Δ* double mutant showed a stronger effect than either single mutant (Fig. 6A and B). The observation that *ltn1Δ* has an effect even in the *ski7Δ* background indicates that Ski7 is not required to generate the Ltn1 substrate. The substrate of Ltn1 consists of the nascent peptide that is still associated with the 60S subunit but dissociated from the 40S subunit (27, 28). Thus, the additive effect of Ski7 and Ltn1 indicates that Ski7 is not required for the disassembly of ribosomes stalled on a nonstop mRNA.

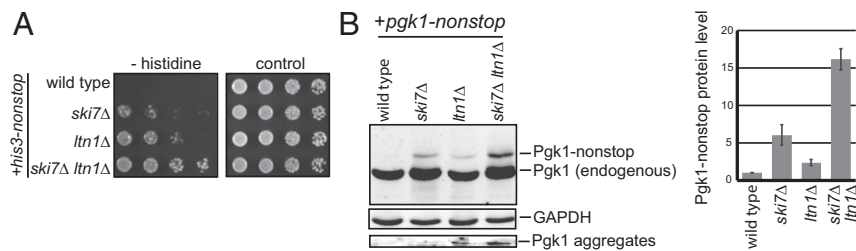


Fig. 6. Ski7 and Ltn1 function independently of each other to limit accumulation of proteins encoded by nonstop mRNAs. (A) A *his3-nonstop* growth assay as in Fig. 5 shows that *ski7Δ* and *ltn1Δ* both affect His3-nonstop expression and that the effects are independent of each other. (B) Western blot analysis shows that *ski7Δ* and *ltn1Δ* both affect Pgk1-nonstop expression and that the effects are independent of each other. The results of three biological replicates were normalized to GAPDH levels and quantitated as in Fig. 5. The *Bottom* blot indicates SDS-resistant amyloid-like aggregates that accumulate in the *ltn1Δ* and *ski7Δ*, *ltn1Δ* strains that migrate very slowly in SDS/PAGE gels.

In addition to ubiquitination by Ltn1, Hbs1-mediated ribosome disassembly also results in C-terminally alanine and threonine tailing (CAT-tailing) of the nascent peptide by Rqc2. Like Ltn1, Rqc2 recognizes the nascent peptide/60S complex (29). If Ltn1-mediated rapid degradation is prevented, CAT-tailed proteins aggregate into amyloid-like aggregates that are resistant to SDS and can be detected as very slow migrating species on SDS/PAGE gels (30, 31). As expected, we detected aggregated Pgk1-nonstop proteins in the *ltn1Δ* strain. Importantly, these aggregates were also seen in the *ski7Δ*, *ltn1Δ* double mutant (Fig. 6B), indicating that Ski7 is not required to generate the Rqc2 substrate. Together these results indicate that Ski7 is not required for ribosome disassembly and that the role of Ski7 in nonstop decay is thus distinct from the known biochemical function of Hbs1 in disassembling stalled ribosomes. This observation explains why Ski7 was able to lose its GTPase activity and Dom34 interaction after gene duplication.

Conclusions

In this study, we show that expression of an Hbs1-like and Ski7-like protein from a single gene through alternative splicing is remarkably well conserved from a common ancestor of animals, fungi, and plants. However, despite this conservation, the Ski7-like protein has changed several times through changes in splicing pattern, such that some eukaryotes express a GTPase domain containing Ski7-like protein, while in other eukaryotes the GTPase domain is not included in the Ski7-like protein. Because this domain is required for nonstop mRNA degradation in yeast, other eukaryotes likely use a different mechanism to degrade these aberrant mRNAs. Although many eukaryotes have a single alternatively spliced *SKI7/HBS1* gene, in multiple lineages duplication of this one gene was followed by loss of alternative splicing, loss of the N-terminal domain, and/or loss of the catalytic histidine. Remarkably, in *Drosophila* and *Rhizopus* this results in pairs of duplicated genes (*Ski7L3* and *Hbs1*) that share no common domains or sequences. Experimental genetic analysis indicates that the features lost after duplication are required for Hbs1 function, but not Ski7 function, explaining why they can repeatedly be lost after gene duplication. In addition, these results indicate that while the GTPase domain is retained in Ski7, its function in nonstop mRNA decay is unlike the GTP hydrolysis-dependent ribosome disassembly function of Hbs1.

Materials and Methods

Transcriptome Analysis. *L. kluyveri* transcriptome sequencing used standard methods. Briefly, strain FM479 [a kind gift of Mark Johnston, University of Colorado, Denver, described previously (11)] was grown in YPD media at 30 °C and RNA was isolated by a hot phenol method (32). Poly(A)⁺ RNA isolation and transcriptome sequencing was performed by LScience which resulted in 47.5 million pairs of 100-nt paired reads. Raw data have been deposited in the GEO database under accession no. SRP129038. Reads were mapped by TopHat (33) to the genome sequence (downloaded from [https://](https://www.ncbi.nlm.nih.gov/bioproject/1445)

www.ncbi.nlm.nih.gov/bioproject/1445) with default parameters except that intron sizes from 30 to 5,000 nt were allowed. All identified introns were inspected manually using the Integrative Genomics Viewer (34).

For other species, raw RNAseq data (*SI Appendix, Table S1*) were downloaded from the European Nucleotide Archive (<https://www.ebi.ac.uk/ena/>), and the genome sequence was downloaded from the National Center for Biotechnology Information (NCBI) (<https://www.ncbi.nlm.nih.gov/>). *SKI7/HBS1* orthologs were identified using BLAST, generally using the Hbs1 sequence from a related species as a query. Transcriptome sequencing reads were mapped with TopHat or HISAT, and the splicing pattern of *SKI7/HBS1* was deduced from sequence coverage and exon junction reads by manual inspection using the Integrative Genomics Viewer (34). The minimum and maximum intron size allowed by TopHat was adjusted to be appropriate for the species. The species included in Fig. 2 produced convincing data with alternative splicing clearly supported by both sequence coverage and junction reads. For some other species, the splicing pattern could not convincingly be deduced, generally because of low-quality or low-depth transcriptome sequence data or gaps in the available genome sequence.

To explore whether the Ski7-like splice isoform of *C. neoformans* was subject to nonsense-mediated decay, duplicate wild-type and *upf1Δ* RNAseq datasets were analyzed by the Tuxedo pipeline, which indicated that the ski7-like splice isoform was up-regulated 1.8-fold in *upf1Δ* (P value = 0.0061).

Yeast Strains and Plasmids. Yeast strains yAv313 (*dcp1-2*, *ski7Δ*), yAv877 (*rps30AΔ*, *hbs1Δ*), yAv987 (*ski7Δ*), and yAv709 (*ltn1Δ*) have been described previously (26, 35, 36). yAv712 (*ski7Δ*) was obtained from Open Biosystems. yAv1777 (*ski7Δ*, *ltn1Δ*) was generated by crossing yAv987 and yAv709. Plasmids were generated using standard methods. Site-directed mutagenesis was performed with the QuikChange Lightning Mutagenesis kit (Agilent). All plasmids were verified by sequencing.

Growth Assays. Cells of the indicated strain and harboring the indicated plasmids were serially diluted (fivefold) and spotted onto SC –Leu, SC –His –Leu, SC –Leu –Ura, and/or SC –Leu –His –Ura (<https://sunrisescience.com/>) and grown at the indicated temperatures.

Northern Blotting and Quantification. *ski7Δ* cells with the indicated plasmids were grown at 30 °C in SC –Leu –Ura media with 2% galactose to induce expression of the *pgk1-nonstop* reporter gene. Total RNA was isolated, separated using formaldehyde-agarose gel electrophoresis, and then blotted onto a Zeta-Probe GT nylon membrane (Bio-Rad). Blots were probed using the ³²P-radiolabeled oligonucleotide probe specific to *pgk1-nonstop* (ATATTGATTAGATCAGGAATTCC) and the RNA subunit of the Signal Recognition Particle (*SCR1*; GTCTAGCCGCGAGGAAGG). Blots were imaged using the Typhoon FLA 7000 phosphor-imaging system (GE Healthcare) and quantification was performed using the ImageQuant NT program (GE Healthcare). A Student's *t* test was used to determine statistical significance between normalized mRNA levels in each strain.

Western Blotting. Protein was isolated from cells growing exponentially at 30 °C using bead vortexing. Western blot analysis was performed using SDS/PAGE and transferring to a nitrocellulose membrane (GE Healthcare). Proteins were visualized using anti-HA (Sigma), anti-PGK1 (Invitrogen), or anti-GAPDH (Sigma) antibodies. Blots were imaged using the ImageQuant LAS4000 mini (GE Healthcare) and quantification was performed using the ImageQuant TL program (GE Healthcare).

ACKNOWLEDGMENTS. We thank Fernando Andrade for performing the RT-PCR analysis of the *Lachancea kluyveri* *YSH1/SYC1* gene and members of the A.v.H. laboratory for critical discussions. This work was supported by NIH R01

Grant GM099790 (to A.v.H.) and a Marilyn and Frederick R. Lummis Jr., M.D. Fellowship in Biomedical Sciences from the Graduate School of Biomedical Sciences (to A.N.M.).

1. Kowalinski E, et al. (2016) Structure of a cytoplasmic 11-subunit RNA exosome complex. *Mol Cell* 63:125–134.
2. Araki Y, et al. (2001) Ski7p G protein interacts with the exosome and the Ski complex for 3'-to-5' mRNA decay in yeast. *EMBO J* 20:4684–4693.
3. van Hoof A, Staples RR, Baker RE, Parker R (2000) Function of the ski4p (Csl4p) and Ski7p proteins in 3'-to-5' degradation of mRNA. *Mol Cell Biol* 20:8230–8243.
4. Frischmeyer PA, et al. (2002) An mRNA surveillance mechanism that eliminates transcripts lacking termination codons. *Science* 295:2258–2261.
5. van den Elzen AM, et al. (2010) Dissection of Dom34-Hbs1 reveals independent functions in two RNA quality control pathways. *Nat Struct Mol Biol* 17:1446–1452.
6. Becker T, et al. (2011) Structure of the no-go mRNA decay complex Dom34-Hbs1 bound to a stalled 80S ribosome. *Nat Struct Mol Biol* 18:715–720.
7. Shoemaker CJ, Eylar DE, Green R (2010) Dom34:Hbs1 promotes subunit dissociation and peptidyl-tRNA drop-off to initiate no-go decay. *Science* 330:369–372.
8. Kellis M, Birren BW, Lander ES (2004) Proof and evolutionary analysis of ancient genome duplication in the yeast *Saccharomyces cerevisiae*. *Nature* 428:617–624.
9. Langkjaer RB, Cliften PF, Johnston M, Piskur J (2003) Yeast genome duplication was followed by asynchronous differentiation of duplicated genes. *Nature* 421:848–852.
10. Marshall AN, Montealegre MC, Jiménez-López C, Lorenz MC, van Hoof A (2013) Alternative splicing and subfunctionalization generates functional diversity in fungal proteomes. *PLoS Genet* 9:e1003376.
11. van Hoof A (2005) Conserved functions of yeast genes support the duplication, degeneration and complementation model for gene duplication. *Genetics* 171:1455–1461.
12. Kalisiak K, et al. (2017) A short splicing isoform of HBS1L links the cytoplasmic exosome and SKI complexes in humans. *Nucleic Acids Res* 45:2068–2080.
13. Gonzalez-Hilarion S, et al. (2016) Intron retention-dependent gene regulation in *Cryptococcus neoformans*. *Sci Rep* 6:32252.
14. Awan AR, Manfredi A, Pleiss JA (2013) Lariat sequencing in a unicellular yeast identifies regulated alternative splicing of exons that are evolutionarily conserved with humans. *Proc Natl Acad Sci USA* 110:12762–12767.
15. Forler D, et al. (2003) An efficient protein complex purification method for functional proteomics in higher eukaryotes. *Nat Biotechnol* 21:89–92.
16. van Hoof A, Frischmeyer PA, Dietz HC, Parker R (2002) Exosome-mediated recognition and degradation of mRNAs lacking a termination codon. *Science* 295:2262–2264.
17. Jaillon O, et al. (2004) Genome duplication in the teleost fish *Tetraodon nigroviridis* reveals the early vertebrate proto-karyotype. *Nature* 431:946–957.
18. Session AM, et al. (2016) Genome evolution in the allotetraploid frog *Xenopus laevis*. *Nature* 538:336–343.
19. Li Z, et al. (2016) Gene duplicability of core genes is highly consistent across all angiosperms. *Plant Cell* 28:326–344.
20. Sinha S, et al. (2017) Insight into the recent genome duplication of the halophilic yeast *Hortaea werneckii*: Combining an improved genome with gene expression and chromatin structure. *G3 (Bethesda)* 7:2015–2022.
21. Rodríguez-Navarro S, Igual JC, Pérez-Ortín JE (2002) SRC1: An intron-containing yeast gene involved in sister chromatid segregation. *Yeast* 19:43–54.
22. Kowalinski E, Schuller A, Green R, Conti E (2015) *Saccharomyces cerevisiae* Ski7 is a GTP-binding protein adopting the characteristic conformation of active translational GTPases. *Structure* 23:1336–1343.
23. Voorhees RM, Schmeing TM, Kelley AC, Ramakrishnan V (2010) The mechanism for activation of GTP hydrolysis on the ribosome. *Science* 330:835–838.
24. Pisareva VP, Skabkin MA, Hellen CU, Pestova TV, Pisarev AV (2011) Dissociation by Pelota, Hbs1 and ABCE1 of mammalian vacuolar 80S ribosomes and stalled elongation complexes. *EMBO J* 30:1804–1817.
25. Brandman O, et al. (2012) A ribosome-bound quality control complex triggers degradation of nascent peptides and signals translation stress. *Cell* 151:1042–1054.
26. Wilson MA, Meaux S, van Hoof A (2007) A genomic screen in yeast reveals novel aspects of nonstop mRNA metabolism. *Genetics* 177:773–784.
27. Bengtson MH, Joazeiro CA (2010) Role of a ribosome-associated E3 ubiquitin ligase in protein quality control. *Nature* 467:470–473.
28. Defenouillère Q, et al. (2013) Cdc48-associated complex bound to 60S particles is required for the clearance of aberrant translation products. *Proc Natl Acad Sci USA* 110:5046–5051.
29. Shen PS, et al. (2015) Protein synthesis. Rqc2p and 60S ribosomal subunits mediate mRNA-independent elongation of nascent chains. *Science* 347:75–78.
30. Yonashiro R, et al. (2016) The Rqc2/Tae2 subunit of the ribosome-associated quality control (RQC) complex marks ribosome-stalled nascent polypeptide chains for aggregation. *eLife* 5:e11794.
31. Choe YJ, et al. (2016) Failure of RQC machinery causes protein aggregation and proteotoxic stress. *Nature* 531:191–195.
32. He F, Amrani N, Johansson MJ, Jacobson A (2008) Chapter 6. Qualitative and quantitative assessment of the activity of the yeast nonsense-mediated mRNA decay pathway. *Methods Enzymol* 449:127–147.
33. Kim D, et al. (2013) TopHat2: Accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. *Genome Biol* 14:R36.
34. Robinson JT, et al. (2011) Integrative genomics viewer. *Nat Biotechnol* 29:24–26.
35. Carr-Schmid A, Pfund C, Craig EA, Kinzy TG (2002) Novel G-protein complex whose requirement is linked to the translational status of the cell. *Mol Cell Biol* 22:2564–2574.
36. van Hoof A, Lennertz P, Parker R (2000) Three conserved members of the RNase D family have unique and overlapping functions in the processing of 5S, 5.8S, U4, U5, RNase MRP and RNase P RNAs in yeast. *EMBO J* 19:1357–1365.
37. Tomecki R, et al. (2010) The human core exosome interacts with differentially localized processive RNases: hDIS3 and hDIS3L. *EMBO J* 29:2342–2357.
38. Ghaemmaghami S, et al. (2003) Global analysis of protein expression in yeast. *Nature* 425:737–741.
39. Kulak NA, Pichler G, Paron I, Nagaraj N, Mann M (2014) Minimal, encapsulated proteomic-sample processing applied to copy-number estimation in eukaryotic cells. *Nat Methods* 11:319–324.
40. Malecki M, et al. (2013) The exoribonuclease Dis3L2 defines a novel eukaryotic RNA degradation pathway. *EMBO J* 32:1842–1854.
41. Liu JJ, et al. (2016) CryoEM structure of yeast cytoplasmic exosome complex. *Cell Res* 26:822–837.