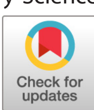
RESEARCH ARTICLE
Ecological and Evolutionary Science



Identification of an Exceptionally Long Intron in the *HAC1* Gene of *Candida parapsilosis*

Elise Iracane,[a] Paul D. Donovan,[a]* Mihaela Ola,[a] Geraldine Butler,[a] Linda M. Holland[a]*

[a]School of Biomedical and Biomolecular Science and UCD Conway Institute of Biomolecular and Biomedical Research, Conway Institute, University College Dublin, Dublin, Ireland

ABSTRACT The unfolded protein response (UPR) in the endoplasmic reticulum (ER) is well conserved in eukaryotes from metazoa to yeast. The transcription factor *HAC1* is a major regulator of the UPR in many eukaryotes. Deleting *HAC1* in the yeast *Candida parapsilosis* rendered cells more sensitive to DTT, a known inducer of the UPR. The deletion strain was also sensitive to Congo red, calcofluor white, and the antifungal drug ketoconazole, indicating that *HAC1* has a role in cell wall maintenance. Transcriptomic analysis revealed that treatment of the wild type with DTT resulted in the increased expression of 368 genes. Comparison with mutant cells treated with DTT reveals that expression of 137 of these genes requires *HAC1*. Enriched GO term analysis includes response to ER stress, cell wall biogenesis and glycosylation. Orthologs of many of these are associated with UPR in *Saccharomyces cerevisiae* and *Candida albicans*. Unconventional splicing of an intron from *HAC1* mRNA is required to produce a functional transcription factor. The spliced intron varies in length from 19 bases in *C. albicans* to 379 bases in *Candida glabrata*, but has not been previously identified in *Candida parapsilosis* and related species. We used RNA-seq data and *in silico* analysis to identify the *HAC1* intron in 12 species in the CTG-Ser1 clade. We show that the intron has undergone major contractions and expansions in this clade, reaching up to 848 bases. Exposure to DTT induced splicing of the long intron in *C. parapsilosis HAC1*, inducing the UPR.

IMPORTANCE The unfolded protein response (UPR) responds to the build-up of misfolded proteins in the endoplasmic reticulum. The UPR has wide-ranging functions from fungal pathogenesis to applications in biotechnology. The UPR is regulated through the splicing of an unconventional intron in the *HAC1* gene. This intron has been described in many fungal species and is of variable length. Until now it was believed that some members of the CTG-Ser1 clade such as *C. parapsilosis* did not contain an intron in *HAC1*, suggesting that the UPR was regulated in a different manner. Here we demonstrate that *HAC1* plays an important role in regulating the UPR in *C. parapsilosis*. We also identified an unusually long intron (626 bp) in *C. parapsilosis HAC1*. Further analysis showed that *HAC1* orthologs in several species in the CTG-Ser1 clade contain long introns.

KEYWORDS *Candida parapsilosis*, Hac1, introns, unfolded protein response

The unfolded protein response (UPR) is activated in response to the build-up of misfolded proteins in the endoplasmic reticulum (ER). Expression of genes required to deal with the ER stress is induced during the UPR (1). The UPR response in fungi was first characterized in the model yeast *Saccharomyces cerevisiae* (2). The UPR is triggered by a transmembrane sensor, Ire1 (inositol requiring enzyme 1), which senses the accumulation of misfolded proteins. Ire1 is an endonuclease that cleaves and removes an atypical intron from *HAC1* mRNA (3). This facilitates translation of the bZIP transcription factor Hac1, which subsequently regulates the expression of genes required

Received 26 September 2018 **Accepted** 10 October 2018 **Published** 7 November 2018

Citation Iracane E, Donovan PD, Ola M, Butler G, Holland LM. 2018. Identification of an exceptionally long intron in the *HAC1* gene of *Candida parapsilosis*. mSphere 3:e00532-18. https://doi.org/10.1128/mSphere.00532-18.

Editor Aaron P. Mitchell, Carnegie Mellon University

Copyright © 2018 Iracane et al. This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International license.

Address correspondence to Linda M. Holland, linda.holland@dcu.ie.

* Present address: Paul D. Donovan, Department of Physiology and Medical Physics, Centre for Systems Medicine, Royal College of Surgeons in Ireland, Dublin, Ireland; Linda M. Holland, School of Biotechnology, Dublin City University, Dublin, Ireland.

for the UPR (1). Hac1 binds to the UPR elements (UPRE) present in the promoter regions of ER-chaperone genes such as *KAR2/*BiP and induces gene transcription (4). The response to unfolded proteins is evolutionarily conserved and plays a central role in the ER stress response in eukaryotes (5).

The UPR is important for fungal pathogenesis (5, 6). In *Candida albicans* Hac1 is required for hyphal formation, which is an important aspect of virulence for this pathogen (7). The UPR is also important for antifungal activity, as UPR-impaired mutants in *C. albicans* are more sensitive to chemicals such as carvacrol (8). The UPR is also required for the virulence and antifungal resistance of *Aspergillus fumigatus* (9). In *Candida glabrata*, Ire1 was found to be required for ER stress but acts independently of Hac1 (10).

The activity of Hac1 has been exploited for biotechnology applications. *Pichia pastoris* (*Komagataella phaffii*) is a widely used system for protein production, and studies have shown that high-level expression of heterologous protein can induce the UPR (11). This can be overcome by overexpressing the spliced form of *HAC1*, increasing the production of heterologous proteins (12, 13). The same method was used in *Aspergillus niger* var. *awamori* and in *Trichoderma reesei* to increase the yield of secreted heterologous protein (14, 15). Moreover, in *T. reesei* protein secretion is regulated not only by the UPR but by another stress response system named REpression under Secretion Stress (RESS) (16).

Ire1-mediated splicing of the *HAC1* intron has been described in many fungi, including *T. reesei*, *A. nidulans*, *C. albicans*, *Yarrowia lipolytica* and *P. pastoris* (4, 7, 17, 18). The overall structure of the intron is well conserved. Common features include two short hairpins at the exon/intron boundaries with the splice sites located within these regions (19). However, the length of the intron varies from 19 nucleotides in *C. albicans* to 379 nucleotides in *C. glabrata* (19).

C. albicans belongs to the CTG-Ser1 clade (species where CTG is translated as serine rather than leucine) (20, 21). Hooks and Griffiths-Jones (19) showed that in some species in the CTG clade (including *C. albicans*, *Candida tropicalis*, and *Candida dubliniensis*) the *HAC1* intron is very short (between 19 bp and 22 bp). However, they could not identify the intron in the other CTG-Ser1 species, including *C. parapsilosis, Lodderomyces elongisporus, Debaryomyces hansenii*, *Scheffersomyces stipitis*, *Clavispora lusitaniae* and *Meyerozyma guilliermondii,* suggesting that these species may use an alternative mechanism to regulate the UPR (19).

Here we describe the role of *HAC1* in the *C. parapsilosis* UPR. Deletion of *HAC1* renders strains susceptible to ER stress. RNA-seq experiments further confirm a role for *C. parapsilosis HAC1* in the ER stress response. We also show that there is an exceptionally long intron (626 nucleotides) in *C. parapsilosis HAC1*, which is spliced under ER stress growth conditions. *HAC1* genes in other CTG-Ser1 clade species also contain unusually long introns.

RESULTS AND DISCUSSION

Functional characterization of *C. parapsilosis HAC1*. A putative *HAC1* gene, *CPAR2_103720*, was identified in the *C. parapsilosis* genome based on sequence similarity to other *HAC1* orthologs (22, 23). Previous studies have shown that *HAC1* has a core role in the UPR (7). Deleting *CPAR2_103720* in *C. parapsilosis* resulted in increased sensitivity to DTT (a strong reducing agent that induces the UPR by preventing disulfide-bond formation) in comparison to the control strain CPRI (24) (Fig. 1A). Although the *hac1*Δ/Δ mutants displayed a growth defect when grown on YPD agar, growth is significantly more reduced in the presence of DTT (Fig. 1B). *C. parapsilosis* Hac1 therefore plays an important role in the UPR, similar to other fungal species.

Deleting *HAC1* also increased sensitivity to Congo red (interferes with glucan synthesis and cross-linking [25]), calcofluor white (interferes with glucan synthesis and cross-linking [26]) and the antifungal ketoconazole (Fig. 1B). Similar phenotypes were observed in *C. albicans hac1* deletions (7–9). These results indicate that *HAC1* plays an essential role in *Candida* species in regulating the response to cell wall stress. Main-

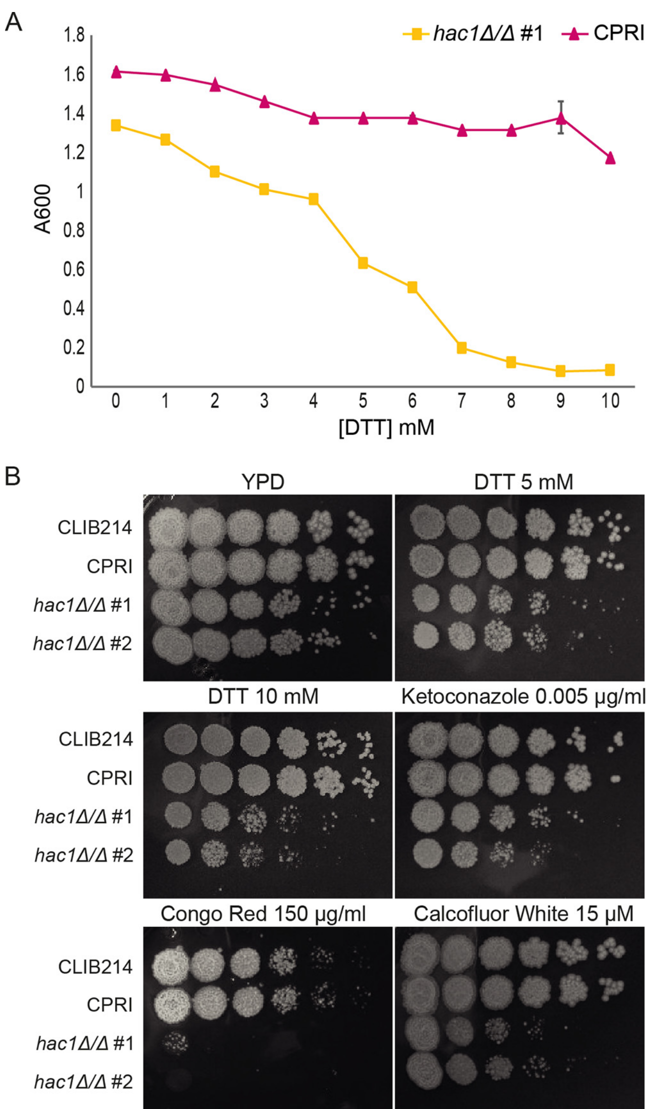


FIG 1 Role of *HAC1* in stress response. (A) Growth at 24 h of the *C. parapsilosis* CPRI control strain and the *hac1*Δ/Δ #1 mutant incubated with DTT (0 to 10 mM) in liquid YPD at 30°C. (B) Growth of *C. parapsilosis* CLIB214, CPRI, *hac1*Δ/Δ #1 and *hac1*Δ/Δ #2 on solid YPD or YPD supplemented with DTT (5 mM and 10 mM), ketoconazole (0.005 μg/ml), Congo red (150 μg/ml), and calcofluor white (15 μM). Plates were incubated at 30°C for 48 h.

taining cell wall integrity is essential for normal cell growth, division, hypha formation, and antifungal tolerance (27).

Transcriptional profiling of the *hac1* deletion. RNA-seq was used to identify the targets of Hac1 in *C. parapsilosis* under UPR. Exponentially growing *C. parapsilosis* CLIB214 (wild type) and *hac1*Δ/Δ strains were grown in YPD with and without exposure to DTT for 1 h. Exposing *C. parapsilosis* CLIB214 to DTT resulted in increased expression of 368 genes and decreased expression of 589 genes (Fig. 2A). Ribosome biogenesis and assembly genes are downregulated, similar to *C. albicans* and *S. cerevisiae* (7). Upregulated genes are enriched for categories such as ERAD pathway and response to endoplasmic reticulum stress, which are associated with the UPR response in *C. albicans* (7). When *HAC1* is deleted, 230 genes remain upregulated when DTT is added. Therefore, expression of these genes is Hac1-independent. We found that expression of 138 genes was no longer induced by DTT when *HAC1* was deleted. From this group of 138 genes, expression of 126 genes (see Table S1 posted at https://figshare.com/s/1f2fc034a73948fe0963) was no longer upregulated in DTT-treated *HAC1* mutant cells,

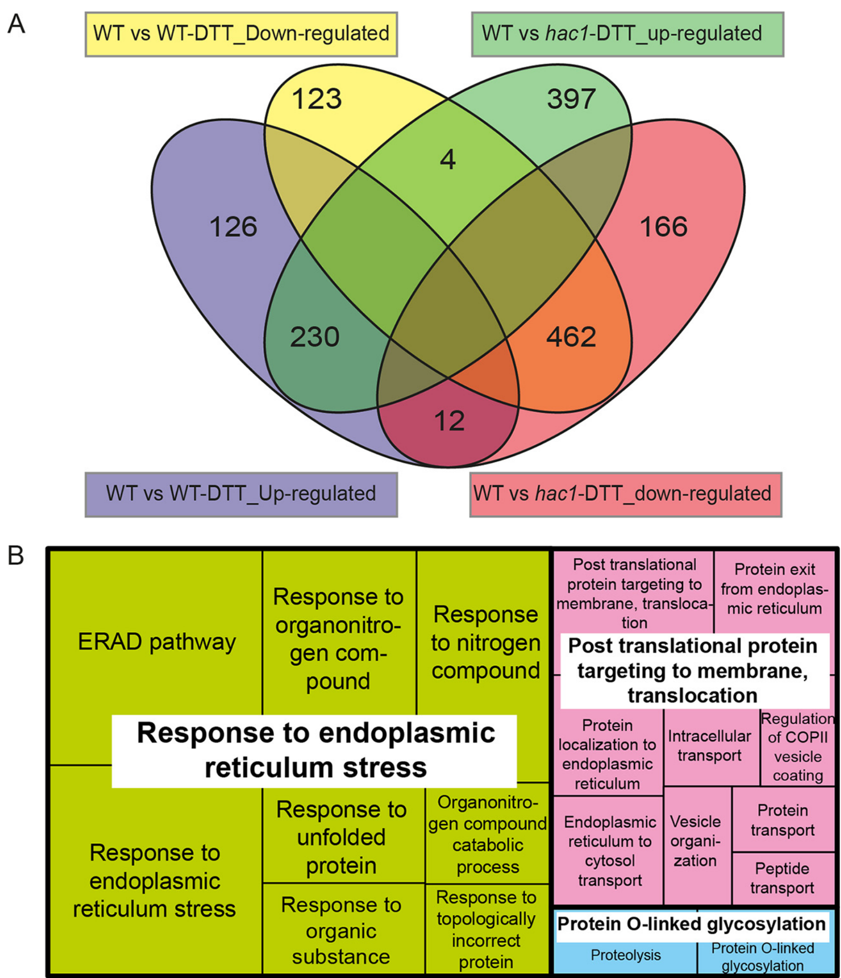


FIG 2 Transcriptomic analysis of UPR induction in CLIB214 and *hac1*Δ/Δ strains. (A) Genes upregulated and downregulated in WT and *hac1*Δ/Δ strains following 1-h exposure to 5 mM DTT. (B) Associated GO terms for biological process of the 138 Hac1-dependent genes. The TreeMap has been generated by REVIGO (48); each tile size is proportional to the absolute $\log_{10}$ of the *P* value of each GO ID.

and expression of 12 genes (including *HAC1* [see Table 3]) was reduced (Fig. 2A). GO term analysis of these 12 Hac1-dependent genes identified enrichment of processes related to response to endoplasmic reticulum stress, posttranslational protein targeting to membrane and protein O-linked glycosylation (Fig. 2B). Orthologs of several of these genes are also *hac1*-dependent in other species such as in *C. albicans* (6 genes [7]), *S. cerevisiae* (7 genes [1, 28, 29]) and *H. polymorpha* (4 genes [30]). Three are required for virulence in *C. albicans*: *PMT1* (*CPAR2_704010*), *PMT4* (*CPAR2_104900*) and *SERP1* (*CPAR2_102440*) (31, 32). *CPAR2_602430*, which is regulated by Hac1 in *C. parapsilosis* but not in *C. albicans*, is upregulated during infection of THP-1 monocytes by *C. parapsilosis* (33). Hac1 is therefore likely to be important for pathogenicity in both species.

Identification of a noncanonical intron in *C. parapsilosis HAC1*. Hooks and Griffiths-Jones (19) carried out a detailed investigation of *HAC1* introns in many fungal species. However, they could find only the hairpin flanking the 5′ splice sites but not the 3′ splice sites of putative introns in *HAC1* in *C. parapsilosis* and four other species in the CTG-Ser1 clade, and they could not identify any intronic features in *C. lusitaniae*. We used a splice-aware tool, HISAT2 (34), to map RNA-seq data for *C. parapsilosis* (35), and identified both splice sites in a putative intron in the last third of the ORF of the gene (Fig. 3A). The intron is unusually long (626 bp), in comparison to that of *C. albicans* (19 nucleotides) or *S. cerevisiae* (252 nucleotides) (7).

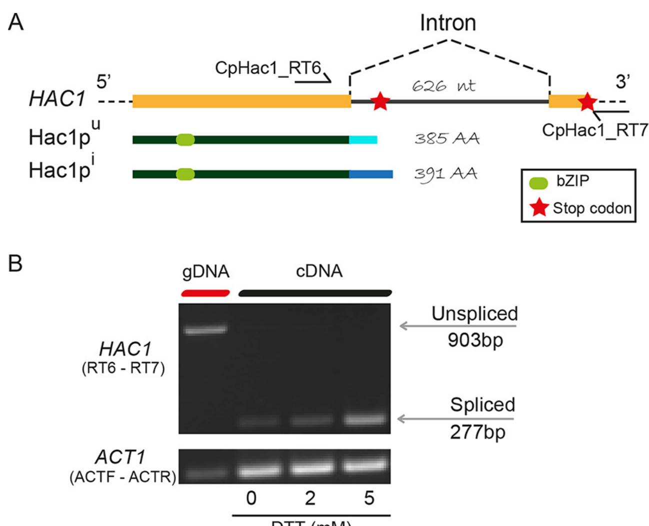


FIG 3 *HAC1* splicing during UPR. (A) Schematic representation of *HAC1*, Hac1p uninduced (Hac1p^u) and Hac1p induced (Hac1p^i). Predicted bZIP domains are represented in light green; stop codon is represented by a red star. (B) RT-PCR of *C. parapsilosis* CLIB214 (WT) cells treated with 0 mM, 2 mM or 5 mM DTT for 1 h. CpHac1_RT6 and CpHac1_RT7 primers (shown in panel A) used to amplify *HAC1*. Actin was amplified using primers ACTF and ACTR (Table 2). A control using genomic DNA is shown in the first lane.

We next used RT-PCR to determine if splicing of the long intron is regulated by the UPR. Genomic DNA was used as a control to amplify the unspliced PCR product (903 bp) (Fig. 3B). Exposure to DTT induced the removal of the intron (Fig. 3B). There is evidence of spliced and unspliced products when cells are grown in the absence of DTT (Fig. 3B). The very low level of unspliced product in the absence of stress has been observed in other yeast species, including *P. pastoris* (18). However, when DTT is added, the amount of spliced product is greatly increased, suggesting that splicing of *C. parapsilosis HAC1* is regulated in a similar manner to other fungi (Fig. 3B) (4, 7, 9).

When the intron is not removed, the predicted Hac1p^u protein (uninduced form of Hac1) is 385 amino acids long because of a premature stop codon in the intron. Splicing of the intron will generate a Hac1p^i protein (induced form) of 391 amino acids (Fig. 3A). Hac1p^i is only 6 amino acids longer than Hac1p^u. However, the C-terminal tail differs by 33 amino acids. The bZIP domain is conserved in both versions of the protein (36) (Fig. 3A). This is the domain that binds to the DNA to activate the target genes (37), suggesting that the presence of the bZIP domain is not enough to activate Hac1p. Cox and Walter (2) showed that in *S. cerevisiae* the Hac1p^u C-terminal tail was extremely unstable, leading to a rapid degradation of the protein.

Identification of the atypical intron in other CTG-Ser1 clade species. We next used RNA-seq data to identify introns in *HAC1* in other species in the CTG-Ser1 clade, including *Candida orthopsilosis*, *Lodderomyces elongisporus* (35), *Scheffersomyces stipitis* (38), and *Clavispora lusitaniae* (39). Introns were predicted by manual inspection in *HAC1* orthologs in *Candida tenuis*, *Meyerozyma guilliermondii*, *Debaryomyces hansenii*, *Candida tanzawaensis*, *Spathaspora passalidarum*, and *Candida metapsilosis* (Fig. 4). All of the features of the intron are present in all species, including two hairpin loops surrounding the 5′ and 3′ splice sites (Fig. 4A). The newly identified introns vary in length from 152 nucleotides in *C. tenuis* to almost 850 nucleotides in *L. elongisporus*. The unexpectedly long intron length and an incorrect annotation of the *HAC1* open reading frame in *C. albicans* orthologs may explain why these introns were not previously identified (19).

Figure 4B shows the distribution of intron length in *HAC1* orthologs in the Saccharomycotina (tree adapted from Pryszcz et al. [40]). Introns fall into three main groupings: small ($<$50 nucleotides), medium (50 to 500 nucleotides), and large ($>$500 nucleotides). Orthologs from related species have introns of similar length. For exam-

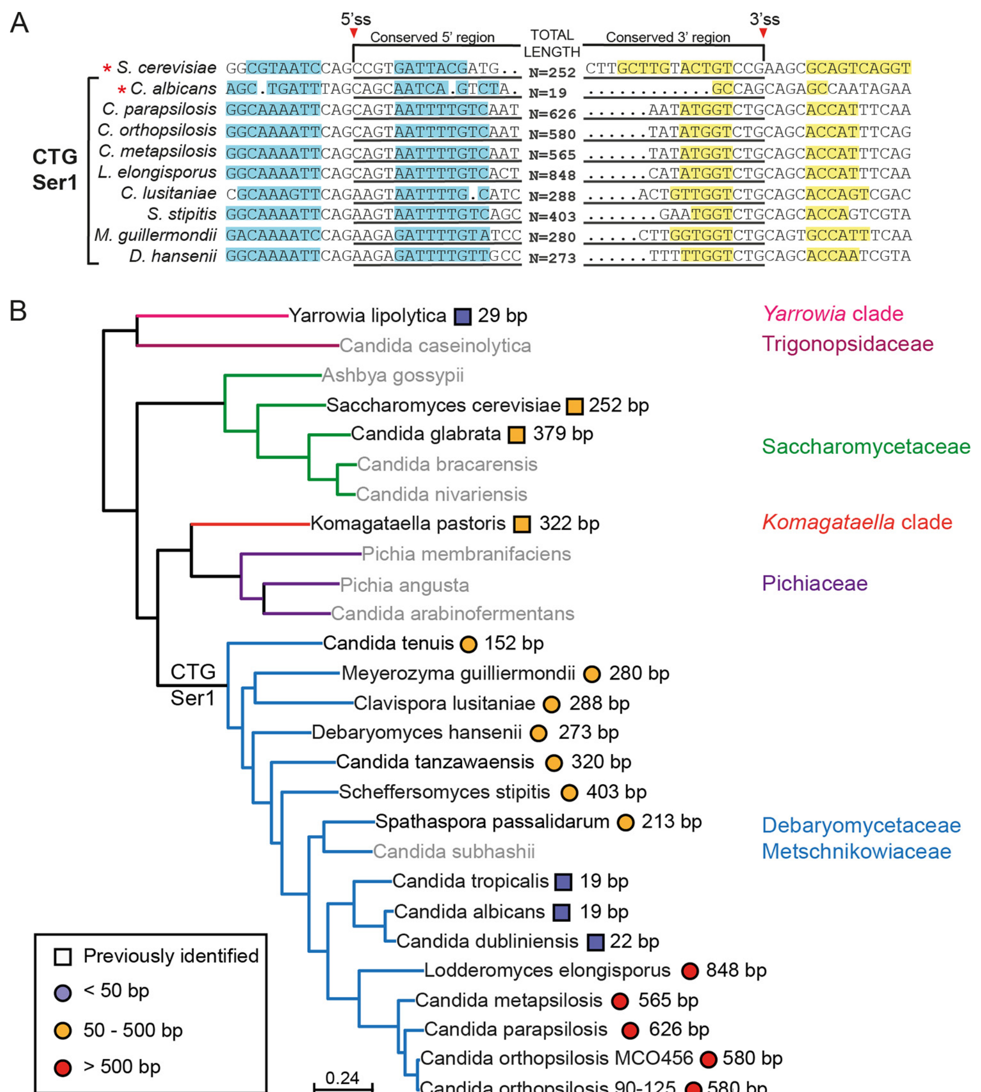


FIG 4 *HAC1* intron across CTG-Ser1 clade. (A) Alignment of *HAC1* from 10 species. The 5′ and 3′ splicing sites are indicated. Red asterisks indicate previously identified introns. The stem-loop complementary regions are highlighted in blue and yellow. (B) Phylogenetic tree represents the *HAC1* intron size through the Saccharomycotina. Species shown in gray were not examined (*Candida caseinolytica*, *Ashbya gossypii*, *Candida bracarensis*, *Candida nivariensis*, *Pichia membranifaciens*, *Pichia angusta*, *Candida arabinofermentans*, and *Candida subhashii*). The species where the intron was previously identified are represented by squares, and those with introns newly identified during this study are represented by circles.

ple, species within the *Saccharomycetaceae* (e.g., *S. cerevisiae* and *C. glabrata*) have medium-length introns. It is likely that the ancestor of the Saccharomycotina also had a medium-length intron, because similar lengths are observed in species in the *Komagataella* clade (*K. pastoris*) and some members of the CTG-Ser1 clade (e.g., *C. lusitaniae*). However, there have been major contractions and expansions in intron size in other CTG-Ser1 clade species. In *C. albicans*, *C. dubliniensis*, and *C. tropicalis* the introns are very short (19 to 22 nucleotides), whereas in *C. parapsilosis* and related species the introns are very long (up to 840 nucleotides).

Conclusion. The UPR is essential for optimizing the response to endoplasmic reticulum stress and the build-up of misfolded proteins, and in many eukaryotes is activated by splicing of a noncanonical intron in *HAC1*. In contrast to a previous report (19), we show that the intron is present in *HAC1* orthologs throughout species in the CTG-Ser1 clade. However, the length of the intron varies substantially within this clade, ranging from 19 nucleotides to 848 nucleotides. The length of the intron does not

TABLE 1 Strains used in the study

Strain	Species	Genotype	Source or reference
CLIB214	*C. parapsilosis*	Wild type	Type strain
CPRI	*C. parapsilosis*	*leu2::FRT/leu2::FRT his1::FRT/his1::FRT frt::CmLEU2/frt::CdHIS1*	41
CPL2H1	*C. parapsilosis*	*leu2::FRT/leu2::FRT his1::FRT/his1::FRT*	41
*hac1*Δ#1	*C. parapsilosis*	*leu2::FRT/leu2::FRT his1::FRT/his1::FRT hac1::CmLEU2/hac1::CdHIS1*	This study
*hac1*Δ#2	*C. parapsilosis*	*leu2::FRT/leu2::FRT his1::FRT/his1::FRT hac1::CmLEU2/hac1::CdHIS1*	This study

appear to alter the function or regulation of *HAC1*; orthologs in *C. albicans* (19 nucleotides) and *C. parapsilosis* (626 nucleotides) are spliced in a similar manner (7), and regulate expression of ERAD and other stress genes.

MATERIALS AND METHODS

Strains, media, and growth. All strains are listed in Table 1, and all primers are listed in Table 2. Yeast strains were grown in liquid YPD (2% glucose, 2% peptone, 1% yeast extract) supplemented with 1 mM to 10 mM dithiothreitol (DTT; Sigma-Aldrich D0632) where indicated. For phenotype analysis, yeast cells from an overnight culture were washed twice in PBS, diluted to an A_{600} of 1 in PBS, and serially diluted 1:5 five times in a 96-well plate. Dilutions were pinned with a 48-pin replicator to YPD agar, supplemented with 5 mM DTT, 10 mM DTT, 0.005 μg/ml ketoconazole, 150 μg/ml Congo red or 15 μM calcofluor white where indicated. Plates were incubated at 30°C for 48 h. *CPAR2_103720* (*HAC1*) was deleted in *C. parapsilosis* CPL2H1 by replacement of one allele with *HIS1* from *C. dubliniensis* and the second with *LEU2* from *Candida maltosa* by homologous recombination as described previously (41). Two *HAC1* deletion strains were generated by deleting the first *HAC1* allele to create a heterozygous mutant. The second allele was deleted in this heterozygous strain and two individual mutants were chosen called *C. parapsilosis hac1*Δ/Δ #1 and *hac1*Δ/Δ #2. CPRI is the control strain with integration of *CdHIS1* and *CmLEU2* at the site of the original *HIS1* alleles. All primers used are listed in Table 2. Upstream and downstream regions were amplified using Q5 High-Fidelity DNA polymerase (New England BioLabs) with, respectively, CpHAC1_KO1 and CpHAC1_KO3 primers, and CpHAC1_KO4 and CpHAC1_KO6 primers. The selectable markers *HIS1* and *LEU2* were amplified using *Ex Taq* polymerase (TaKaRa) with CpHAC1_KO2 and CpHAC1_KO5. The CpHAC1_KO5 primer introduces a unique barcode into the deletion strain (Table 2). The upstream region, one of the selectable markers and the downstream region were fused by PCR using *Ex Taq* Polymerase (TaKaRa) and the resulting disruption cassette was transformed into *C. parapsilosis* CPL2H1 by chemical transformation as described previously (41). Correct insertion of marker was confirmed by PCR using CpHAC1_5′Check_2 and LEU_Check_1/HIS_Check_1 in the 5′ region and CpHAC1_3′Check and LEU_Check_2/HIS_Check_2 in the 3′ region. Complete open reading frame deletion was confirmed using primers CpHac1_ORF_F and CpHac1_Check_R.

RNA extraction. *C. parapsilosis* CLIB214, *hac1*Δ/Δ #1, and *hac1*Δ/Δ #2 were grown overnight in YPD, inoculated in 60 ml YPD to an A_{600} of 0.2, and incubated at 30°C with shaking until A_{600} of 0.6 was reached. Fifteen milliliters of this culture was then supplemented with H_2O (control) or 5 mM DTT and incubated for 1 h at 30°C. The cells were collected, resuspended in 200 μl RNAlater, snap-frozen in liquid

TABLE 2 Primers used in the study

Primer	Sequence
RT-PCR	
CpHac1_RT6	TGGGAAACTTTTCACAAAATACG
CpHac1-RT7	TCACACCATAAATCAATCCAACTC
ACTF	GAAGCTTTGTTCCGTCCAGC
ACTR	TGATGGAGCCAAAGCAGTGA
HAC1 deletion in *Candida parapsilosis*	
CpHAC1_KO1	ATACCCCCTTTGGATCAATT
CpHAC1_KO3	CACGGCGCGCCTAGCAGCGGGACTAGTATGTGTGGGCTTA
CpHAC1_KO2	CCGCTGCTAGGCGCGCCGTGACCAGTGTGATGGATATCTGC (universal primer)
CpHAC1_KO5	GCAGGGATGCGGCCGCTGAC**GCGCAACCTTCCGGAGTAT**AGCTCGGATCCACTAGTAACG[a]
CpHAC1_KO4	GTCAGCGGCCGCATCCCTGCAATAATCAAGTTATTTTTAG
CpHAC1_KO6	CCTCATTCAGTGGGAGTG
CpHAC1_5′Check_2	CGATGAAACGCAGTAGCAAA
CpHAC1_3′Check	TATAACACAAGAAAACAATC
LEU check 1	GAAGTTGGTGACGCGATTGT
LEU check 2	TTCCCCTTCAATGTATGCAA
HIS check 1	AAAATCAATGGGCATTCTCG
HIS check 2	TGGGAAGCAGACATTCAACA
CpHac1_check_R	TTTCCACCTCTTCTTGAACCA
CpHac1_ORF_F	CCACCTAGGAAGAGAGCCAAG

[a]The barcode present in primer CpHAC1_KO5 is shown in bold and underlined.

TABLE 3 List of 12 genes downregulated in *hac1*Δ/Δ strain involved in UPR[a]

Gene name	WT-DTT	*Hac1*-DTT	Orthologs in *C. albicans*	Orthologs in *S. cerevisiae*	Orthologs in *H. polymorpha*	Description based on orthologs (from Candida Genome Database (47) or cited references)
CPAR2_212260	2.541366	-1.43765	Orf19.5831	/	/	Hypothetical protein
CPAR2_213780	1.906718	-2.03439	*KAR2*	*KAR2*	*KAR2*	Similar to chaperones of Hsp70p family. In *C.albicans* role in translocation of proteins into the ER (49)
CPAR2_102440	1.544158	-1.43188	Orf19.242.2	*YSY6 / SERP1*	/	Endoplasmic reticulum localization. Potential virulence factor in *C. albicans* (32).
CPAR2_103720	1.539826	-7.03222	*HAC1*	*HAC1*	*HAC1*	bZIP TF, role in UPR; In *C. albicans* : control of morphology , atypical splicing at C-term under ER stress, induced during mating and by caspofungin.
CPAR2_400620	1.49627	-1.63123	Orf19.4579	*ERV29*	/	Putative SURF4 family member, role in ER to Golgi vesicle-mediated transport. Plasma membrane localized in *C. albicans*.
CPAR2_602430	1.402542	-1.29599	Orf19.3438	*SCJ1*	*SCJ1*	Chaperone binding activity, role in ER-associated protein catabolic process, protein folding in ER, role in UPR, ER lumen localization. Works with Kar2 in *S. cerevisiae* (50).
CPAR2_212250	1.400748	-1.52207	Orf19.5830	*LHS1/CER1/SSI1*	*LHS1*	Similar to *S. cerevisiae* Hsp70p, ATP binding, adenyl-nucleotide exchange factor activity, role to UPR, role in post-translational protein targeting to membrane (51). In *C. albicans* predicted Kex2p substrate, possibly essential.
CPAR2_205310	1.399552	-1.17637	Orf19.6630	*LCL2*	/	ER-associated protein catabolic process
CPAR2_808040	1.207646	-1.56278	*SEC61*	*SEC61*	*SEC61*	Putative ER protein-translocation complex subunit. Essential in *C. albicans* and have 10 predicted transmembrane regions.
CPAR2_603570	1.145137	-1.30171	*DPM2*	/	/	Enzyme part of the complex of dolichol phosphate mannose synthase (DPMS), important for the cell wall composition. (52)
CPAR2_104900	1.111397	-1.23252	*PMT4*	*PMT4*	*PMT4*	Protein mannosyltransferase (PMT) required for normal cell wall composition and full virulence in mouse systemic infection. (31)
CPAR2_704010	1.042138	-1.28259	*PMT1*	*PMT1*	*PMT1*	PMT protein, required for virulence in mouse systemic infection and for adhesion to epithelial cells.(31)

[a]Differential gene expression in fold change in WT cells exposed to DTT (column 2) and in *hac1*Δ/Δ cells exposed to DTT (column 3) compared to WT untreated cells. Color for log fold change: 1> and <1.5, yellow; 1.5> and <2, orange; 2>, red; −1> and <−1.5, light green; −1.5> and <−2, medium green; −2>, dark green. Orthologs found to be Hac1 dependent in *C. albicans*, *S. cerevisiae*, or *H. polymorpha* are underlined.

nitrogen, and stored at −80°C. RNA was extracted using a RiboPure RNA purification kit (yeast) (Ambion; catalog no. AM1926). Quality was assessed by Agilent 2100 Bioanalyzer instrument.

RT-PCR. For each sample 2 μg of RNA was treated with 2 U of DNase I (Invitrogen 18068015) in a total volume of 20 μl for 5 min at room temperature, followed by DNase inactivation by adding 1 μl of 25 mM EDTA and incubating at 65°C for 10 min. To synthesize cDNA, 5 μl of DNase-treated RNA was incubated at 70°C for 10 min with oligo(dT) (Promega) at a final concentration of 20 μg/ml. One microliter RNasin (40 U/ml), 4 μl 5× MMLV-RT buffer, 1 μl dNTPs (10 mM), 1 μl MMLV-RT enzyme and 7 μl RNase free H_2O were added to the cDNA mix and incubated for 1 h at 37°C and then for 2 min at 95°C. The cDNA was then amplified using primers CpHac1_RT6/CpHac1_RT7 (Table 2) to detect splicing of the intron in Hac1 RNA and primers ACTF/ACTR as a control to amplify actin.

RNA sequencing and analysis. RNA extracted from 12 samples was sequenced by BGI Global Genomics Services (100 bases, paired-end reads; over 10 million reads per sample). The samples were 3 biological replicates of CLIB214 (WT) incubated for 1 h with 5 mM DTT or with H_2O, 2 biological replicates of *hac1*Δ/Δ #1 incubated 1 h with 5 mM DTT or with H_2O, and 1 replicate of *hac1*Δ/Δ #2 incubated 1 h with 5 mM DTT or with H_2O. Data were analyzed using established bioinformatic protocols (42). Raw paired-end sequenced reads were trimmed using Skewer v0.1.120 (43) and mapped to the *Candida parapsilosis* CDC317 genome using TopHat v2.0.12 (44). Transcripts were counted using htseq-count from HTSeq v0.6.1 (45), and differentially expressed genes were identified using the Bioconductor package DESeq2 (46). A log_2FC of >1 or <−1 and an adjusted *P* value of <0.001 were used as cutoff values (Table 3).

Gene ontology was found by CGD Gene Ontology Term Finder using default setting (47). REVIGO (48) has been used to generate TreeMap using the default settings (allowed similarity: medium [0.7], semantic similarity measure used: SimRel) in the database with GO term sizes: *Saccharomyces cerevisiae*.

Identification of intron in *HAC1* in several species. HISAT2 (v2.0.4) (34), a splice-aware RNA-seq mapping tool, was used to map RNA-seq data to complete genomes where available, including *C. parapsilosis*, *Candida orthopsilosis* and *Lodderomyces elongisporus* (SRP077251) (35), *Pichia (Scheffersomyces) stipitis* (SRX135712) (38), and *Clavispora lusitaniae* (SRX1131478) (39). Intron predictions were manually inspected by comparing against known *HAC1* intron structures from related species. Where no RNA-seq data were available, the *HAC1* intron structure was identified by manual alignment with known *HAC1* genes (*Candida tenuis*, *Pichia [Meyerozyma] guilliermondii*, *Debaryomyces hansenii*, *Candida tanzawaensis*, *Spathaspora passalidarum*, and *Candida metapsilosis*).

Data availability. RNA-seq data sets used for intron identification for *C. parapsilosis*, *Candida orthopsilosis* and *Lodderomyces elongisporus* are available using the accession number SRP077251 (35), for *Pichia (Scheffersomyces) stipitis* at SRX135712 (38), and for *Clavispora lusitaniae* at SRX1131478 (39). The raw gene expression data for the *C. parapsilosis HAC1* RNA-seq experiment are available at the Gene Expression Omnibus database under accession number GSE120094. Table S1 is available at https://figshare.com/s/1f2fc034a73948fe0963.

ACKNOWLEDGMENTS

L.M.H. was funded by an Irish Research Council, ELEVATE International Career Development fellowship—cofunded by Marie Curie Actions (ELEVATE/2013/84). This work was also supported by grants to G.B. from Science Foundation Ireland (12/IA/1343, https://www.sfi.ie), the Wellcome Trust (102406/Z/13/Z, https://wellcome.ac.uk) and the European Union's Horizon 2020 research and innovation program under the Marie Sklodowska-Curie grant agreement no. H2020-MSCA-ITN-2014-642095.

REFERENCES

1. Travers KJ, Patil CK, Wodicka L, Lockhart DJ, Weissman JS, Walter P. 2000. Functional and genomic analyses reveal an essential coordination between the unfolded protein response and ER-associated degradation. Cell 101:249–258. https://doi.org/10.1016/S0092-8674(00)80835-1.
2. Cox JS, Walter P. 1996. A novel mechanism for regulating activity of a transcription factor that controls the unfolded protein response. Cell 87:391–404. https://doi.org/10.1016/S0092-8674(00)81360-4.
3. Sidrauski C, Walter P. 1997. The transmembrane kinase Ire1p is a site-specific endonuclease that initiates mRNA splicing in the unfolded protein response. Cell 90:1031–1039. https://doi.org/10.1016/S0092-8674(00)80369-4.
4. Oh MH, Cheon SA, Kang HA, Kim J-Y. 2010. Functional characterization of the unconventional splicing of *Yarrowia lipolytica HAC1* mRNA induced by unfolded protein response. Yeast 27:443–452. https://doi.org/10.1002/yea.1762.
5. Hollien J. 2013. Evolution of the unfolded protein response. Biochim Biophys Acta 1833:2458–2463. https://doi.org/10.1016/j.bbamcr.2013.01.016.
6. Krishnan K, Askew DS. 2014. Endoplasmic reticulum stress and fungal pathogenesis. Fungal Biol Rev 28:29–35. https://doi.org/10.1016/j.fbr.2014.07.001.
7. Wimalasena TT, Enjalbert B, Guillemette T, Plumridge A, Budge S, Yin Z, Brown AJP, Archer DB. 2008. Impact of the unfolded protein response upon genome-wide expression patterns, and the role of Hac1 in the polarized growth, of *Candida albicans*. Fungal Genet Biol 45:1235–1247. https://doi.org/10.1016/j.fgb.2008.06.001.
8. Chaillot J, Tebbji F, Remmal A, Boone C, Brown GW, Bellaoui M, Sellam A. 2015. The monoterpene carvacrol generates endoplasmic reticulum stress in the pathogenic fungus *Candida albicans*. Antimicrob Agents Chemother 59:4584–4592. https://doi.org/10.1128/AAC.00551-15.
9. Richie DL, Hartl L, Aimanianda V, Winters MS, Fuller KK, Miley MD, White S, McCarthy JW, Latgé J-P, Feldmesser M, Rhodes JC, Askew DS. 2009. A role for the unfolded protein response (UPR) in virulence and antifungal susceptibility in *Aspergillus fumigatus*. PLoS Pathog 5:e1000258. https://doi.org/10.1371/journal.ppat.1000258.
10. Miyazaki T, Nakayama H, Nagayoshi Y, Kakeya H, Kohno S. 2013. Dissection of Ire1 functions reveals stress response mechanisms uniquely evolved in *Candida glabrata*. PLoS Pathog 9:e1003160. https://doi.org/10.1371/journal.ppat.1003160.
11. Whyteside G, Alcocer MJC, Kumita JR, Dobson CM, Lazarou M, Pleass RJ, Archer DB. 2011. Native-state stability determines the extent of degradation relative to secretion of protein variants from *Pichia pastoris*. PLoS One 6:e22692. https://doi.org/10.1371/journal.pone.0022692.
12. Huang M, Gao Y, Zhou X, Zhang Y, Cai M. 2017. Regulating unfolded protein response activator HAC1p for production of thermostable raw-starch hydrolyzing α-amylase in *Pichia pastoris*. Bioprocess Biosyst Eng 40:341–350. https://doi.org/10.1007/s00449-016-1701-y.
13. Li C, Lin Y, Zheng X, Pang N, Liao X, Liu X, Huang Y, Liang S. 2015. Combined strategies for improving expression of *Citrobacter amalonaticus* phytase in *Pichia pastoris*. BMC Biotechnol 15:88. https://doi.org/10.1186/s12896-015-0204-2.
14. Wu Y, Sun X, Xue X, Luo H, Yao B, Xie X, Su X. 2017. Overexpressing key component genes of the secretion pathway for enhanced secretion of an *Aspergillus niger* glucose oxidase in *Trichoderma reesei*. Enzyme Microb Technol 106:83–87. https://doi.org/10.1016/j.enzmictec.2017.07.007.
15. Valkonen M, Ward M, Wang H, Penttilä M, Saloheimo M. 2003. Improvement of foreign-protein production in *Aspergillus niger* var. awamori by constitutive induction of the unfolded-protein response. Appl Environ Microbiol 69:6979–6986. https://doi.org/10.1128/AEM.69.12.6979-6986.2003.
16. Saloheimo M, Pakula TM. 2012. The cargo and the transport system: secreted proteins and protein secretion in *Trichoderma reesei* (*Hypocrea jecorina*). Microbiology 158:46–57. https://doi.org/10.1099/mic.0.053132-0.
17. Saloheimo M, Valkonen M, Penttilä M. 2003. Activation mechanisms of the *HAC1*-mediated unfolded protein response in filamentous fungi. Mol Microbiol 47:1149–1161. https://doi.org/10.1046/j.1365-2958.2003.03363.x.
18. Guerfal M, Ryckaert S, Jacobs PP, Ameloot P, Van Craenenbroeck K, Derycke R, Callewaert N. 2010. The *HAC1* gene from *Pichia pastoris*: characterization and effect of its overexpression on the production of secreted, surface displayed and membrane proteins. Microb Cell Fact 9:49. https://doi.org/10.1186/1475-2859-9-49.
19. Hooks KB, Griffiths-Jones S. 2011. Conserved RNA structures in the non-canonical *Hac1/Xbp1* intron. RNA Biol 8:552–556. https://doi.org/10.4161/rna.8.4.15396.
20. Ohama T, Suzuki T, Mori M, Osawa S, Ueda T, Watanabe K, Nakase T. 1993. Non-universal decoding of the leucine codon CUG in several Candida species. Nucleic Acids Res 21:4039–4045. https://doi.org/10.1093/nar/21.17.4039.
21. Krassowski T, Coughlan AY, Shen X-X, Zhou X, Kominek J, Opulente DA, Riley R, Grigoriev IV, Maheshwari N, Shields DC, Kurtzman CP, Hittinger CT, Rokas A, Wolfe KH. 2018. Evolutionary instability of CUG-Leu in the genetic code of budding yeasts. Nat Commun 9:1887. https://doi.org/10.1038/s41467-018-04374-7.
22. Maguire SL, ÓhÉigeartaigh SS, Byrne KP, Schröder MS, O'Gaora P, Wolfe KH, Butler G. 2013. Comparative genome analysis and gene finding in *Candida* species using CGOB. Mol Biol Evol 30:1281–1291. https://doi.org/10.1093/molbev/mst042.
23. Fitzpatrick DA, O'Gaora P, Byrne KP, Butler G. 2010. Analysis of gene evolution and metabolic pathways using the Candida Gene Order Browser. BMC Genomics 11:290. https://doi.org/10.1186/1471-2164-11-290.
24. Jocelyn PC. 1987. Chemical reduction of disulfides. Methods Enzymol 143:246–256. https://doi.org/10.1016/0076-6879(87)43048-6.
25. Kopecká M, Gabriel M. 1992. The influence of Congo red on the cell wall and (1→3)-β-D-glucan microfibril biogenesis in *Saccharomyces cerevisiae*. Arch Microbiol 158:115–126. https://doi.org/10.1007/BF00245214.
26. Roncero C, Durán A. 1985. Effect of Calcofluor white and Congo red on fungal cell wall morphogenesis: in vivo activation of chitin polymerization. J Bacteriol 163:1180–1185.
27. Levin DE. 2011. Regulation of cell wall biogenesis in *Saccharomyces cerevisiae*: the cell wall integrity signalling pathway. Genetics 189:1145–1175. https://doi.org/10.1534/genetics.111.128264.
28. Umebayashi K, Hirata A, Horiuchi H, Ohta A, Takagi M. 1999. Unfolded protein response-induced BiP/Kar2p production protects cell growth against accumulation of misfolded protein aggregates in the yeast endoplasmic reticulum. Eur J Cell Physiol 78:726–738. https://doi.org/10.1016/S0171-9335(99)80041-7.
29. Nikawa J, Akiyoshi M, Hirata S, Fukuda T. 1996. *Saccharomyces cerevisiae IRE2/HAC1* is involved in *IRE1*-mediated *KAR2* expression. Nucleic Acids Res 24:4222–4226. https://doi.org/10.1093/nar/24.21.4222.
30. Moon H-Y, Cheon SA, Kim H, Agaphonov MO, Kwon O, Oh D-B, Kim J-Y, Kang HA. 2015. *Hansenula polymorpha* Hac1p is critical to protein N-glycosylation activity modulation, as revealed by functional and transcriptomic analyses. Appl Environ Microbiol 81:6982–6993. https://doi.org/10.1128/AEM.01440-15.
31. Rouabhia M, Schaller M, Corbucci C, Vecchiarelli A, Prill SK-H, Giasson L, Ernst JF. 2005. Virulence of the fungal pathogen *Candida albicans* requires the five isoforms of protein mannosyltransferases. Infect Immun 73:4571–4580. https://doi.org/10.1128/IAI.73.8.4571-4580.2005.

32. Li J, Yu Q, Zhang B, Xiao C, Ma T, Yi X, Liang C, Li M. 2018. Stress-associated endoplasmic reticulum protein 1 (SERP1) and Atg8 synergistically regulate unfolded protein response (UPR) that is independent on autophagy in *Candida albicans*. Int J Med Microbiol 308:378–386. https://doi.org/10.1016/j.ijmm.2018.03.004.
33. Tóth R, Cabral V, Thuer E, Bohner F, Németh T, Papp C, Nimrichter L, Molnár G, Vágvölgyi C, Gabaldón T, Nosanchuk JD, Gácser A. 2018. Investigation of Candida parapsilosis virulence regulatory factors during host-pathogen interaction. Sci Rep 8:1346. https://doi.org/10.1038/s41598-018-19453-4.
34. Kim D, Langmead B, Salzberg SL. 2015. HISAT: a fast spliced aligner with low memory requirements. Nat Methods 12:357–360. https://doi.org/10.1038/nmeth.3317.
35. Donovan PD, Schröder MS, Higgins DG, Butler G. 2016. Identification of non-coding RNAs in the *Candida parapsilosis* species group. PLoS One 11:e0163235. https://doi.org/10.1371/journal.pone.0163235.
36. de Castro E, Sigrist CJA, Gattiker A, Bulliard V, Langendijk-Genevaux PS, Gasteiger E, Bairoch A, Hulo N. 2006. ScanProsite: detection of PROSITE signature matches and ProRule-associated functional and structural residues in proteins. Nucleic Acids Res 34:W362–W365. https://doi.org/10.1093/nar/gkl124.
37. Nojima H, Leem S-H, Araki H, Sakai A, Nakashima N, Kanaoka Y, Ono Y. 1994. Hac1: a novel yeast bZIP protein binding to the CRE motif is a multicopy suppressor for cdcW mutant of *Schizosaccharomyces pombe*. Nucleic Acids Res 22:5279–5288. https://doi.org/10.1093/nar/22.24.5279.
38. Papini M, Nookaew I, Uhlén M, Nielsen J. 2012. *Scheffersomyces stipitis*: a comparative systems biology study with the Crabtree positive yeast *Saccharomyces cerevisiae*. Microb Cell Fact 11:136. https://doi.org/10.1186/1475-2859-11-136.
39. Kapoor S, Zhu L, Froyd C, Liu T, Rusche LN. 2015. Regional centromeres in the yeast *Candida lusitaniae* lack pericentromeric heterochromatin. Proc Natl Acad Sci U S A 112:12139–12144. https://doi.org/10.1073/pnas.1508749112.
40. Pryszcz LP, Németh T, Saus E, Ksiezopolska E, Hegedűsová E, Nosek J, Wolfe KH, Gacser A, Gabaldón T. 2015. The genomic aftermath of hybridization in the opportunistic pathogen *Candida metapsilosis*. PLoS Genet 11:e1005626. https://doi.org/10.1371/journal.pgen.1005626.
41. Holland LM, Schröder MS, Turner SA, Taff H, Andes D, Grózer Z, Gácser A, Ames L, Haynes K, Higgins DG, Butler G. 2014. Comparative phenotypic analysis of the major fungal pathogens *Candida parapsilosis* and *Candida albicans*. PLoS Pathog 10:e1004365. https://doi.org/10.1371/journal.ppat.1004365.
42. Wang C, Schröder MS, Hammel S, Butler G. 2016. Using RNA-seq for analysis of differential gene expression in fungal species. Methods Mol Biol 1361:1–40. https://doi.org/10.1007/978-1-4939-3079-1_1.
43. Jiang H, Lei R, Ding S-W, Zhu S. 2014. Skewer: a fast and accurate adapter trimmer for next-generation sequencing paired-end reads. BMC Bioinformatics 15:182. https://doi.org/10.1186/1471-2105-15-182.
44. Kim D, Pertea G, Trapnell C, Pimentel H, Kelley R, Salzberg SL. 2013. TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. Genome Biol 14:R36. https://doi.org/10.1186/gb-2013-14-4-r36.
45. Anders S, Huber W. 2010. Differential expression analysis for sequence count data. Genome Biol 11:R106. https://doi.org/10.1186/gb-2010-11-10-r106.
46. Love MI, Huber W, Anders S. 2014. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol 15:550. https://doi.org/10.1186/s13059-014-0550-8.
47. Skrzypek MS, Binkley J, Binkley G, Miyasato SR, Simison M, Sherlock G. 2017. The Candida Genome Database (CGD): incorporation of Assembly 22, systematic identifiers and visualization of high throughput sequencing data. Nucleic Acids Res 45:D592–D596. https://doi.org/10.1093/nar/gkw924.
48. Supek F, Bošnjak M, Škunca N, Šmuc T. 2011. REVIGO summarizes and visualizes long lists of gene ontology terms. PLoS One 6:e21800. https://doi.org/10.1371/journal.pone.0021800.
49. Morrow MW, Janke MR, Lund K, Morrison EP, Paulson BA. 2011. The *Candida albicans* Kar2 protein is essential and functions during the translocation of proteins into the endoplasmic reticulum. Curr Genet 57:25–37. https://doi.org/10.1007/s00294-010-0323-1.
50. Silberstein S, Schlenstedt G, Silver PA, Gilmore R. 1998. A role for the DnaJ homologue Scj1p in protein folding in the yeast endoplasmic reticulum. J Cell Physiol 143:921–933.
51. Hamilton TG, Norris TB, Tsuruda PR, Flynn GC. 1999. Cer1p functions as a molecular chaperone in the endoplasmic reticulum of *Saccharomyces cerevisiae*. Mol Cell Biol 19:5298–5307. https://doi.org/10.1128/MCB.19.8.5298.
52. Juchimiuk M, Kruszewska J, Palamarczyk G. 2015. Dolichol phosphate mannose synthase from the pathogenic yeast *Candida albicans* is a multimeric enzyme. Biochim Biophys Acta 1850:2265–2275. https://doi.org/10.1016/j.bbagen.2015.08.012.