Supplemental data

Figure S1. Multiple sequence alignment of the REPULS domains. Positions where hydrophobic and aromatic characters are particularly conserved are shaded green and orange, respectively. Other striking similarities are in blue and pink (basic and acidic residues, respectively), as well as in yellow (small residues or loop-forming residues). Predicted secondary structures are reported up to the alignment. REPULS domains which were particularly difficult to detect are reported at bottom. UniProt identifiers : *Saccharomyces cerevisiae* (Q08562), *Candida glabrata* (Q6FSM2), *Kazachstania africana* (H2AY18), *Naumovozyma dairensis* (G0W9C7), *Naumovozyma castelli* (G0VIR1), *Tetrapisispora phaffii* (G8BZA4), *Vanderwaltozyma polyspora* (A7TPE3)), *Zygosaccharomyces rouxii* (C5DWI1), *Torulospora delbrueckii* (G8ZXX1), *Kluyveromyces waltii* (Kwal14.1868 (from YOGB, Byrne and Wolfe 2005)), *Lachancea thermotolerans* (C5E2Q8), *Kluyveromyces lactis* (Q6CUF0), *Ashbya gossypii* (Q75EC7), *Eremothecium cymbalariae* (G8JR28).

The central region of Uls1 (aa 550 to 900), located after the SIMs and before the Swi2/Snf2-like translocase domain, harbors typical features of globular domains, as deduced from Hydrophobic Cluster Analysis (HCA) (Callebaut et al. 1997). Indeed this region contains ~1/3 of strong hydrophobic amino acids (V I L F M Y W), which are distributed on the HCA two-dimensional plot into hydrophobic clusters, which mainly correspond to the regular secondary structures forming the core of globular domains. Using this Uls1 fragment as query, we searched the sequence databases (non redundant (nr) database at NCBI, 17,919,084 sequences) using PSI-BLAST (Altschul et al. 1997). We deduced from the PSI-BLAST results, obtained at convergence by iteration 3, that Uls1 contains a repeated domain, included between amino acids 677 and 761 (first repeat) and 773 and 855 (second repeat). We named these repeats REPULS after REPeat in ULS1.

REPULS domains are present in Uls1 sequences from yeasts of the *Saccharomyces* complex, also named *Saccharomycetaceae*. *Saccharomycetaceae* are hemiascomycetes that have HML/HMR-like cassettes, which appear after the split from the *Yarrowia lipolytica* and *Debaromyces hansenii/ Candida albicans* branch (Génolevures Consortium 2009, Dujon 2010). No REPULS domains could indeed be detected in these last yeast species, belonging to other subdivisions of hemiascomycetes, or in other more distant species. Strict tandem repeats are

observed in Uls1 from Saccharomyces cerevisae to Zygosaccharomyces rouxii, whereas Kluyveromyces waltii and Lachancea thermotolerans Uls1 have only one REPULS domain, and Kluyveromyces lactis, Ashbya gossypii and Eremothecium cymbalariae have multiple copies.

Secondary structure predictions indicate that the REPULS domains are likely made of four alpha-helices. No obvious similarity has been detected with known 3D structures, using sequence similarity searches (see before) or fold recognition programs (e.g. Phyre : Bennett-Lovsey et al. 2008). Positions in which the hydrophobic character is conserved (in green) likely participate in the hydrophobic core of the domain. Positions mainly occupied by aromatic amino acids (orange) might play a role either in the packing of the hydrophobic core or in a specific function, at the surface of the domain. REPULS domains contain several conserved basic amino acids and are characterized by a striking sequence signature (G-[LV]-K-M-[PDEN]), located in the loop linking helices C and D. Worth noting is that these sequence features of REPULS domains are well conserved in Uls1 proteins of species containing only two (tandem) or one (single) REPULS domain(s). In species having more than two domains, some of the additional REPULS domains lose some of these sequence characteristics, so that they are difficult to detect. This is particularly the case for the degenerated REPULS repeats D of Ashbya gossypii and Eremothecium cymbalariae and repeat F of Kluyveromyces lactis. The sequence of the repeat D of Eremothecium cymbalariae was the only one that could be detected from the PSI-BLAST results. The relationships between the REPULS repeat D of Ashbya gossypii and Eremothecium cymbalariae and the REPULS repeat F of Kluyveromyces lactis were then detected using HCA.

Figure S2. Telomere length in cells lacking Uls1, Rif2 and Sir4. Genomic DNA from strains Lev346 (WT), 184-48c (*uls1-\Delta*), 183-30d (*rif2-\Delta*), Lev575 (*sir4-\Delta*), 184-14c (*rif2-\Delta sir4-\Delta*), 183-10c (*uls1-\Delta sir4-\Delta*), 183-2d (*uls1-\Delta rif2-\Delta*) and 183-25a (*uls1-\Delta rif2-\Delta sir4-\Delta*) was digested with XhoI, separated in a 0.9% agarose gel and probed with a Y' distal fragment.

References

Altschul, S.F., T.L. Madden, A.A. Schaffer, J. Zhang, Z. Zhang, W. Miller and D.J. Lipman, Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res 25: 3389-402.

- Bennett-Lovsey, R.M., A.D. Herbert, M.J.E. Sternberg and L.A. Kelley. 2008. Exploring the extremes of sequence/structure space with ensemble fold recognition in the program Phyre., *Proteins* 70: 611-25.
- Butler, G., C. Kenny, A. Fagan, C. Kurischko, C. Gaillardin and K.H. Wolfe. 2004. Evolution of the MAT locus and its Ho endonuclease in yeast species. *Proc Natl Acad Sci U S A* 101: 1632-7.
- Byrne, K.P. and K.H. Wolfe. 2005. The Yeast Gene Order Browser: combining curated homology and syntenic context reveals gene fate in polyploidy species. *Genome Res* 15:1456-61.
- Callebaut, I., G. Labesse, P. Durand, A. Poupon, L. Canard, J. Chomilier, B. Henrissat and J.P. Mornon. 1997. Deciphering protein sequence information through hydrophobic cluster analysis (HCA): current status and perspectives. *Cell Mol Life Sci* 53: 621-45.
- Dujon, B. 2010. Yeast evolutionary genomics. *Nat Rev Genet* 11 : 512-24.
- Génolevures Consortium, Souciet, J.L., B. Dujon, C. Gaillardin, M. Johnston, P.V. Baret, P. Cliften, D.J. Sherman, J. Weissenbach, E. Westhof, P. Wincker, C. Jubin, J. Poulain, V. Barbe, B. Ségurens, F. Artiguenave, V. Anthouard, B. Vacherie, M.E. Val, R.S. Fulton, P. Minx, R. Wilson, P. Durrens, G. Jean, C. Marck, T. Martin, M. Nikolski, T. Rolland, M.L. Seret, S. Casarégola, L. Despons, C. Fairhead, G. Fischer, I. Lafontaine, V. Leh, M. Lemaire, J. de Montigny, C. Neuvéglise, A. Thierry, I. Blanc-Lenfle, C. Bleykasten, J. Diffels, E. Fritsch, L. Frangeul, A. Goëffon, N. Jauniaux, R. Kachouri-Lafond, C. Payen, S. Potier, L. Pribylova, C. Ozanne, G.F. Richard, C. Sacerdot, M.L. Straub and E. Talla. 2009. Comparative genomics of protoploid Saccharomycetaceae. *Genome Res* 19:1696-709.

		Α	B
Saccharomyces_cerevisiae_A Saccharomyces_cerevisiae_B Candida_glabrata_A Candida_glabrata_B Kazachstania_africana_A Kazachstania_africana_B Naumovozyma_dairenensis_A Naumovozyma_castellii_A Naumovozyma_castellii_B Tetrapisispora_phaffii_A Tetrapisispora_phaffii_B Vanderwaltozyma_polyspora_A	677 773 479 574 680 775 526 621 636 621 636 731 471 566 557 652	DLFIKS.IDTAKDLIAKNTSR AQYFKS.IEVARDLISKSTR ULFKKS.IDTARELLLKNTTR VQYFKS.IEVAMDLIQKSDR DLFENS.IKTAKDLLAKNTTR DQFFKT.LEVAKDLFSKSSR AQYFKS.LDTAKNLLAKNSSR AQYFKS.LDTAKALLAKNSSR KLFNKA.INSAKDLLSKSSR EQFFKS.INLARALISKSSR KLFDKS.LSIARELVKNSAR	TEMTKRILYRHLDNLVSYKNFFEDGR. SEDAKRKITRFLNIIEEFRKDIDTGF. SVITKNQLNQYLNTVLIYKQNFESGK. SMDVKWRLLEMLKQLERFRIDIDEGK. TEMTKRMLYGHLDRLLKYKQNFEHGR. SSDVKSAVVKILDYFAILRKDIDVGI. TELTKRMLYSHMDLLKNYKKIFEDGR. TTDVKFKIYDSLTSLEQFRSSIDTGI. TEFIKRMLYSNLDLRSYKEYFEANK. PPDVKFKIYDSLKVLEELRSIDAGI. TDTLKLQMISHLDYILKYKHEFESGN. PLEIKYQYFDNLLMIENFRKGIDSGI. SVATKQQLYDNINRLKYKENFEDGN.
Zygosaccharomyces_rouxii_A Zygosaccharomyces_rouxii_B Torulospora_delbrueckii_B Torulospora_delbrueckii_B Kluyveromyces waltii A	274 369 275 370 498	NGFDRS.LNTAVDLLQKNETR TQYYKS.IDVARDLVRNSNR DPFTKS.LDTAKELLA.KNTSR AOYYKS.LEVARELVRNSNR TIFYKS.INVARDLIQ.KNSVR	SLPVKRQLYQQMDVLKRFKINFENGA. ELSVKAR TELLNILKNLRQYIDAGL. TEMTKRTLYQQLDVLRNYRDHFQRGR. HFAIKVRTADSLNVLLDLROSIDAGF. SEQNKRLMNQHLDILESFKRSCDRGV.
Lachancea_thermotolerans_A Kluyveromyces_lactis_A Kluyveromyces_lactis_B Kluyveromyces_lactis_C Kluyveromyces_lactis_C Kluyveromyces_lactis_E Ashbya_gossypii_B Ashbya_gossypii_B Ashbya_gossypii_C Eremothecium_cymbalariae_B Eremothecium_cymbalariae_D Ashbya_gossypii_D Kluyveromyces_lactis_F	506 222 317 407 591 407 528 615 426 547 634 747 726 740	DIFFKS.IGVARDLILKNTVR NLFVET.INKVKELLRNTSR SKFKSH.IDQVFSLADKIHR TKVFQEIHAVKSSIL.EKSLS EYFLQS.ITKMKRRLH.ELKSSR GIFKRT.CMKALELLEASDR AAFFYT.ISKVRSILASSNR TGALKN.ISDARDLIAQNTKR LVFNNV.QKVRLLLGANR KAFFQT.ISKVRSILASLNR HGAMRN.IIDARKLIESNTKR KYYYQS.IQNAVDIYNTPFNQSR MRYFQS.IQQAVSTYSTALNRAR LRKLQDEIEEAILTLSTNHSLDK	SDQNKMLMNÄYLDVVENFKRSCDKGV. SEDNKIY NTLCDRLLAFESEIFYEQ. PDKEKHQTKHYASIVLDGYRNRSQYP. SDRASKLFKDLDLLGFLEFTEGEE. SPADMSRILQSLDCLVFFIEKYHEDKQ TIHEKANTLTLIQTLEKGIIRAMDMS. SQENRQAVQRLIKEVIRLEEHFLSK. SPELKTDINNNLLAIEAYQQLVFCGL. SAETKQKTYNLDEVEKTIKETMQNV. SAENKQLIRDKIAVIVGFENNIHLKK. PNDTKAL NKHLSVIESYQELIDNGL. TLEAKQQTYKLLDSIRNAVIETMAGV. LGHEKTKQIILALECFNYFRKKFSA. LGYDRVRILLALEALRAYRELFHRA. MAQDTIVSALSNILSLLKISEEFNST.
		C	D
Saccharomyces_cerevisiae_A Saccharomyces_cerevisiae_B Candida_glabrata_A Candida_glabrata_B Kazachstania_africana_A Kazachstania_africana_B Naumovozyma_dairenensis_A Naumovozyma_dairenensis_B Naumovozyma_castellii_A Naumovozyma_castellii_B Tetrapisispora_phaffii_A Tetrapisispora_phaffii_B Vanderwaltozyma_polyspora_A Vanderwaltozyma_polyspora_B		C SLIDINRRHVAHESAQILFT PPTPLKREGVGKAVVGLRQ YCDRPMRRACADATQALFH PPTTSDKENIGRNVIQLKQ NCPQFMMNECRDSAELFS PPSYEEKIKASQAVLLLKK PCSVDLRKRARESAETLFS PPTLESKIKIGKEIQELKA FISRELRTKVREAAELLFA PLTFESKRKVSTAVLSLKD RLSAHLRWSCKDGAEHLFL PPNYMLKGHIGKAVVELKE PPSHVLKGHIGKAIIRLKE	D NGVKMFIVFETLQDYGIKF 761 QGLKMDRLYENLRRYKIPI 855 NGVKMFUVFETLQDYGIKY 562 QGLKMEKLYENLKIYGVPI 656 NGVKMFVVFETLEDYGITF 763 HGLKMEKLYENLKRYKIAT 857 NGVKMFIVFETLQDYGIRF 609 QGLKMEKLYANL EXGVCT 703 NGVKMFIVFETLQDYGVFY 659 HGLKMNKLYMNLEKYGIAI 813 NGLKMFIVTELLQDYGIIF 554 QGLKMEKLYENLKVYSIPT 648 NGVKMFVVNELLQDYGIVF 640 LGLKMEKLYDNLAVYGVPS 734
Saccharomyces_cerevisiae_A Saccharomyces_cerevisiae_B Candida_glabrata_A Candida_glabrata_B Kazachstania_africana_A Kazachstania_africana_B Naumovozyma_dairenensis_B Naumovozyma_castellii_A Naumovozyma_castellii_B Tetrapisispora_phaffii_A Tetrapisispora_phaffii_B Vanderwaltozyma_polyspora_B Zygosaccharomyces_rouxii_A Zygosaccharomyces_rouxii_B Torulospora_delbrueckii_B Kluyveromyces_waltii_A		C SLIDINRRHVAHESAQILFT PPTPLKREGVGKAVVGLRQ YCDRPMRRACADATQALFH PPTTSDKENIGRNVIQLKQ NCPQFMMNECRDSAELFS PPSYEEKIKASQAVLLLKK PCSVDLRKRARESAETLFS PPTLESKIKIGKEIQELKA FISRELRKVREARELFA PLTFESKRKVSTAVLSLKD RLSAHLRWSCKDGAEHLFL PPNYMLKGHIGKAIVELKE PPSHVLKGHIGKAIIRIKE ASDILQRSRVRDAAEFLFR PPTTLLQQVGKAVLELKD PPKPAFKLEVGRSCVELOE PPKPAFKLEVGRSCVELOE	D NGVKMFIVFETLQDYGIKF 761 QGLKMDRLYENLRRYKIPI 855 NGVKMFLVFERLQDYGIKY 562 QGLKMEKLYENLKIYGVPI 656 NGVKMFVFETLEDYGITF 763 HGLKMEKLYENLKRYKIAT 857 NGVKMFIVFETLQDYGIRF 609 QGLKMEKLYENLKRYKIAT 857 NGVKMFIVFETLQDYGIRF 609 GLKMEKLYENLKYGIAI 813 NGVKMFIVFETLQDYGIFF 554 QGLKMEKLYENLKVYSIPT 648 NGVKMFLVTELLQDYGIFF 554 QGLKMEKLYENLKVYSIPT 648 NGVKMFLVYETLQDYGIVF 640 LGLKMEKLYENLAVGINF 357 QGLKMEKLYENLAVGINF 357 QGLKMEKLYENLAVGINF 451 NGVKMFLVYETLQDYGIVF 452 HGVKMFLYENLAVGINF 452

Figure S1



Supplementary table 1. Yeast strains used in this study. All strains are from the W303-1a background (*ade2-1 trp1-1 ura3-1 leu2-3,112 his3-11,15 can1-100 RAD5*).

_

Strains	Genotype
Lev346	MATa
RL71	$MATa$ uls1- Δ :: $klLEU2$
212-12a	MATa hml-Δ::NAT
213-4a	MATa hml-Δ::NAT
213-7b	$MATa \ hml-\Delta::NAT \ uls1-\Delta::klLEU2$
213-9c	$MATa \ hml-\Delta::NAT \ uls1-\Delta::klLEU2$
195-28a	MATa uls1-Δ::klLEU2 lif1-Δ::skHIS3
210-3d	$MATa \ hml-\Delta::NAT$
211-1a	$MATa \ hml-\Delta::NAT$
211-8d	$MATa \ hml-\Delta::NAT \ uls1-\Delta::klLEU2$
211-10b	$MATa hml-\Delta::NAT uls1-\Delta::klLEU2$
209-1c	$MATa hml-\Delta::NAT pACE1-UBR1 pACE1-ROX1 rap1-(\Delta)::KANr$
209-2b	MATa pACE1-UBR1 pACE1-ROX1 rap1-(Δ)::KANr
206-2b	$MATa \ hml-\Delta::NAT$
205-9a	$MATa \ hml-\Delta::NAT$
206-1d	$MATa hml-\Delta::NAT uls1-\Delta::klLEU2$
205-14c	$MATa hml-\Delta::NAT uls1-\Delta::klLEU2$
199-3a	MATalpha lys2::pGAL1-ISCEI uls1-C1333S
200-2c	MATalpha uls1-K975R
200-5d	MAT a uls1-K975R

- 196-11c MATalpha bar1- Δ
- 196-13b MATalpha bar1- Δ
- 196-5a MATalpha bar1- Δ uls1- Δ ::klLEU2
- 196-6a MATalpha bar1- Δ uls1- Δ ::klLEU2
- Lev791 MATalpha bar1- Δ uls1- Δ 677-855
- Lev792 MATalpha bar1- Δ uls1- Δ 677-855
- 210-5b MATa hml-Δ::NAT uls1-Δ::klLEU2
- RL179 MATa bar1-Δ lys2::GAL-ISCEI ISceI-URA3-ISceI
- RL183 MATa bar1-Δ lys2::GAL-ISCEI ISceI-URA3-ISceI uls1-Δ::klLEU2
- RL185 MATa bar1-Δ lys2::GAL-ISCEI ISceI-URA3-ISceI uls1-Δ::klLEU2 smt3-3R::TRP1
- RL181 MATa bar1-Δ lys2::GAL-ISCEI ISceI-URA3-ISceI smt3-3R::TRP1
- RL266 $MATa \ bar1-\Delta \ lys2::GAL-ISCEI \ ISceI-URA3-ISceI \ uls1-\Delta::klLEU2 \ siz1-\Delta::HPH$
- RL267 MATa bar1-Δ lys2::GAL-ISCEI ISceI-URA3-ISceI uls1-Δ::klLEU2 siz2-Δ::NAT
- RL268 MATa bar1-Δ lys2::GAL-ISCEI ISceI-URA3-ISceI siz1-Δ::HPH
- RL269 MATa bar1-Δ lys2::GAL-ISCEI ISceI-URA3-ISceI siz2-Δ::NAT
- 169-1c *MATa pACE1-UBR1 pACE1-ROX1 rap1-(Δ)::KANr*
- 169-15c $MATa \ bar1-\Delta \ pACE1-UBR1 \ pACE1-ROX1 \ rap1-(\Delta)::KANr$
- 169-6b $MATa \ pACE1-UBR1 \ pACE1-ROX1 \ rap1-(\Delta)::KANr \ smt3-3R::TRP1$
- 169-11a MATa bar1- Δ pACE1-UBR1 pACE1-ROX1 rap1-(Δ)::KANr smt3-3R::TRP1
- 195-19d MATalpha
- 195-3a MATalpha rap1-DAmp::KANr
- 195-7b *MATalpha uls1-*Δ::klLEU2
- 195-8a MATalpha rap1-DAmp::KANr uls1-Δ::klLEU2
- 195-6b MATalpha rap1-DAmp::KANr uls1- Δ ::klLEU2 lif1- Δ ::skHIS3
- 195-1d MATalpha rap1-DAmp::KANr uls1-Δ::klLEU2 smt3-3R::TRP1

- 198-16a *MATa* uls1-Δ::klLEU2 rap1-Δ2-228
- 198-19a *MATa* uls1-Δ::klLEU2 rap1-Δ2-228
- 196-1a MATalpha bar1- Δ uls1- Δ ::klLEU2 rap1- Δ 2-309
- 196-7c MATalpha bar1- Δ uls1- Δ ::klLEU2 rap1- Δ 2-309
- 210-2d *МАТа hml-Δ*::NAT
- 210-3d *МАТа hml-Δ*::*NAT*
- 210-5b MATa hml-Δ::NAT uls1-Δ::klLEU2
- 210-4c $MATalpha hml-\Delta::NAT uls1-\Delta::klLEU2$
- 210-10b $MATa hml \Delta :: NAT uls 1 \Delta :: klLEU2 rap 1 K246R$
- 210-5d MATalpha hml- Δ ::NAT uls1- Δ ::klLEU2 rap1-K246R
- 210-1b *MATa hml-Δ::NAT rap1-K246R*
- 210-7b *MATa hml-*Δ::*NAT rap1-K246R*
- 212-2d *МАТа hml-Δ*::NAT
- 212-10b *МАТа hml-Δ*::*NAT*
- 212-1a MATalpha hml- Δ ::NAT uls1- Δ ::klLEU2
- 212-8a $MATa hml-\Delta::NAT uls1-\Delta::klLEU2$
- 212-5b $MATalpha hml-\Delta::NAT uls1-\Delta::klLEU2 rap1-K240R, K246R$
- 212-17d *MATa hml-Δ::NAT uls1-Δ::klLEU2 rap1-K240R*, *K246R*
- 212-13c *MATa* hml-Δ::NAT uls1-Δ::klLEU2 rap1-K240R, K246R
- 212-3b *MATa hml-Δ::NAT rap1-K240R*, *K246R*
- 212-4c *MATa hml-Δ::NAT rap1-K240R*, *K246R*
- 184-48c MATalpha bar1- Δ hmr- Δ ::NAT uls1- Δ ::klLEU2
- 184-9d MATalpha bar1- Δ hmr- Δ ::NAT uls1- Δ ::klLEU2
- 183-30d MATalpha bar1- Δ hmr- Δ ::NAT rif2- Δ ::HPH
- Lev 575 $MATa hml \Delta :: NAT sir 4 \Delta :: HPH$

- Lev576 $MATalpha hmr-\Delta::NAT sir4-\Delta::HPH$
- 184-14c MATalpha bar1- Δ hmr- Δ ::NAT rif2- Δ ::HPH sir4- Δ ::HPH
- Lev601 $MATa hml-\Delta::NAT rif2-\Delta::skHIS3 sir4-\Delta::HPH$
- 183-10c MATalpha bar1- Δ hmr- Δ ::NAT uls1- Δ ::klLEU2 sir4- Δ ::HPH
- 183-2d MATalpha bar1- Δ hmr- Δ ::NAT uls1- Δ ::klLEU2 rif2- Δ ::HPH
- 183-25a MATalpha bar1- Δ hmr- Δ ::NAT uls1- Δ ::klLEU2 rif2- Δ ::HPH sir4- Δ ::HPH
- 184-49a MATalpha bar1- Δ hmr- Δ ::NAT uls1- Δ ::klLEU2 rif2- Δ ::HPH sir4- Δ ::HPH
- 205-9a *MATa hml-*Δ::*NAT*
- 205-14c $MATa hml-\Delta::NAT uls1-\Delta::klLEU2$
- 212-3c $MATa hml \Delta::NAT uls1 \Delta::klLEU2 sir4 \Delta::HPH$
- 212-4b $MATa hml-\Delta::NAT uls1-\Delta::klLEU2 sir4-\Delta::HPH$
- 205-16a $MATa hml-\Delta::NAT uls1-\Delta::klLEU2 sir4-\Delta::HPH smt3-3R::TRP1$
- 212-1b $MATa hml-\Delta::NAT uls1-\Delta::klLEU2 sir4-\Delta::HPH rap1-K240R, K246R$
- 212-2a $MATa hml-\Delta::NAT uls1-\Delta::klLEU2 sir4-\Delta::HPH rap1-K240R, K246R$
- 207-14d MATa hml- Δ ::NAT uls1- Δ ::klLEU2 sir4- Δ ::HPH rap1- Δ 2-309