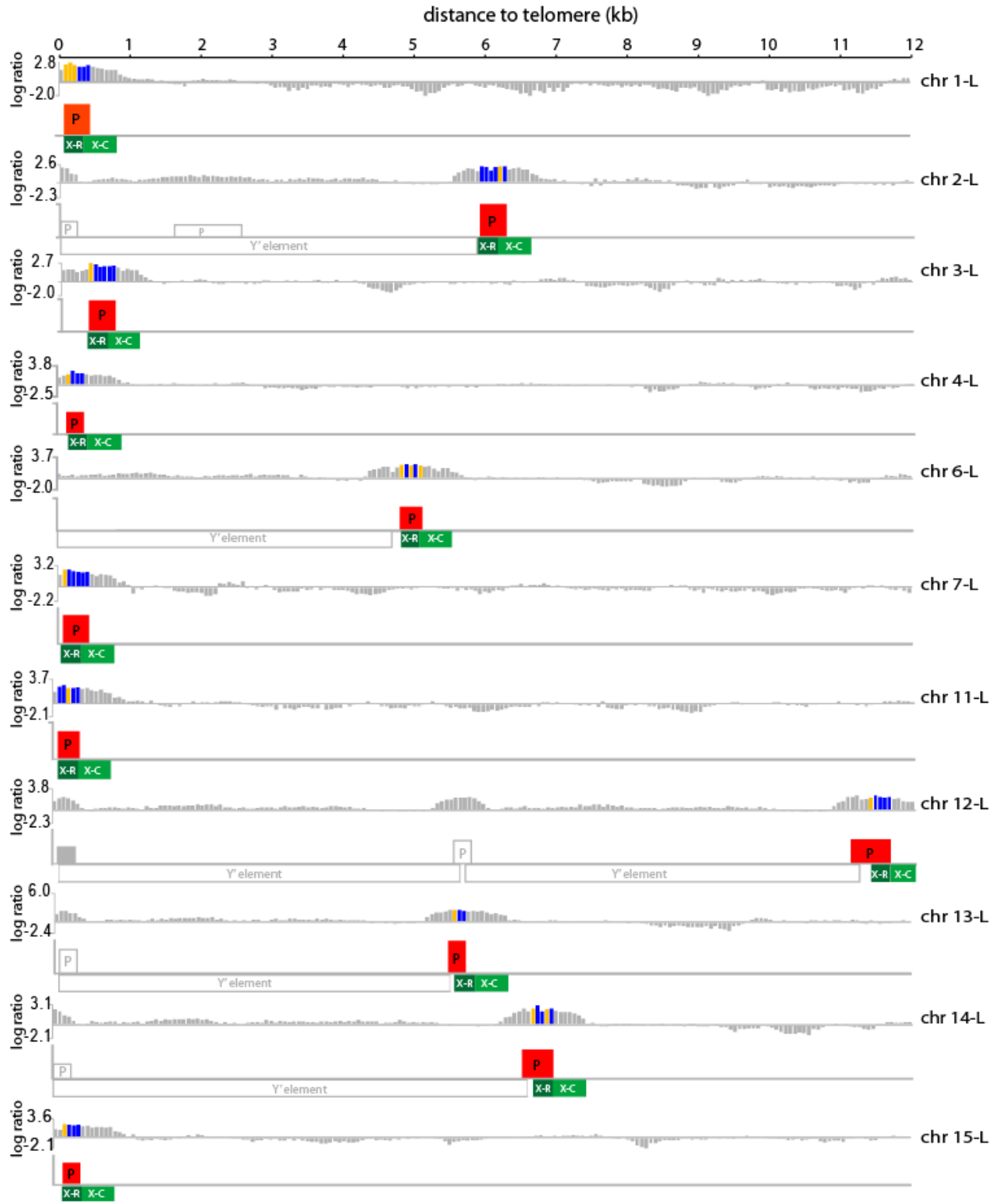Supplementary Information for Smith *et al*.

Environment-responsive transcription factors bind subtelomeric elements and regulate gene silencing

Table of Contents

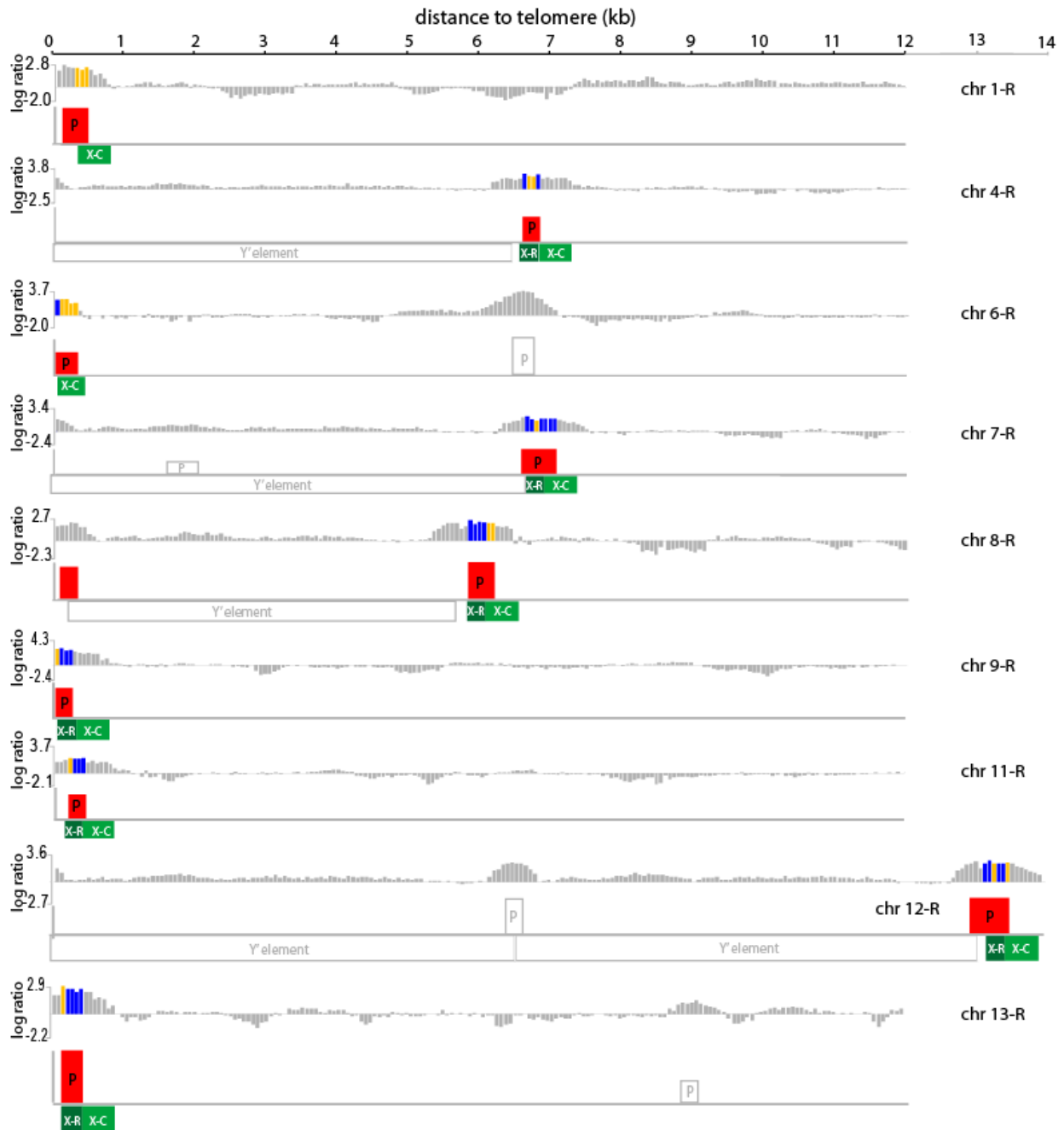Smith_Figure S1: Analysis of Oaf1p binding specificity to X elements (related to Figure 4)

Figure S1. Analysis of Oaf1p binding specificity to X elements (Related to Fig. 4). X elements have homology with each other; therefore, we investigated experimental noise in the microarray data. 50mer tiling array probes that are in peak regions of Oaf1p binding (*red box marked P*; determined using Peak Find of Nimblescan software) and overlap with X element repeat or core regions (*green boxes* marked *X-R* and *X-C*) were tested for homology to the entire yeast genome using FASTA (at SGD website). Log$_2$ Oaf1p enrichment ratios for probes that were tested are colored and those with homology levels that are predicted to have no significant background binding to other genomic sequences by tiling microarray analysis (Deng *et al*., 2008) are *gold* colored. The 20 X elements shown contain probes predicted to bind specifically, suggesting that Oaf1p interaction with X elements is specific and not due to experimental noise.
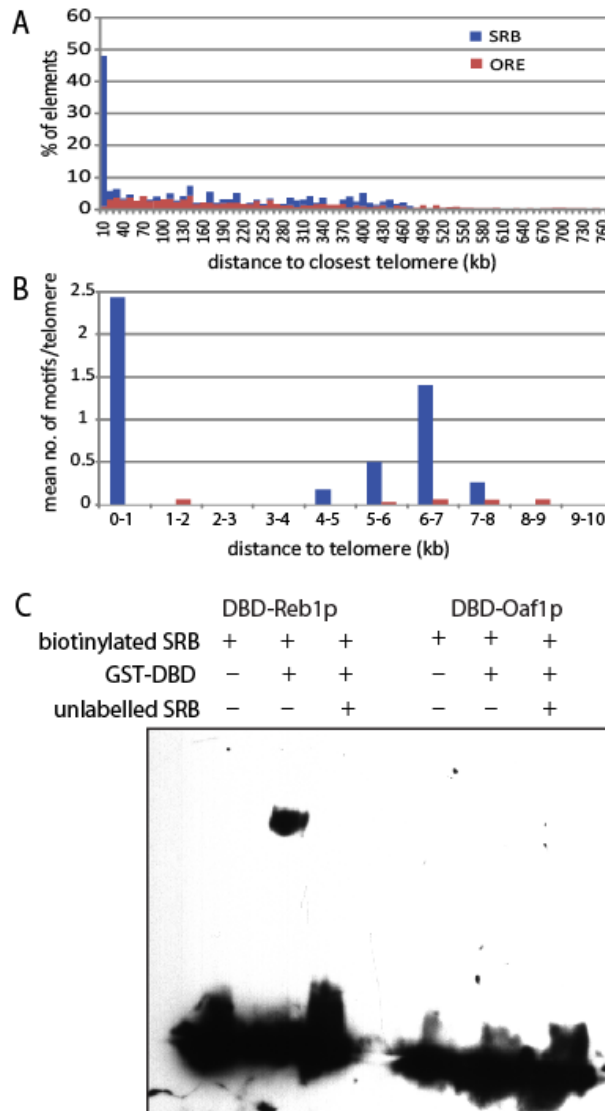
Figure S2. Characterization of context-specific binding of Oaf1p. High resolution ChIP-chip data were used to identify a putative DNA-recognition sequence for Oaf1p in the context of negative regulation, which was termed SRB motif (see Materials and Methods section in the manuscript and Supplemental Experimental procedures section below). *A and B* Profile of positions of subtelomeric Reb1-binding (SRB) motifs matches profile of X element positions. Histograms are shown of distance to closest telomere for oleate response elements (OREs) that bind Oaf1p-Pip2p heterodimers (Rottensteiner *et al*, 2003) and SRB motifs in the entire genome (*A*) and within 10 kb of telomeres (*B*). The mean number of OREs and SRBs in the whole genome are 0.067 elements/kb and 0.020 elements/kb, respectively. (*C*) Electromobility shift assay (EMSA) showing that the DNA-binding domain of Reb1p bound to the biotinylated SRB domain resulting in a shift in the mobility of the probe. This shift was not observed for the DNA-binding domain of Oaf1p.

Smith_Figure S3. Conditional regulation of proto-silencing by X element-binding TFs with direct targets omitted from analysis (Related to Fig. 5)
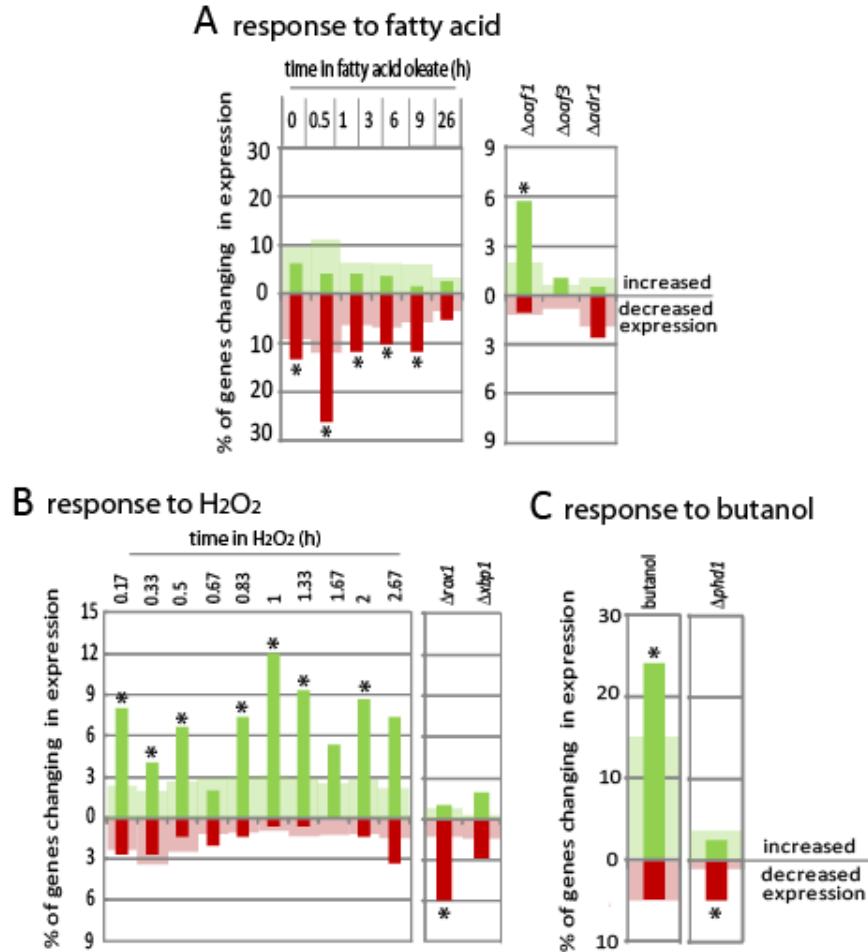

Figure S3. Conditional regulation of proto-silencing by X element-binding TFs with direct targets omitted from analysis (Related to Fig. 5). Microarray expression data were analyzed to determine the effects of environmental stresses on the expression of subtelomeric genes that are centromere-proximal to X elements and the roles of X-element binding factors in this regulation. The analysis was done as in Fig. 5 except that genes with start sites adjacent to X elements and direct targets of factors implicated in subtelomeric gene regulation were omitted from the analysis. Targets that were omitted in addition to those with start sites adjacent to X elements are listed below: For *panel A*, direct targets of Adr1p, Oaf1p and Oaf3p, the three factors with enriched binding at subtelomeres in the presence of fatty acids (Fig 1*A*), were omitted. For the H_2O_2 time course in *panel B*, direct targets of factors that enrich at subtelomeres in the presence of 4 mM H_2O_2 (Rox1p, Mal33p, Yap6p and Yjl206p) or 0.4 mM H_2O_2 (Aft2p, Xbp1p, Nrg1p, Yap6p and Yjl206p) (Mak *et al*, 2009) were omitted. For analysis of deletion of *ROX1* and *XBP1*, direct targets of each factor in H_2O_2 were omitted. For *panel C*, direct targets of Phd1p, which enriches at subtelomeres in the presence of 1% butanol (Mak *et al*, 2009) were omitted. For description of Figure see the legend of Fig. 5, except *asterisks* denote bars with Sudent's *t*-test *P* values < 0.01.
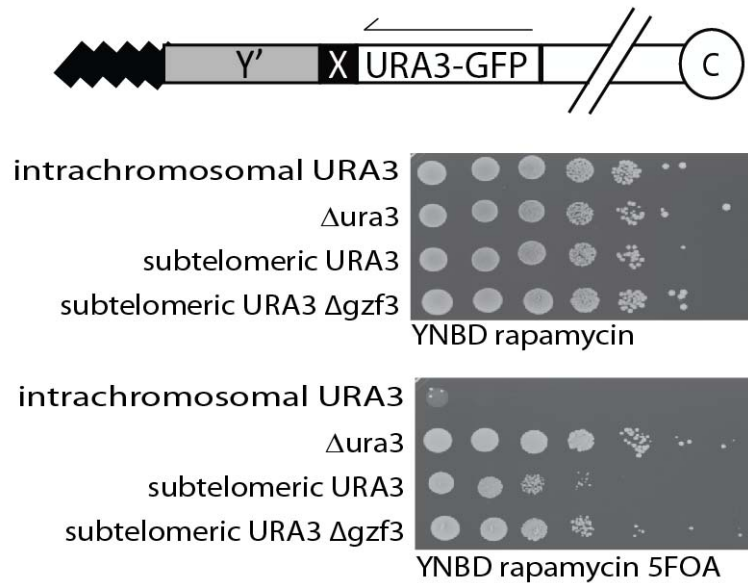
Figure S4. 5FOA subtelomeric silencing assay showing Δ*gzf3* increases subtelomeric silencing in the presence of rapamycin. Gzf3p, a TF that negatively regulates nitrogen catabolic gene expression (Soussi-Boudekou et al., 1997), conditionally interacts with X elements in the presence of rapamycin (Fig. 2). A strain with Ura3-GFP encoded on subtelomere XI-L (FEP318-19) is compared to an isogenic strain with Δ*gzf3*, and control strains with and without intrachromosomal URA-GFP (PIY125 and FYBL1-8B, respectively). 10 fold serial dilutions of each strain were spotted onto rapamycin-containing medium in the presence and absence of 5FOA. Deletion of *GZF3* improved colony viability suggesting that interaction of Gzf3p with X elements reduces subtelomeric silencing. Deletion of *GZF3* did not improve the growth of the strain with intrachromosomal URA3-GFP (PIY125) (*data not shown*), demonstrating that the phenotype is specific to the role of Gzf3p in subtelomeric gene expression.

Supplemental Experimental Procedures

Analysis of experimental noise in ChIP-chip data (Procedure for Fig. S1)

X elements have homology with each other (76 - 93 % identity using Megablast at NCBI website), which might result in experimental noise in the microarray data; therefore background signal in the ChIP-chip data was investigated to determine if the interaction of X elements with Oaf1p is specific. The tiling arrays are made up of short, densely packed probes, enabling the identification and removal of regions that might cause background from the analysis, while, still retaining data within X elements as described below: A comparison of 50mer probe sequences to the entire yeast genome using FASTA program (at SGD website) determined that Oaf1p binding peaks at X elements usually include probes that are sufficiently unique to have no significant background binding to other genomic fragments in the experiment. Probes were considered to be unique if they shared < 92% identity with any other 50 base sequences in the entire genome. This cutoff was established empirically in previously published microarray control studies using the same 50mer probe length and hybridization temperature of 42°C (Deng et al., 2008). These data (shown in Fig. S1) indicate that the interaction is not due to experimental noise, and support the conclusion that Oaf1p indeed binds most X elements. This conclusion is supported by the fact that genes modulated by oleate or *OAF1* deletion (in the analysis of Fig. 5) were identified on almost all of the 32 chromosome ends (*data not shown*).

Characterization of context-specific binding of Oaf1p (procedure for Fig. S2)

Notably, chromosome arms with Y' elements have interesting Oaf1p binding profiles. In addition to peaks coinciding with X elements, these chromosome arms also display peaks at the telomere-proximal end of the subtelomeric region (for example, see regions shown in Fig. 6*A*). This pattern is consistent with the hypothesis that interactions can form between X element binding proteins and telomeres (through Rap1p/Sir complex interactions), causing telomeres to form a folded structure that can both insulate Y' elements from silencing and facilitate spreading of silencing towards centromeres (Pryde and Louis, 1999; Strahl-Bolsinger et al., 1997). These data suggest that indirect interactions are detectable by this method of analysis and raise the issue that the Oaf1p interaction with X elements might be mediated other X element binding proteins.

To gain more insight into the nature of the interaction, the tiling array data were analyzed to identify a potential Oaf1p binding motif in the context of negative regulation (the activity suggested for Oaf1p at X elements in Table I, and Figures 5 and S3).

Oaf1p is known to heterodimerize with Pip2p and bind to oleate response elements (OREs) of the sequence $CGGNNNTN(A/G)N_{8-12}CCG$ (Rottensteiner et al., 2003). In this context, Oaf1p is an activator; however, the DNA motif recognized by Oaf1p in the context as a repressor (without Pip2p) has not been characterized. To characterize the negative regulatory motif, we used network structure data (Smith et al., 2007) to narrow the search space of the motif analysis of high density Oaf1p ChIP-chip data as described below. DNA sequences corresponding to Oaf1p binding peaks identified from one biological replicate (using Nimblescan software with FDR threshold <0.001) were determined. Next, duplicate peaks with the lowest peak value score were removed, and then peaks were selected that overlapped with intergenic regions previously shown to interact with Oaf1p in the context of negative regulation (i.e. bound by Oaf1p, Adr1p and Oaf3p, but not Pip2p in the presence of fatty acid). These 44 sequences were used for motif searches with AlignACE (Roth et al., 1998) (with default parameters) and MEME version 4.1.1 (Bailey and Elkan, 1994) (with expected frequency of 0-1 sequences per peak, a maximum of 6 motifs identified, and motif widths of 6-24). The first motif identified by AlignACE (with map score of 254), overlapped with 3 of the 6 motifs identified by MEME (with E-values $< 1 \times 10^{-24}$). The consensus sequence derived from the four motifs (G[not A]AGGGTAANNNNN[not C][not C]) was chosen as the putative Oaf1p recognition sequence and termed SRB motifs. The entire genome was searched for this motif using Fuzznuc of Emboss software (Rice et al., 2000). The overlap of the motif with Oaf1p binding peaks was determined with the Map Peaks tool of Nimblescan software. Of the 238 SRB motifs in the entire genome, 155 (65%) were within 100 bp of an Oaf1p binding position identified in the tiling array data in the presence of fatty acids.

Chromosome position analysis was performed on all ORE and SRB elements in the genome (Fig. S2). Almost 50% of SRBs were found within 10 kb of telomeres, whereas ORE elements showed no specific obvious positional enrichment (*panel A*). A higher resolution histogram of the subtelomeric region shows enrichment of SRBs within 1 kb and also between 5-7 kb from telomeres (*panel B*); matching the Oaf1p binding profile in subtelomeric regions (Fig. 4). This analysis also revealed that SRB motifs are found in clusters of 4-7 motifs, whereas OREs are

usually found singly (for example see Figure 6*A*). This configuration of multiple binding sites for sequence specific factors is also found for some GRF that regulate heterochromatin formation (Gasser, 2001).

EMSA was performed using LightShift Chemiluminescent EMSA Kit (Thermoscientific, Rockford, IL) with previously reported conditions for Reb1p binding (Chasman et al., 1990). DBDs of Oaf1p (227 aa) and Reb1p (483 aa) used in the analysis were generated as GST-fusion proteins and purified using GST Gene Fusion system (GE Healthcare (Piscataway, NJ). To construct expression plasmids, Oaf1p and Reb1p DNA fragments were amplified from genomic DNA by PCR with oligonucleotides (AAGGGATCCGGAAATGATGATAATA and TTGCTCGAGGGTATCATCGTGTT) and (AAGGGATCCCTCAACAAATCTAG and TTGCTCGAGGGAATTAATTTTCTG), respectively, and ligated into pGEX-4T1 with BamHI and XhoI. Double stranded target DNA in the EMSA was ACCTCC<u>CCACTCGTTACCCTGC</u>CCCACT, which is found in Oaf1p binding peak in X element on chromosome arm 14R and contains an SRB domain (*underlined*).

Analysis of conditional regulation of proto-silencing by X element-binding TFs with direct targets omitted (Procedure for Fig. S3)

The motive for this revised analysis is that some TFs conditionally enrich at subtelomeres in response to stimuli (Mak *et al*, 2009) and may regulate gene expression through direct interaction with a promoter region, rather than through regulation of proto-silencing at X elements. The analysis was performed exactly as described in the Experimental Procedures Section for Fig. 5 except the subtelomeric gene group (within 20 kb centromere-proximal to X elements) excluded genes with start sites adjacent to X elements and genes that are direct targets of factors implicated in conditional control of subtelomeric genes as described below. Genes with start sites adjacent to X elements were: YDL248W, YFL062W, YGL263W, YML132W, YNL336W, YOL166W-A, YBR302C, YCR108C, YER188C-A, YGR295C, YIR042C, YJR161C, YOR394C-A, and YPR202W. Additional genes that were omitted are listed below: For *panel A*, direct targets of Adr1p, Oaf1p and Oaf3p, the three factors with enriched binding at subtelomeres in the presence of 0.15% fatty acids (Fig 1*A*), were omitted. Direct targets of these factors in the presence of fatty acids were those published previously (Smith et al., 2007). For

panels *B* and *C*, targets of factors previously shown to conditionally enrich at subtelomeric regions (Mak et al., 2009) that are relevant to the perturbation analyzed were omitted. For the H_2O_2 time course in *panel B*, these genes include direct targets of Rox1p, Mal33p, Yap6p or Yjl206p in the presence of 4 mM H_2O_2 and direct targets of Aft2p, Xbp1p, Nrg1p, Yap6p or Yjl206p in the presence of 0.4 mM H_2O_2. For analysis of deletion of *ROX1* and *XBP1*, direct targets of Rox1p in 4 mM H_2O_2 and Xbp1p in 0.4 mM H_2O_2, respectively, were omitted. For *panel C*, direct targets of Phd1p in the presence of 1% butanol were omitted. For *panels B* and *C*, direct targets of factors were those identified previously (with P values < 0.01) (Harbison et al., 2004) and were taken from Pvalbyintergenic_9.2_forpaper file (at http://jura.wi.mit.edu/young_public/regulatory_code/files_for_paper.zip). ORFs with adjacent start sites were annotated as published previously (Smith et al., 2007). Analysis of statistical significance of gene expression profiles was performed as described in the Experimental Procedures Section for Fig. 5, except that conditions were scored as having a significant effect on subtelomeric gene expression if they had Student's *t*-test *P* values < 0.01 (because of smaller sample size).

5-FOA viability assay of subtelomeric silencing (Procedure for Fig. S4)

Strains for this assay were a kind gift of Dr. Edward Louis and were described previously (Loney *et al*, 2009). Silencing assays were performed as described previously (Gottschling *et al*, 1990). Briefly, the open reading frame of *GZF3* was disrupted with a Kan-MX cassette in strain FEP318-19 (with subtelomeric URA3-GFP as shown in Figure S4) by homologous recombination of a PCR fragment from Δ*GZF3* (BY4742 deletion library; Invitrogen, Carlsbad, CA). FEP318-19 and FEP318-19 Δ*GZF3* were grown in 5 mL YEPD overnight to saturation along with control strains with and without intrachromosomal URA-GFP (PIY125 and FYBL1-8B, respectively). 10-fold serial dilutions of each strain were spotted (2 μL/spot) on YNBD rapamycin plates (1.7 g yeast nitrogen base without amino acids/L; 5g ammonium sulfate/L, 2% dextrose, 20 mg uracil/L, 20 mg L-histidine-HCl/L, 60 mg L-leucine/L, 50 mg L-lysine/L, 20 g agar/L, and 100 nM rapamycin) with and without 1 g/L 5-FOA (Bio 101, Inc). Images were taken after 72 h of growth at 30°C. The same assay was performed with deletion strains *XPB1*, *YAP6* and *ROX1* in the presence of 2 mM H_2O_2 and *CHA4* in the presence of 0.2 mg sulfometuron methyl/L in place of rapamycin.