HHS Public Access

Author manuscript

Nat Cell Biol. Author manuscript; available in PMC 2010 August 01.

Published in final edited form as:

Nat Cell Biol. 2010 February ; 12(2): 111–118. doi:10.1038/ncb2011.

DNA zip codes control an ancient mechanism for gene targeting to the nuclear periphery

Sara Ahmed[1], **Donna G. Brickner**[1], **William H. Light**[1], **Ivelisse Cajigas**[1], **Michele McDonough**[2], **Alexander B. Froyshteter**[1], **Tom Volpe**[2], and **Jason H. Brickner**[1,3]

[1]Department of Biochemistry, Molecular Biology and Cell Biology, Northwestern University, Evanston, IL 60208

[2]Department of Cell and Molecular Biology, Feinberg School of Medicine, Northwestern University, Chicago, IL 60611

Abstract

Many genes are recruited to the nuclear periphery upon transcriptional activation in *Saccharomyces cerevisiae*. We have identified two Gene Recruitment Sequences (GRS I and II) from the promoter of the *INO1* gene that target the gene to the nuclear periphery. These GRSs function as DNA zip codes; they are sufficient to target a nucleoplasmic locus to the nuclear periphery. Targeting requires components of the nuclear pore complex (NPC) and a GRS is sufficient to confer a physical interaction with the NPC. GRS I elements are enriched in promoters of genes that interact with the NPC and genes that are induced by protein folding stress. Full transcriptional activation of *INO1* and another GRS-containing gene requires GRS-mediated targeting of the promoter to the nuclear periphery. Finally, GRS I also functions as a DNA zip code in *Schizosaccharomyces pombe*, suggesting that this mechanism of targeting to the nuclear periphery has been conserved over approximately one billion years of evolution.

The spatial organization of DNA within the nucleus compartmentalizes the genome into different subnuclear environments that may affect gene expression. Both transcriptionally active and inactive genes localize at the nuclear periphery1–3. The localization of individual genes within the nucleus can also be dynamically controlled. For example, in *Saccharomyces cerevisiae*, many inducible genes rapidly relocalize from the nucleoplasm to the nuclear periphery upon activation4–9.

How genes are targeted from one location to another within the nucleus is unclear. Localization could simply reflect changes in transcriptional status, chromatin structure or the production of nascent RNA. The targeting of certain genes seems to involve nascent RNA

[3]Corresponding author: j-brickner@northwestern.edu.

Author Contributions

S.A., D.G.B, T.V. and J.H.B designed the experiments, S.A., D.G.B., W.H.L., M.M., I.C., A.B.F. and J.H.B performed the experiments, S.A. and J.H.B wrote the manuscript.

Competing Interests

The authors declare no competing interests.

transcripts that might mediate recruitment to the nuclear periphery in yeast6, 7, 10. Alternatively, changes in localization could represent gene targeting, controlled by *cis*-acting DNA elements. Consistent with this idea, peripheral targeting of certain genes in budding yeast is independent of transcription11, 12.

We have approached this problem by studying the mechanism of recruitment of the yeast *INO1* gene to the nuclear periphery. This gene is targeted to the nuclear periphery upon activation5. Localization of *INO1* to the nuclear periphery is controlled by two *cis*-acting DNA elements. These elements function as DNA zip codes that are sufficient to target an ectopic locus to the nuclear periphery. One of these elements is also sufficient to target an ectopic locus to the nuclear periphery in the highly divergent fission yeast, suggesting that this mechanism of targeting is ancient. Finally, we show that the full transcriptional activation of *INO1* requires one of these DNA zip codes in the promoter of the gene. This suggests that the genome can encode for its own spatial organization and that this impacts gene expression.

Results

Identification of a DNA element required for recruitment of *INO1* to the nuclear periphery

We monitored gene localization with respect to the nuclear periphery by expressing the lac-repressor fused to GFP in a strain having an array of lac repressor binding sites integrated at the chromosomal locus of interest (Fig. 1a13, 14). We then quantified the fraction of population in which the GFP spot colocalizes with the nuclear envelope5. A nucleoplasmic locus like *URA3* colocalizes with the nuclear envelope marker in ~27% of cells5 (Fig. 1b, indicated as a hatched blue line throughout). Upon activation by inositol starvation, *INO1* colocalizes with the nuclear envelope in ~60% of cells in the population (Fig. 1b). We first asked if targeting of *INO1* is dependent on chromosomal context by integrating the *INO1* gene and the lac repressor array beside the *URA3* gene (*URA3:INO1*; Fig. 1a). We found that this hybrid locus was targeted to the nuclear periphery upon inositol starvation (Fig. 1b). Therefore, the *INO1* gene was sufficient to confer peripheral targeting of the *URA3* locus.

To identify *cis*-acting subnuclear targeting element(s), we deleted several 100bp segments from the *INO1* promoter sequence (Fig. 1c) and tested their ability to target *URA3* to the nuclear periphery (Fig. 1d). Loss of segment 4 resulted in unregulated peripheral targeting of *URA3:INO1* (Fig. 1d) and unregulated, modest *INO1* transcription (similar to the *INO1-100* mutant15; data not shown). Loss of segment 3 blocked targeting of *URA3:INO1* to the nuclear periphery (Fig. 1d), suggesting that segment 3 contains DNA sequences necessary for targeting *URA3:INO1* to the nuclear periphery.

A DNA element that is required for *INO1* targeting to the nuclear periphery functions as a DNA zip code

We next integrated the 100 base pair segment 3 alone beside *URA3* and found that it was sufficient to target *URA3* to the nuclear periphery (Fig. 1e and 1f). Therefore, segment 3 functioned as a DNA zip-code: a DNA sequence that is sufficient to target an ectopic locus to a particular subnuclear location. When removed from the *INO1* promoter, segment 3-

mediated peripheral localization was no longer regulated by inositol (Fig. 1f). This suggests that the peripheral targeting element is ordinarily negatively regulated in the context of the *INO1* promoter.

To identify a minimal Gene Recruitment Sequence (GRS), we integrated a series of smaller fragments from segment 3 at *URA3* and determined their peripheral targeting activity (summarized in Fig. 1e; complete data in Fig. S1). All of the DNA fragments that were active for peripheral targeting contained a common eight base pair sequence (Fig. 1e). When this eight base pair fragment (GRS I; see below) was integrated in either orientation beside *URA3*, it functioned to target *URA3* to the nuclear periphery (Fig. 1f).

To verify that GRS I is responsible for peripheral targeting of full-length *INO1*, we introduced a transition mutations in GRS I in the *INO1* promoter and tested the effect of this mutation on the peripheral targeting of *URA3:INO1*. Mutation of GRS I blocked targeting of *URA3:INO1* to the nuclear periphery, confirming that it was the element responsible for this relocalization (Fig. 2a).

Peripheral targeting of the endogenous *INO1* gene is mediated by two redundant DNA zip codes

We then introduced the *grs I* mutation into the promoter of the endogenous *INO1* gene. This mutation did not block targeting of *INO1* to the nuclear periphery (Fig. 2b). We hypothesized that additional, redundant targeting elements also contribute to peripheral targeting of endogenous *INO1*. We found that deletion of a 943 base pair region upstream of *INO1* (Δ*up INO1*; Fig. 2d) led to GRS I-dependent targeting of *INO1* (Fig. 2c). This suggested that additional targeting elements exist within this 943 base pair region.

We integrated a series of fragments from this 943 base pair region at *URA3* and tested their targeting activity (Fig. 2d; complete data in Fig. S1b and S1c). We identified a second DNA zip code, GRS II, embedded within the upstream *SNA3* gene. GRS I and GRS II are redundant; mutation of either element alone had no effect on peripheral targeting of *INO1* (Fig. S2a). However, loss of both GRS I and GRS II blocked targeting of endogenous *INO1* to the nuclear periphery (Fig. 2e).

GRS I-containing genes in the *S. cerevisiae* genome

The sequences of GRS I, 5’-GGGTTGGA-3’ and GRS II, 5’-GAATGATTGCTGGGAAGAAT-3’ are not obviously related. GRS I does not correspond to any known binding site16, 17. A sequence within GRS II (5’-TGCTGG-3’) resembles the binding site for the daughter-specific transcription factor Ace218. However, peripheral targeting does not correlate with Ace2-dependent transcription and we have not observed an interaction of the *INO1* promoter with Ace2 *in vivo* using chromatin immunoprecipitation (ChIP; not shown). Therefore, these elements most likely represent previously uncharacterized DNA binding sites.

Perfect matches of GRS I appear 280 times in the yeast genome as a whole and 97 times in 94 promoters (within 1000 bp 5’ of the transcription initiation site). Among these genes, the most significantly overrepresented gene ontology class was “cellular response to heat”

(corrected hypergeometric19, 20 $P = 0.007$). The *INO1* gene is transcriptionally induced not only by inositol starvation, but also by unfolded protein stress in the endoplasmic reticulum (ER)21, heat shock22 and nitrogen starvation22. We asked if genes containing GRS I elements were co-regulated with *INO1* under any of these conditions (Table S1). We observed a significant enrichment of GRS I-containing genes among the genes most highly induced under heat shock and ER stress conditions23 but no significant enrichment of GRS I genes among those highly induced by inositol starvation or nitrogen deprivation (Table S1). The greatest enrichment of GRS I-containing genes was among the >90th percentile of genes induced under combined heat shock + ER stress conditions24. This enrichment was more significant if we limited our analysis to genes in which the GRS I element is less than 775 base pairs upstream of the translational start site (Table I). This suggests that GRS I-containing promoters are significantly enriched for genes that are co-regulated by protein folding stress in the ER and the cytoplasm.

One perfect match of GRS I exists in the promoter of the *TSA2* gene, which encodes an inducible thioredoxin peroxidase that is activated by heat shock and oxidative stress25. We localized the *TSA2* gene and found that it localized in the nucleoplasm in the absence of stress (peripheral in 37% ± 2% of cells; Fig. 3a). In the presence of oxidative stress, *TSA2* localized to the nuclear periphery in 73% ± 3% of cells (Fig. 3a). When we introduced the *grs I* mutation in the *TSA2* promoter, targeting of *TSA2* to the nuclear periphery was blocked (Fig. 3a). This strongly suggests that GRS-mediated targeting is a general mechanism employed by genes in *Saccharomyces cerevisiae*.

GRS I is functional in *Schizosaccharomyces* pombe

We next asked if GRS-mediated peripheral targeting is an evolutionarily conserved mechanism. To address this question, we tested if the GRS I element was sufficient to direct peripheral targeting in the fission yeast *Schizosaccharomyces pombe*. *S. pombe* is distantly related to *S. cerevisiae*, having diverged from a common ancestor between four hundred million and one billion years ago26. We integrated the Lac operator array plasmid with or without a single copy of GRS I at the *ura4* locus in *S. pombe* (Methods). We then quantified the fraction of the population in which the lac operator array colocalized with the nuclear pore protein Nup120 (Fig. 3b). In the strain without the GRS I element, the *ura4* locus was nucleoplasmic, colocalizing with the nuclear envelope in 35% ± 1% of the cells in the population (Fig. 3c). However, in the strain with the GRS I element integrated at *ura4*, we observed an increase in the colocalization of the *ura4* locus with the nuclear envelope to 50% ± 3% of the cells in the population (Fig. 3c). Although this level of peripheral localization was not as high as we had observed for *URA3* in budding yeast, it represents a significant change in the localization of *ura4* ($P = 0.002$ using a one-tailed *t* test). This suggests that the mechanism of GRS I targeting is ancient and conserved between two highly divergent species.

GRS I targets chromosomal loci to the nuclear pore complex

Many studies have suggested that genes that undergo gene recruitment upon activation are targeted to the nuclear pore complex (NPC)4, 6–8, 11, 12, 27–30. We monitored peripheral targeting of *INO1* in a collection of 30 viable null mutant yeast strains lacking proteins that

make up the NPC or that associate with the nuclear periphery (summarized in Fig. 4a; complete data in Fig. S3). The majority of the proteins that make up the core channel31 of the NPC were dispensable for peripheral targeting of *INO1* (Fig. 4a). In contrast, most of the proteins associated with the nucleoplasmic face of the NPC were required for peripheral localization of *INO1* (Fig. 4a). Nup1, proteins in the SAGA complex and proteins in the TREX2 complex are also required for localization of the *GAL1-10* locus to the nuclear periphery upon galactose induction11, 12, 28. However, Mlp1, which is required for recruitment of *GAL1-10*27, *GAL2* and *HSP104*7 to the nuclear periphery, was not required for recruitment of *INO1* to the nuclear periphery (Fig. 4a). Instead, Mlp2, a homologous protein, was required for *INO1* targeting to the nuclear periphery. This suggests that different genes may use overlapping, but distinct, targeting mechanisms.

Tthe NPC protein Nup2 physically interacts with active genes that localize at the nuclear periphery4, 11. To test if the *INO1* promoter physically associates with the NPC under activating conditions, we used chromatin immunoprecipitation (ChIP). Nup2-TAP co-immunoprecipitated with the *INO1* promoter when the gene was active (Fig. 4b). We did not observe an interaction of Nup2-TAP with either the repressed *INO1* promoter or with a nearby intergenic region (Fig. 4b). The interaction of *INO1* with the NPC requires the GRS elements; we did not observe an interaction of the *INO1* promoter with Nup2-TAP when both GRS I and GRS II were mutated (Figs. S2c & S2d).

NPC mutants that blocked targeting of *INO1* to the nuclear periphery also blocked targeting of *URA3* to the nuclear periphery by GRS I alone (Fig. 4c). Furthermore, the GRS I element at *URA3* was sufficient to confer an interaction with Nup2-TAP by ChIP (Fig. 4d). This suggests that the GRS elements control targeting of *INO1* to the NPC and that the interaction of *INO1* with the NPC observed by ChIP is mediated by DNA elements and is not a result of post-transcriptional interaction with nascent mRNA.

We next asked if GRS I-containing promoters were enriched among genes that have been shown to physically associate with the NPC by ChIP4. Using Chi square analysis, we found that GRS-containing promoters were significantly enriched among genes that associate with the nuclear pore proteins Nup2, Mlp1, Mlp2, Nup60 and Nup116 and the transport factors Cse1 and Xpo1 (Table II). As a control, we performed the same analysis with a reversed GRS I (GRS I_{rev}; 5'-AGGTGGG-3'). We observed no enrichment of GRS I_{rev} containing promoters among genes that interact with NPC proteins (not shown). Furthermore, we noticed that the GRS I from *INO1* and *TSA2* is related to a sequence motif that was previously found to be overrepresented in promoters of NPC-associated genes4 (Fig. 5). This suggests that GRS I-like elements control the interaction of many genes with the NPC.

Peripheral targeting enhances transcription of *INO1* and *TSA2*

We next tested the functional significance of peripheral localization for transcriptional activation of *INO1*. Mutation of the GRS I element of *URA3:INO1* resulted in poor accumulation of *INO1* mRNA after 3 hours of induction (Fig. 6a). At steady state, we observed a two-fold difference in mRNA levels (Fig. S4a). This decrease did not correlate with a difference in the rate of mRNA decay, suggesting that it is due to a difference in

transcription (Fig. S4b). Furthermore, we also observed a similar defect in the activation of *grs I* mutant *TSA2* relative to wild type *TSA2* (Fig. S4c).

We have noticed that plasmid-borne *INO1*, integrated either at *URA3* or in place of the endogenous gene, is regulated normally by inositol starvation but is expressed to higher levels than endogenous *INO1* (Fig. S4d). For this reason, we also compared the transcription of wild type with *grs* mutant forms of *INO1* after introducing chromosomal mutations to remove GRS I and II at the endogenous locus. We observed a similar decrease in *INO1* mRNA levels at the endogenous *INO1* locus in strains lacking both GRS elements (Fig. 6c). Peripheral targeting correlated with transcription; expression of endogenous *INO1* was not significantly affected by the *grs I* mutation or *grs II* mutation alone (Fig. S2b). Therefore, the GRS elements are redundant for both *INO1* localization and transcription and full activation of *INO1* and *TSA2* requires DNA zip codes that confer peripheral localization.

Transcription of *INO1* is enhanced by promoter targeting to the nuclear periphery

We next asked if the part of the gene that is targeted to the nuclear periphery is important for transcription. We reintroduced the GRS I element either upstream or downstream of the coding sequence of the *grs I* mutant *URA3:INO1* and quantified the *INO1* mRNA levels upon activation. Reintroduction of GRS I ~450 bp upstream (*i.e.* 5') of *grs I* mutant *INO1* restored regulated targeting to the nuclear periphery (Fig. 6c) and largely suppressed the defect in transcription (Fig. 6d). In contrast, reintroduction of GRS I at the 3' end of *grs I* mutant *INO1* caused the gene to localize to the nuclear periphery constitutively, but it did not suppress the defect in transcription (Fig. 6c & 6d). This suggests that the targeting of the promoter, not the gene *per se*, to the nuclear periphery is associated with full transcriptional activation.

We also tested if GRS I was sufficient to promote transcriptional activation. We introduced GRS I upstream of a crippled *CYC1* promoter (*CYC1**) driving a β-galactosidase (*LacZ*) reporter gene. The well-established unfolded protein response element (UPRE) functions as an enhancer in this context and was sufficient to promote LacZ expression. However, the GRS I element did not enhance transcription of *CYC1*- LacZ* (Fig. S5a). Therefore, GRS I may not simply be an enhancer and its important role might be in controlling promoter interactions with factors at the nuclear envelope or the NPC that promote transcription of certain genes.

Discussion

The GRS elements are small, well-defined DNA elements with the ability to target ectopic chromosomal loci to a particular subnuclear location. The existence of such DNA zip codes suggests that genomes code for their own spatial organization. The DNA zip codes that we have identified are negatively regulated when *INO1* is repressed. Previous work has implicated transcriptional regulators in promoting peripheral targeting of genes5, 11 and it may be that this is why the regulation of GRS-mediated targeting requires that they are located within the promoter. Regulation of peripheral targeting was lost either when expression was constitutive (Fig. 1; Δ4 mutant) or when the element was introduced downstream of *INO1* (Fig. 4c). We also observed unregulated peripheral targeting when

GRS I was introduced downstream of *GAL1*, another regulated gene that is recruited to the nuclear periphery (Fig. S6). Therefore, transcription factors may regulate both transcription and gene targeting to the nuclear periphery. This suggests that the spatial organization coded by DNA can be dynamic and regulated.

Our previous work has shown that peripheral localization of genes can establish "transcriptional memory", which promotes the reactivation of genes like *INO1* and *GAL1* after they are repressed12. We show here that targeting to the nuclear periphery is also important for full expression of *INO1* and *TSA2*. It is still unclear how localization promotes activation. We find that peripheral targeting of the promoter, but not the 3'end of the gene, promotes *INO1* transcription. Likewise, introduction of the GRS I element downstream of the *GAL1* gene had no effect on its activation (Fig. S6). The GRS elements might promote transcription by recruiting transcription factors that both activate transcription and target genes to the nuclear periphery. However, the GRS I element was not sufficient to promote transcription from a crippled promoter. Therefore, it also remains possible that the important activity of the GRS elements is as DNA zip codes and that the expression of certain genes is promoted by protein-DNA interactions at the nuclear periphery or the NPC.

The GRS I element functions as a DNA zip code in an organism that is approximately one billion years diverged from the organism in which it was identified. Perfect matches of GRS I occur 112 times in the *S. pombe* genome, 19 of which are clearly in promoters (Table S2). It will be interesting to determine if these elements control the subnuclear localization of these genes. Our work suggests that GRS-mediated targeting to the nuclear periphery is an ancient mechanism that may be shared by many eukaryotes. We conclude that DNA zip codes represent an additional level of genetic information that controls the spatial organization of the genome and affects gene expression.

Methods

Chemicals and Reagents

Unless stated otherwise, chemicals were from Sigma Aldrich, DNA oligonucleotides were from Operon and Integrated DNA Technologies, restriction enzymes were from New England Biolabs and yeast media components were from Q-Biogene. Antibodies against GFP, fluorescent secondary antibodies and Human Pan Mouse IgG dynabeads were from Invitrogen and the antibody against myc was from Santa Cruz Biotechnology.

Plasmid construction

Plasmids pRS306-INO15, pRS304-Sec63-myc12 and pAFS14413 have been described. pRS306-INO1' contains the same *INO1* insert as pRS306-INO1 but in the opposite orientation with the 5' end of the gene near to the unique *Xho*I site in pRS30632. Plasmids pRS306-*grs I*mutINO1 and pRS306 *grs I*mutINO1' are derivatived from these two plasmids and were generated by Quick Change site directed mutagenesis to convert the GRS I sequence in the *INO1* promoter from 5'-GGGTTGGA-3' to 5'-AAACCAAA-3'. To reintroduce GRS I at the 5' or 3' end of *INO1*, GRS I (41–60) was cloned into pRS306-*grs I*mutINO1 and pRS306-*grs I*mutINO1' using *Xma*I and *Not*I. Plasmids were digested with

either *Stu*I to integrate at *URA3* or *Bgl*II to integrate at *INO1*. The Δ2 (−351 to −450), Δ3 (−251 to −350), Δ4 (−151 to −250) and Δ5 (−50 to −150) mutants were generated by deletion of non-overlapping regions of the *INO1* promoter in pRS306-INO1'. For DNA localization experiments, a fragment of 128 Lac operator repeats was moved from p6LacO1285 into each of these plasmids using *Hin*dIII and *Xho*I.

GRS I and GRS II mapping plasmids were created in p6LacO1285. Fragments larger than 150bp were generated by PCR amplification from genomic DNA then cloned into p6LacO128. Smaller fragments were cloned as 5' phosphorylated oligonucleotides with one of the following combinations of overhangs 1) *Sac*I and *Spe*I 2) *Xho*I at both ends 3) *Xma*I and *Not*I or 4) *BamH*I and *Not*I. Sequences for all oligonucleotides are described in supplementary Table S3.

pRS305Nup2-TAP was generated by cloning of Nup2-TAP into pCR2.1 (Invitrogen) using PCR amplification from BY4741 Nup2-TAP33 genomic DNA. Nup2-Tap was ligated into pRS305 using *Not*I and *Spe*I and the resulting plasmid was digested with *Swa*I and integrated at the *LEU2* locus in yeast strains WLY53 and WLY54.

The UPRE was removed from pJC00234 by digestion with *Xho*I and re-ligation of the cut plasmid to generate pJC002CYC1*LacZ. GRS I (41–60) was cloned into *Xho*I cut pJC002 using 5' phosphorylated oligonucleotides to generate pJC002GRSCYC1*LacZ.

Yeast strains

Yeast strains used in this study are described in supplementary Table S4. Except for BY4741 Nup2-TAP 33 all strains were constructed from CRY1 (*ade2-1 can1-100 his3-11,15 leu2-3,112 trp1-1 ura3-1* Mat **a**). Deletions in yeast proteins were made using the PCR based deletion system 35. Deletions were confirmed by PCR using genomic DNA.

Yeast strains for GRS mapping were generated by integration of *Stu*I-digested plasmids containing the LacO array and relevant GRS fragments at *URA3*. For integration of smaller GRS fragments ≤ 10bp the PCR based integration system35 was adapted to integrate the relevant GRS fragment along with the kanMX6 marker at *URA3* as follows. Plasmid p6LacO128 was first integrated at *URA3*. GRS fragments were included in the primers used to amplify the kanMX6 gene. The 5' 45bp of the primers had homology to the β-lactamase gene in pRS306. As a control, we integrated the kanMX6 marker alone at *URA3*. This integration had no effect on the localization of *URA3*. This PCR-based integration strategy was also used to introduce GRS I at the 3' end of *GAL1* using a strain containing p6LacO*GAL1* integrated downstream of *GAL1*12.

Chromosomal mutations in GRS I and GRS II at endogenous *INO1* locus were made using homologous recombination. Fragments including the entire *INO1* gene, promoter and 3'UTR that were mutant for *grs I* (AAACCAAA), *grs II* (deletion of the central TGCTGG sequence) or both, were transformed into a *pro*Δ*ino1* mutant strain. This strain lacks the 450 bp upstream of the *INO1* transcriptional start site and a part of the coding sequence, which is replaced by the kanMX6 marker35. We selected for strains that had recovered *INO1* by selecting for inositol prototrophy. Ino+ transformants that had lost the kanMX6 marker were

confirmed by DNA sequencing. As a control for this approach, wild type *INO1* was also recreated in this way and was used as the wild type control.

The *TSA2 grs I* mutant was generated in the chromosome using the *delitto perfetto* strategy36. To introduce GRS I into *Schizosaccharomyces pombe*, we took advantage of the ability of *URA3* from *Saccharomyces cerevisiae* to complement the *ura4* mutation in the orthologous gene from *Schizosaccharomyces pombe*37. The *Ura4+* gene was replaced with a non-functional fragment of the *URA3* gene from *Saccharomyces cerevisiae* called *ura3.1*. We amplified the *ura3.1* mutant by two rounds of PCR. First, using *S. pombe* genomic DNA as template, we carried out two reactions using either the ura4up + ura3.1up primer pair or the ura3.1down + ura4down primer pair. These reactions generated two products of 215 base pairs, each having 25 base pairs of homology to *URA3* at one end. These products were then used as primers to amplify 835 basepairs of *URA3* corresponding the coding sequence, but lacking the promoter using pRS306 as template32. The PCR product was transformed into strain 972 h- and 5 fluoroorotic acid resistant transformants were isolated to create strain MM16037. Strain MM160 was transformed with *Stu*I-digested p6LacO128 or p6LacO (41–75) to generate strains MM162 and MM163, respectively. MM169 (*NUP120myc, ura4-Δ18, ade6+, his7+::LacI-GFP*) was then crossed to MM162 and MM163 to obtain strains MM170 and MM171, respectively. PCR analyses confirmed expected insertions within all strains.

Reverse transcriptase real-time quantitative PCR

RNA preparation and RT Q-PCR analysis was performed as described12.

Chromatin immunoprecipitation

Chromatin immunoprecipitation was performed as described12, except that TAP-tagged Nup2 was recovered using Pan Mouse IgG dynabeads (Invitrogen). DNA was quantified by real-time quantitative PCR5.

Chromatin localization assay

Samples were visualized on a Zeiss LSM510 confocal microscope in the Northwestern University Biological Imaging Facility. Chromatin localization experiments in *Saccharomyces cerevisiae* were performed as described5, 12. Briefly, methanol fixed, spheroplasted, detergent-extracted cells were probed with 1:200 monoclonal anti-myc (to detect Sec63-myc) and 1:1000 rabbit polyclonal anti-GFP (to detect GFP-Lac repressor). Secondary antibodies were diluted 1:200. A single z slice through each cell having the brightest and most focused anti-GFP spot was collected. Cells in which this anti-GFP spot colocalized with Sec63-*myc* nuclear membrane staining were scored as peripheral and all other cells were scored as nucleoplasmic. For each biological replicate, the fraction of cells in a population of 30–50 cells that scored as peripheral was determined. Error bars represent the standard error of the mean between biological replicates.

For experiments in *Schizosaccharomyces pombe*, 10ml of cells were grown to mid-log phase in YEA, transferred to YEA + 2.4M Sorbitol for 30 minutes followed by fixation with 3.5% formaldehyde for 1 hour. Cells were washed twice with 1ml PEMS (100mM PIPES, 1mM

EGTA, 1mM $MgSO_4$, 1.2M Sorbitol pH6.9) and resuspended in 1ml PEMS containing 0.2% BME and 1mg per ml Zymolyase 100T. Spheroplasting was checked under the microscope by mixing 0.5µl of 20% SDS and 9.5µl of cells. Spheroplasts were spun down, washed three times with 1ml PEMS, resuspended in PEMS + 1% Triton, resuspended in 1ml PEMBAL (100mM PIPES, 1mM EGTA, 1mM $MgSO_4$, 1% BSA, 0.1% sodium azide, 100mM L-lysine hydrochloride) and rotated for 30 minutes at room temperature. Spheroplasts were spun down, and resuspended in 100µl of PEMBAL containing 1:100 myc antibody and 1:500 GFP antibody 3 hours or overnight. Spheroplasts were washed three times with 1ml PEMBAL, resuspended in 100µl of PEMBAL containing a 1:100 dilution of Alexa Fluor 594 goat anti-mouse IgG and Alexa Fluor 488 goat anti-rabbit IgG and incubated for 3 hours. Spheroplasts were washed three times with 1ml PEMBAL and once with PBS + 0.1% sodium azide. 10µl of cells were spotted onto polylysine-treated slides and sealed with 2µl of mounting media.

Statistical methods

Table I, Table II and Table S1 used the Chi square test. For Table I and supplementary Table S1: each expression dataset was ranked according to signal (mRNA levels relative to untreated control). The number of GRS I-containing genes that were above and below the top 10% of the range was compared with the predicted number (10% and 90% of the total within each set, respectively) by Chi square test. For Table II: the number of non-GRS I or GRS I genes was calculated as described in the footnote of Table II and compared with the observed number of non-GRS I and GRS I genes in each ChIP by Chi square test.

β-galactosidase Miller assay

10 ml of cells were grown to log phase and an equal number of cells was pelleted and resuspended in 2 ml Z buffer (40mM NaH_2PO_4, 60mM Na_2HPO_4). 0.5 ml of cells were used for each reaction. 20µl of 0.1% SDS and 2 drops of chloroform were added and the reaction incubated at 30° C for 15 minutes. 160µl of 4mg/ml ortho-Nitrophenyl-β-galactoside was added and the reaction incubated at 30° C until a pale yellow color developed. The development time was noted and the reaction quenched by addition of 400µl of 1M Na_2CO_3. Absorbance values at 420 nm and 550nm were read. Miller units were calculated using the following formula38:

$$\text{Miller units} = \frac{1000 \times [(A_{420}) - (1.75 \times A_{550})}{\text{time} \times \text{volume of cells (0.5ml)} \times A_{600}}$$

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

We acknowledge Robert Lamb for generously sharing his confocal microscope. Also, we thank Audrey Gasch, Richard Carthew, Rick Gaber, Sandy Westerheide, Jonathan Widom, Susan Wente, Michael Rout and members of the Brickner lab for helpful discussions. This work was supported by the Searle Leadership Fund at Northwestern

University, a gift of The Searle Funds at The Chicago Community Trust (J.H.B.), an Institutional Research Grant from the American Cancer Society (J.H.B.) and National Institutes of Health grant GM080484 (J.H.B.).

References

1. Akhtar A, Gasser SM. The nuclear envelope and transcriptional control. Nat Rev Genet. 2007; 8:507–517. [PubMed: 17549064]
2. Ahmed S, Brickner JH. Regulation and epigenetic control of transcription at the nuclear periphery. Trends Genet. 2007; 23:396–402. [PubMed: 17566592]
3. Taddei A. Active genes at the nuclear pore complex. Curr Opin Cell Biol. 2007; 19:305–310. [PubMed: 17467257]
4. Casolari JM, et al. Genome-wide localization of the nuclear transport machinery couples transcriptional status and nuclear organization. Cell. 2004; 117:427–439. [PubMed: 15137937]
5. Brickner JH, Walter P. Gene recruitment of the activated INO1 locus to the nuclear membrane. PLoS Biol. 2004; 2:e342. [PubMed: 15455074]
6. Casolari JM, Brown CR, Drubin DA, Rando OJ, Silver PA. Developmentally induced changes in transcriptional program alter spatial organization across chromosomes. Genes Dev. 2005; 19:1188–1198. [PubMed: 15905407]
7. Dieppois G, Iglesias N, Stutz F. Cotranscriptional recruitment to the mRNA export receptor Mex67p contributes to nuclear pore anchoring of activated genes. Mol Cell Biol. 2006; 26:7858–7870. [PubMed: 16954382]
8. Taddei A, et al. Nuclear pore association confers optimal expression levels for an inducible yeast gene. Nature. 2006; 441:774–778. [PubMed: 16760983]
9. Sarma NJ, et al. Glucose-responsive regulators of gene expression in Saccharomyces cerevisiae function at the nuclear periphery via a reverse recruitment mechanism. Genetics. 2007; 175:1127–1135. [PubMed: 17237508]
10. Abruzzi KC, Belostotsky DA, Chekanova JA, Dower K, Rosbash M. 3'-end formation signals modulate the association of genes with the nuclear periphery as well as mRNP dot formation. Embo J. 2006; 25:4253–4262. [PubMed: 16946703]
11. Schmid M, et al. Nup-PI: the nucleopore-promoter interaction of genes in yeast. Mol Cell. 2006; 21:379–391. [PubMed: 16455493]
12. Brickner DG, et al. H2A.Z-mediated localization of genes at the nuclear periphery confers epigenetic memory of previous transcriptional state. PLoS Biol. 2007; 5:e81. [PubMed: 17373856]
13. Straight AF, Belmont AS, Robinett CC, Murray AW. GFP tagging of budding yeast chromosomes reveals that protein-protein interactions can mediate sister chromatid cohesion. Curr Biol. 1996; 6:1599–1608. [PubMed: 8994824]
14. Robinett CC, et al. In vivo localization of DNA sequences and visualization of large-scale chromatin organization using lac operator/repressor recognition. J Cell Biol. 1996; 135:1685–1700. [PubMed: 8991083]
15. Swift S, McGraw P. INO1-100: an allele of the Saccharomyces cerevisiae INO1 gene that is transcribed without the action of the positive factors encoded by the INO2, INO4, SWI1, SWI2 and SWI3 genes. Nucleic Acids Res. 1995; 23:1426–1433. [PubMed: 7753636]
16. Harbison CT, et al. Transcriptional regulatory code of a eukaryotic genome. Nature. 2004; 431:99–104. [PubMed: 15343339]
17. Lee TI, et al. Transcriptional regulatory networks in Saccharomyces cerevisiae. Science. 2002; 298:799–804. [PubMed: 12399584]
18. Colman-Lerner A, Chin TE, Brent R. Yeast Cbk1 and Mob2 activate daughter-specific genetic programs to induce asymmetric cell fates. Cell. 2001; 107:739–750. [PubMed: 11747810]
19. Carmona-Saez P, Chagoyen M, Tirado F, Carazo JM, Pascual-Montano A. GENECODIS: a web-based tool for finding significant concurrent annotations in gene lists. Genome Biol. 2007; 8:R3. [PubMed: 17204154]
20. Nogales-Cadenas R, et al. GeneCodis: interpreting gene lists through enrichment analysis and integration of diverse biological information. Nucleic Acids Res. 2009; 37:W317–W322. [PubMed: 19465387]

21. Cox JS, Chapman RE, Walter P. The unfolded protein response coordinates the production of endoplasmic reticulum protein and endoplasmic reticulum membrane. Mol Biol Cell. 1997; 8:1805–1814. [PubMed: 9307975]
22. Gasch AP, et al. Genomic expression programs in the response of yeast cells to environmental changes. Mol Biol Cell. 2000; 11:4241–4257. [PubMed: 11102521]
23. Travers KJ, et al. Functional and genomic analyses reveal an essential coordination between the unfolded protein response and ER-associated degradation. Cell. 2000; 101:249–258. [PubMed: 10847680]
24. Leber JH, Bernales S, Walter P. IRE1-independent gain control of the unfolded protein response. PLoS Biol. 2004; 2:E235. [PubMed: 15314654]
25. Wong CM, Ching YP, Zhou Y, Kung HF, Jin DY. Transcriptional regulation of yeast peroxiredoxin gene TSA2 through Hap1p, Rox1p, and Hap2/3/5p. Free Radic Biol Med. 2003; 34:585–597. [PubMed: 12614847]
26. Heckman DS, et al. Molecular evidence for the early colonization of land by fungi and plants. Science. 2001; 293:1129–1133. [PubMed: 11498589]
27. Luthra R, et al. Actively transcribed GAL genes can be physically linked to the nuclear pore by the SAGA chromatin modifying complex. J Biol Chem. 2007; 282:3042–3049. [PubMed: 17158105]
28. Cabal GG, et al. SAGA interacting factors confine sub-diffusion of transcribed genes to the nuclear envelope. Nature. 2006; 441:770–773. [PubMed: 16760982]
29. Menon BB, et al. Reverse recruitment: the Nup84 nuclear pore subcomplex mediates Rap1/Gcr1/Gcr2 transcriptional activation. Proc Natl Acad Sci U S A. 2005; 102:5749–5754. [PubMed: 15817685]
30. Chekanova JA, Abruzzi KC, Rosbash M, Belostotsky DA. Sus1, Sac3, and Thp1 mediate post-transcriptional tethering of active genes to the nuclear rim as well as to non-nascent mRNP. RNA. 2008; 14:66–77. [PubMed: 18003937]
31. Alber F, et al. The molecular architecture of the nuclear pore complex. Nature. 2007; 450:695–701. [PubMed: 18046406]
32. Sikorski RS, Hieter P. A system of shuttle vectors and yeast host strains designed for efficient manipulation of DNA in Saccharomyces cerevisiae. Genetics. 1989; 122:19–27. [PubMed: 2659436]
33. Ghaemmaghami S, et al. Global analysis of protein expression in yeast. Nature. 2003; 425:737–741. [PubMed: 14562106]
34. Cox JS, Walter P. A novel mechanism for regulating activity of a transcription factor that controls the unfolded protein response. Cell. 1996; 87:391–404. [PubMed: 8898193]
35. Longtine MS, et al. Additional modules for versatile and economical PCR-based gene deletion and modification in Saccharomyces cerevisiae. Yeast. 1998; 14:953–961. [PubMed: 9717241]
36. Storici F, Durham CL, Gordenin DA, Resnick MA. Chromosomal site-specific double-strand breaks are efficiently targeted for repair by oligonucleotides in yeast. Proc Natl Acad Sci U S A. 2003; 100:14994–14999. [PubMed: 14630945]
37. Grimm C, Kohli J, Murray J, Maundrell K. Genetic engineering of Schizosaccharomyces pombe: a system for gene disruption and replacement using the ura4 gene as a selectable marker. Mol Gen Genet. 1988; 215:81–86. [PubMed: 3241624]
38. MIller, JH. Experiments in Molecular Genetics. Cold Spring Harbor, NY, USA: Cold Spring Harbor Laboratory; 1972.

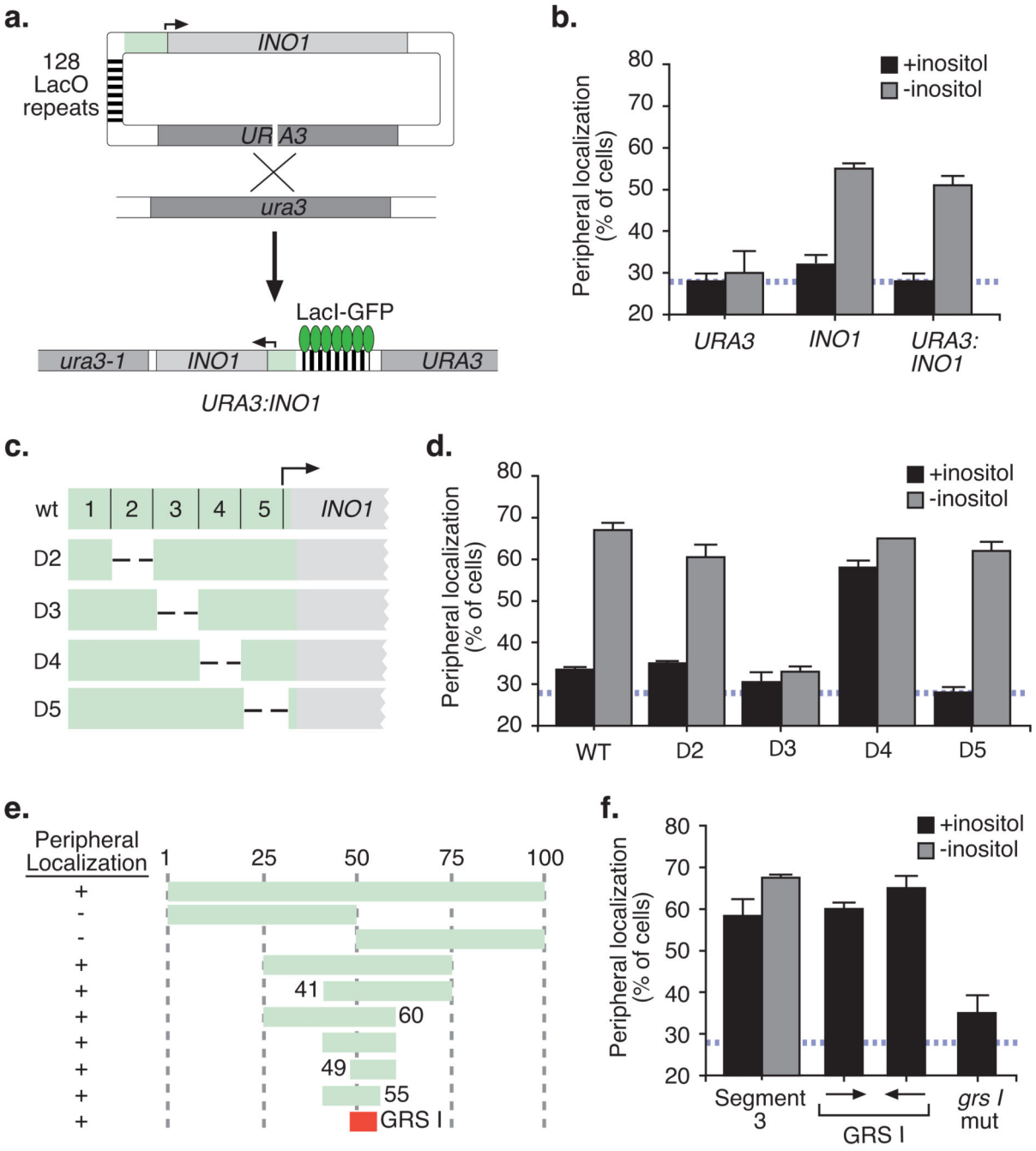

Figure 1. Identification of a Gene Recruitment Sequence (GRS) in the *INO1* promoter

(**a**) Integration scheme for integrating *INO1* and the lac repressor array at *URA3* by homologous recombination. The *INO1* gene included 504bp upstream and 685bp downstream of the coding sequence. (**b**) The fraction of the population in which the GFP spot colocalized with the nuclear envelope marker Sec63-myc for cells grown in the presence or absence of 100μM *myo*-inositol. The hatched blue line represents the mean peripheral localization for *URA3*. The maximal peripheral localization observed using this assay is ~80% of cells for a gene that is artificially tethered to the nuclear envelope5.

Therefore, the dynamic range of this assay is 20% – 80%. Data represent the mean and s.e.m from 5 biological replicates (30–50 cells were analyzed per replicate). Map (**c**) and peripheral localization (**d**) of 100bp non-overlapping deletions in the *INO1* promoter integrated at *URA3*. (**e**) Map of fragments within segment 3 that were integrated at *URA3* and their peripheral localization (complete data in Fig. S1a). (**f**) Peripheral localization of segment 3 in the presence and absence of inositol, the 8bp GRS I in either orientation or a mutant version of GRS I integrated at *URA3* ($n \geq 3$, 30–50 cells per biological replicate).

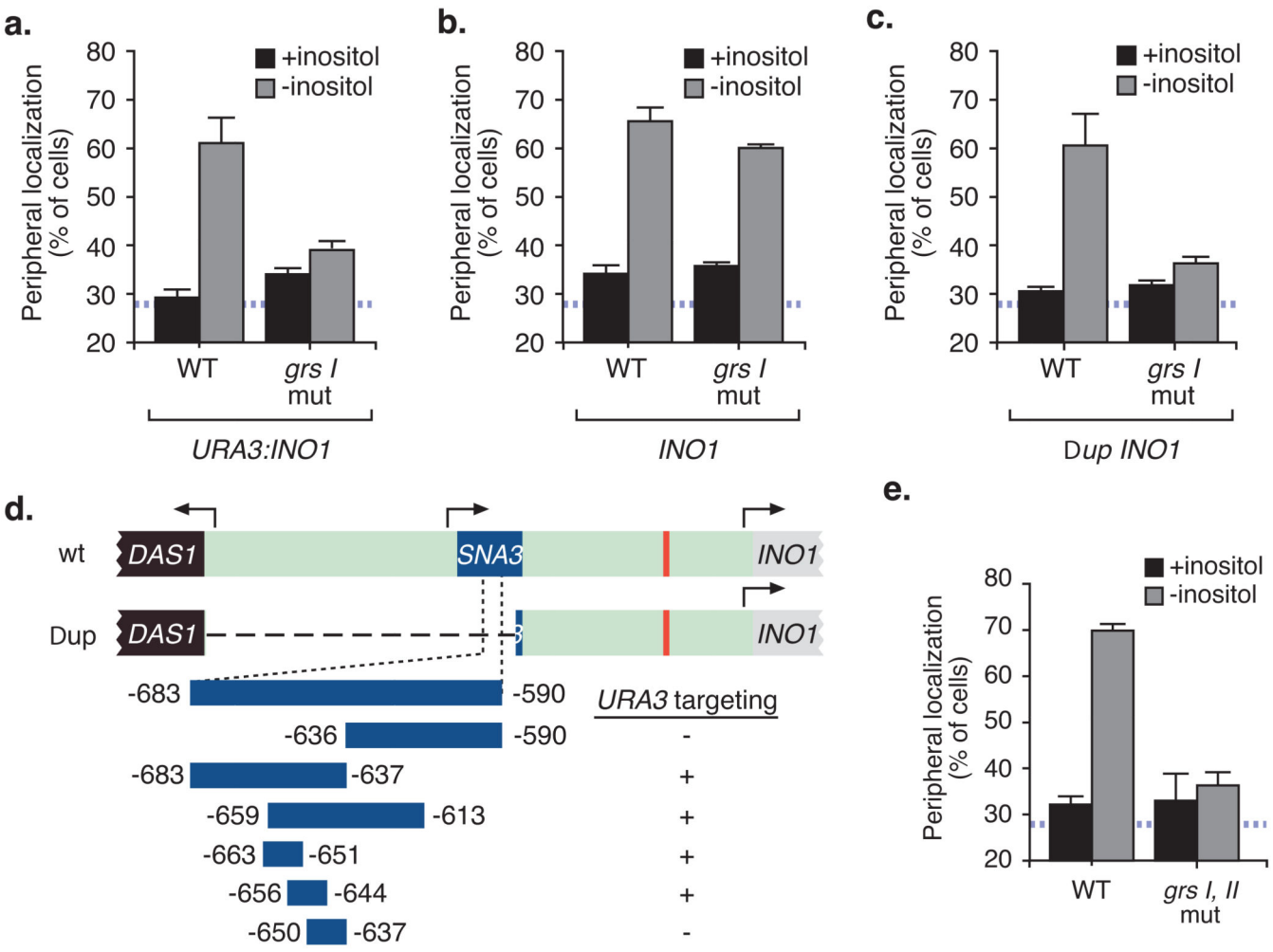

Figure 2. Targeting of the endogenous *INO1* gene is mediated by two redundant DNA zip codes

(**a**) Localization of wild type *INO1* and GRS I mutant *INO1* integrated either at *URA3* (**a**) or at *INO1* (**b**). (**c**) Localization of wild type or *grs I* mutant *INO1* at endogenous *INO1* in a strain lacking the 943bp upstream of GRS I as depicted in panel (**d**) (for panels **a** – **d**, $n = 3$, 30–50 cells per replicate). (**d**) Map of fragments used to identify GRS II by integration at *URA3* and their peripheral localization (complete data in Figs. S1b and S1c). (**e**) Peripheral localization of a combined mutation in GRS I and GRS II at endogenous *INO1* ($n \geq 3$, 30–50 cells per replicate).

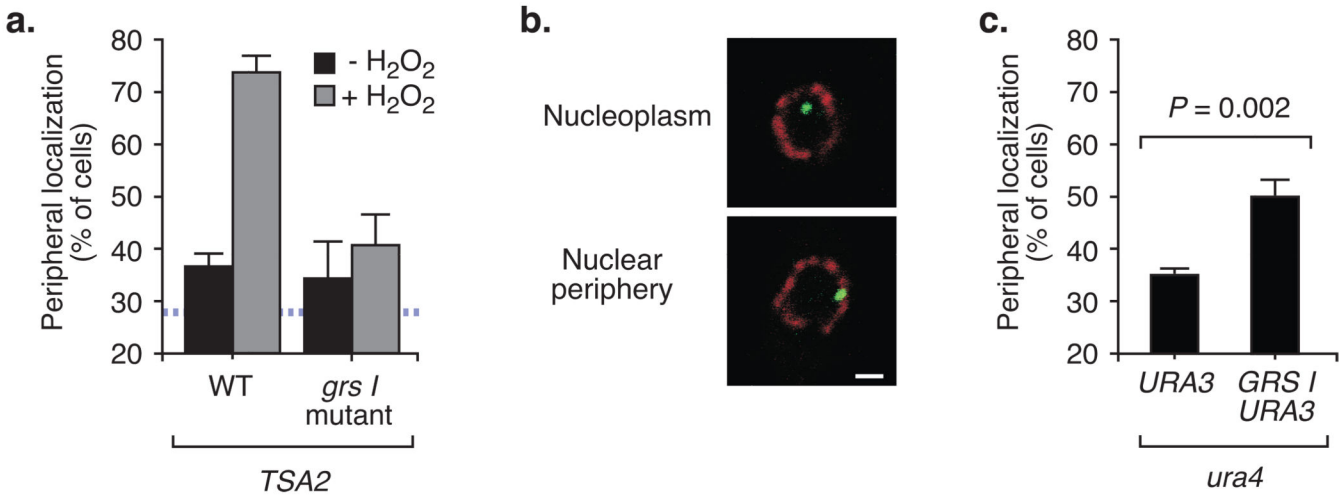

Figure 3. GRS I mediated targeting to the nuclear periphery is general and ancient

(**a**) Peripheral localization of the wild type or *grs I* mutant *TSA2* gene was determined for a population of cells for cells grown in YPD (-H_2O_2 repressing conditions) or in YPD + 0.5mM H_2O_2 (activating conditions) (*n* = 3, 30–50 cells per replicate). (**b**) Introduction of the GRS into *S. pombe*. Representative confocal micrographs of immunofluorescence against LacI-GFP (green) and Nup120-myc (red) in *S. pombe* that were scored as either nucleoplasmic or peripheral. The scale bar equals 1μm. (**c**) The fraction of the population in which *ura4* colocalized with Nup120-myc in cells with either the lac repressor array plasmid or the lac repressor array plasmid with a single copy of GRS I (*n* = 4, 60–100 cells per replicate).

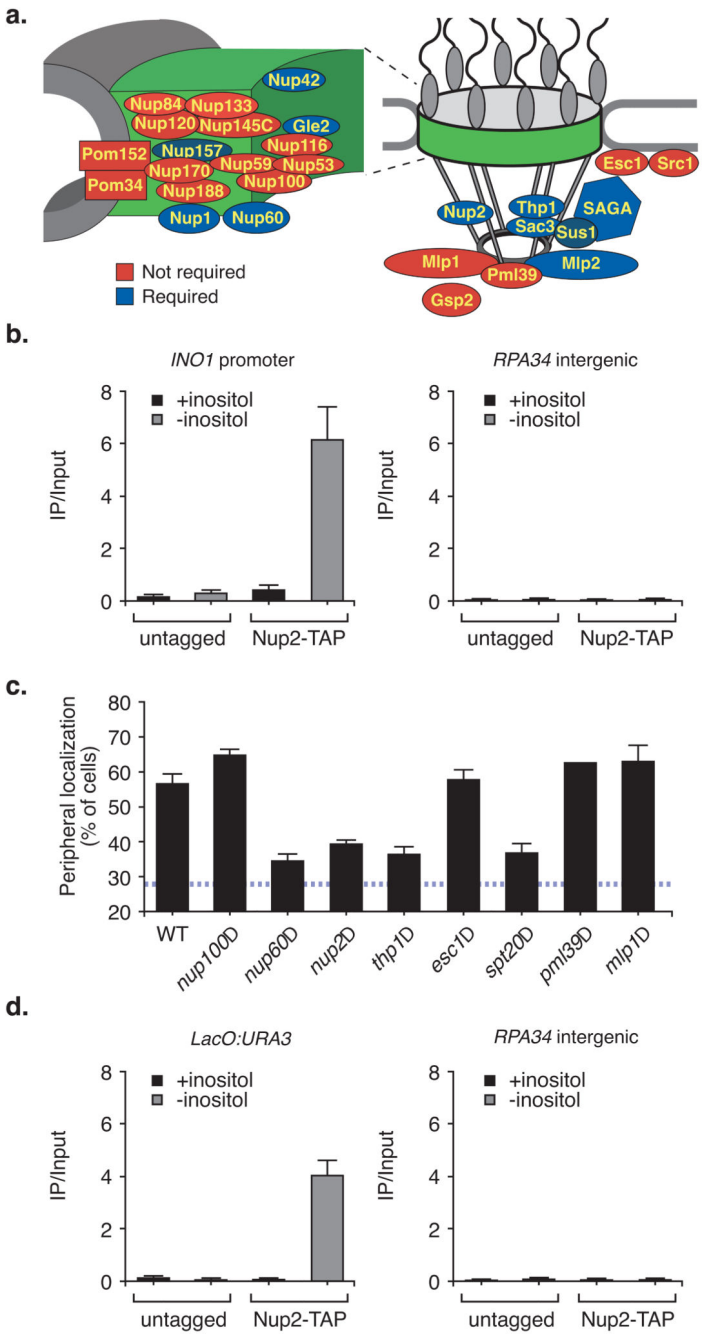

Figure 4. *INO1* recruitment to the nuclear periphery requires components of the nuclear pore complex (NPC) and associated factors

(**a**) Summary of nuclear pore proteins and associated factors required for *INO1* targeting to the nuclear periphery (complete data in Fig. S3; data for Nup2 has previously been published12). Proteins filled in blue were required for *INO1* peripheral targeting. Proteins filled in red were not required for *INO1* peripheral targeting. The positions of proteins within the pore and contacts between them are approximations based on a structural model of the NPC structure31. Left: an expanded view of the core channel of the NPC. (**b**)

Chromatin immunoprecipitation of Nup2-TAP under repressing and activating conditions. Recovery of the *INO1* promoter (left panel) or the *RPA34* intergenic region ~5000 bp upstream (right panel) with IgG magnetic beads (Invitrogen) was quantified relative to the input by real-time quantitative PCR. (**c**) Nuclear pore requirements for GRS I mediated targeting of *URA3* ($n \geq 2$, 30–50 cells per replicate). Cells were grown in the presence of inositol. (**d**) Recovery of *URA3* or *RPA34* with Nup2 (as in panel **b**) from strains with or without GRS I at the *URA3* locus (see Methods). For panels **b** and **d**, data represent the mean and the s.e.m. ($n = 3$).

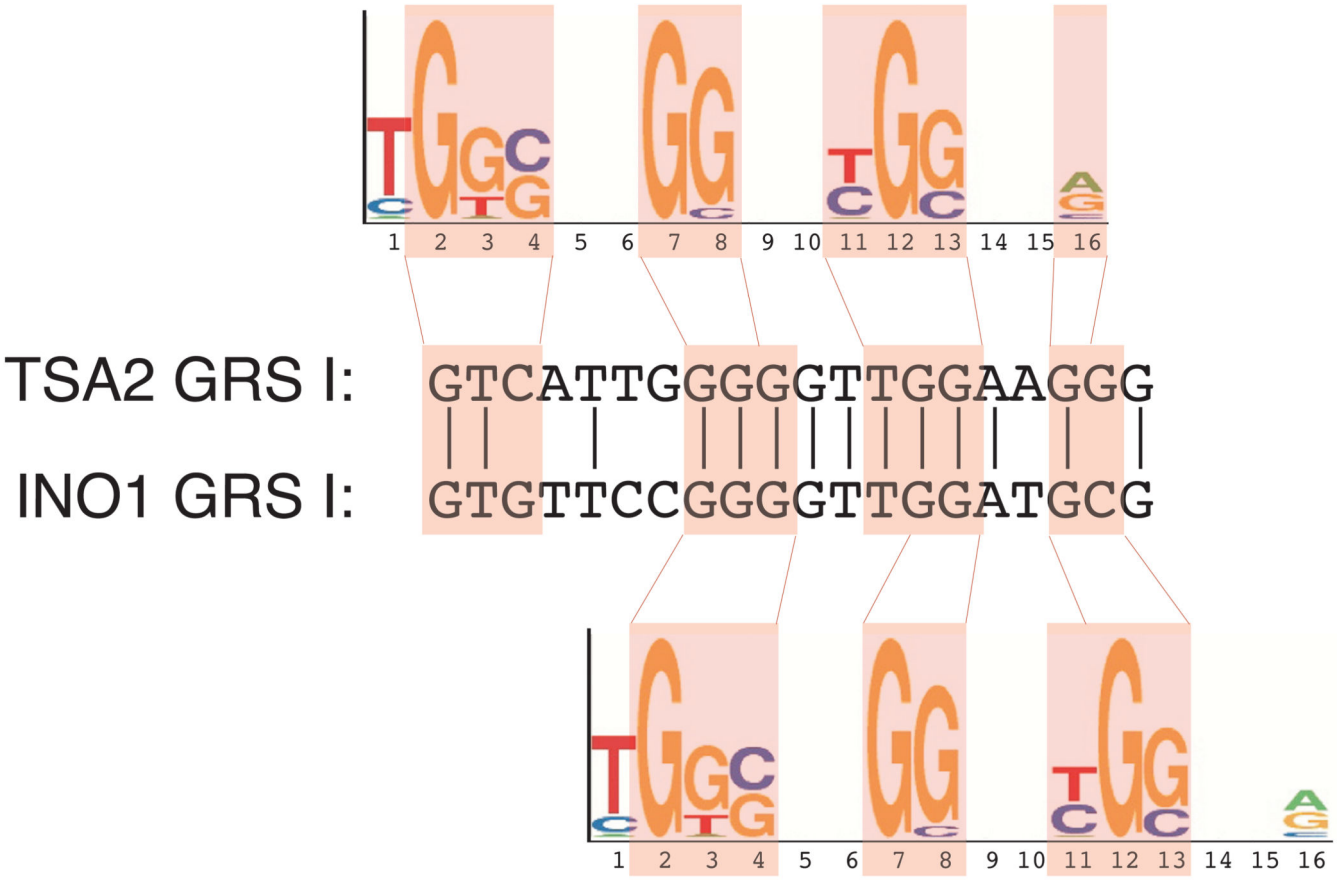

Figure 5. GRS I is enriched among genes that interact with many nuclear pore proteins

Comparison of GRS I from *INO1* and *TSA2* to a motif, identified by *Casolari et al*.4, that is overrepresented in genes associated with Mlp1, Mlp2 or Nic96. Two alternative alignments are shown.

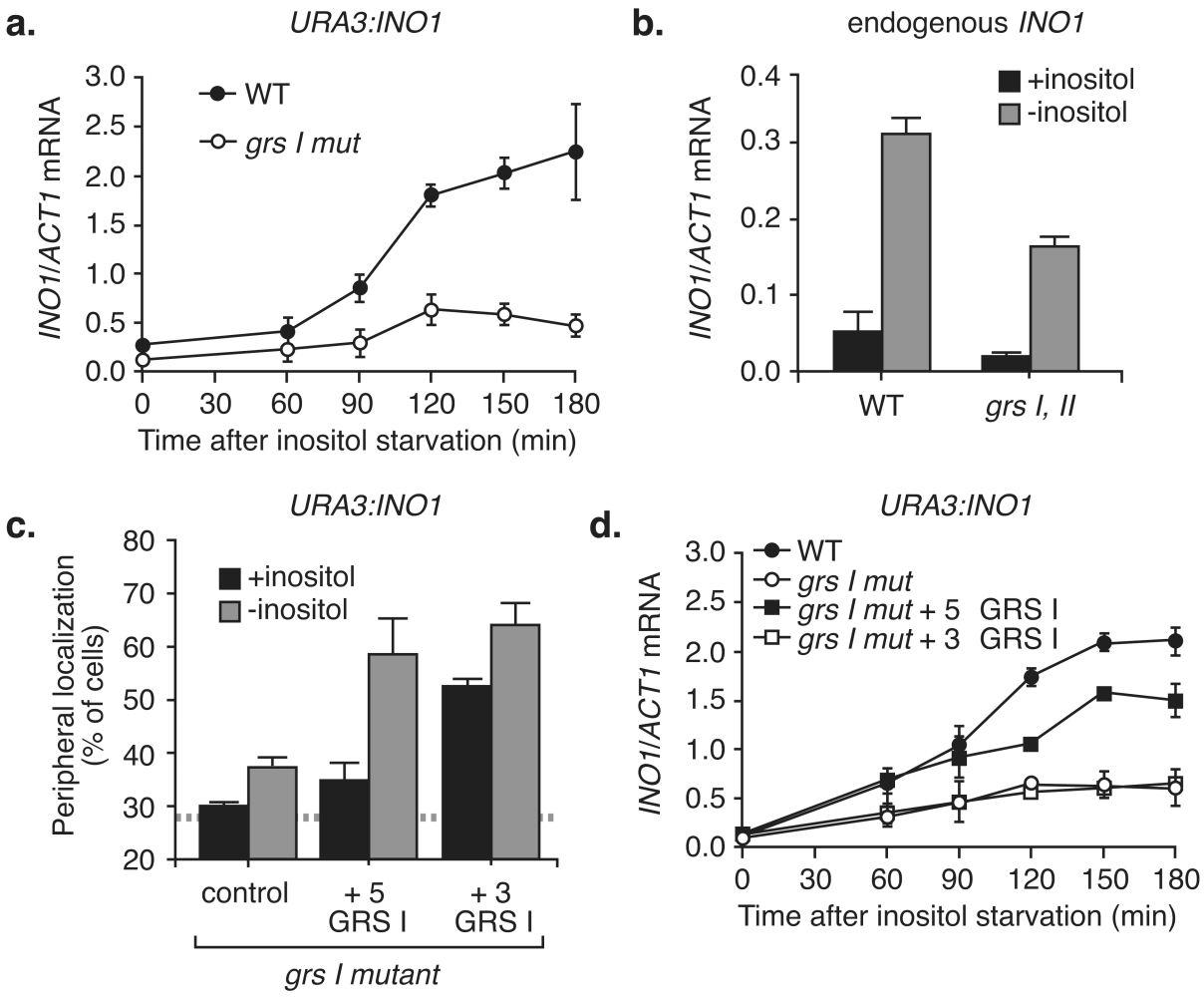

Figure 6. Localization at the nuclear periphery enhances transcription of *INO1*

(**a**) *INO1* mRNA levels were quantified by RT-qPCR, relative to *ACT1*, following induction by inositol starvation from strains having either plasmid-borne wild type *INO1* or *grs I* mutant *INO1* integrated at *URA3*. (**b**) Strains were constructed in which the *grs I, II* mutations were introduced at the chromosomal *INO1* locus. Wild type and *grs I, II* mutant strains were grown overnight in the presence or absence of inositol and the mRNA levels quantified as in (**a**). (**c** & **d**) GRS I was reintroduced either at the 5' end or the 3' end of *grs I* mutant *INO1* and these plasmids were integrated at *URA3*. These strains were compared with wild type and *grs I* mutant INO1 for localization (**c**) and transcription (**d**). For panels **a** – **d**, $n = 4$.

Table I

Coregulation of GRS I-element genes [ζ]

	< 90th percentile[¬]		> 90th percentile[)]			
Condition	**Exp.**	**Obs.**	**Exp.** [Ψ]	**Obs.**	χ^2	***P***
Tm [ι], 60min	54	49	6	11	4.630	0.0314*
DTT[©], 60min	54	49	6	11	4.630	0.0314*
DTT + HS[(]	58	49	7	16	12.968	0.0003***
*opi1*Δ[^]	67	63	7	11	2.525	0.1121
HS 5min[√]	52	47	6	11	4.647	0.0311*
HS 10min[√]	56	49	6	13	9.042	0.0026**
HS 15min[√]	59	50	7	16	12.944	0.0003***
HS 20min[√]	53	47	6	12	6.679	0.0098**
HS 30min[√]	57	51	6	12	6.632	0.0100*
N Depl., 8h[√]	64	61	7	10	1.426	0.2324
N Depl., 12h[√]	64	60	7	11	2.536	0.1113

[ζ]Genes having a GRS I element within 775bp of the translational start site.

[¬]Expected < 90th percentile = 0.9 × [number of GRS I genes in expression data]

[)]Expected > 90th percentile = 0.1 × [number of GRS I genes in expression data]

[ι]Tunicamycin, averaged from two replicates from Leber et al.24

[©]Dithiothreitol, from Leber et al.24

[(]Dithiothreitol + Heat shock, from Leber et al.24

[^]From Travers et al.23; This strain lacks the repressor of *INO1* transcription and overexpresses inositol-repressed genes.

[√]From Gasch et al.22

***Extremely significant enrichment

**Very significant enrichment

* Significant enrichment

Nat Cell Biol. Author manuscript; available in PMC 2010 August 01.

Table II

GRS I genes interact with the nuclear pore complex

	Non-GRS I [ζ]		GRS I [¬]			
ChIP	**Exp.**	**Obs.**	**Exp.**	**Obs.**	χ^2	***P***
Cse1	454	443	7	18	17.552	0.0001***
Nup116	528	518	8	18	12.689	0.0004***
Nup2	290	283	4	11	12.419	0.0004***
Mlp1	392	384	6	14	10.830	0.0010***
Xpo1	384	376	6	14	10.833	0.0010***
Nup60	169	164	3	8	8.481	0.0036**
Nup100	1383	1371	21	33	6.961	0.0083**
Mlp2	431	425	7	13	5.226	0.0222*
Nup84	54	52	1	3	4.074	0.0435*
Nup145	216	214	3	5	1.352	0.2450
Kap95	94	93	1	2	1.011	0.3147
Nsp1	114	113	2	3	0.509	0.4757
Nic96	583	582	9	10	0.113	0.7369

[ζ] Expected Non-GRS I genes = [number of bound targets] − [expected GRS I genes]

[¬] Expected GRS I genes= [number of GRS I genes] × ([number of bound targets] / [number of microarray probes])

*** Extremely significant enrichment

** Very significant enrichment

* Significant enrichment