The Basis for Evolution of DNA-Binding Specificity of the Aft1 Transcription Factor in Yeasts

Isabelle R. Gonçalves,^{*,1} Natalia Conde e Silva,^{†,1} Cesar La Torre Garay,[‡] Emmanuel Lesuisse,[‡]
Jean Michel Camadro,[‡] and Pierre Louis Blaiseau^{*,§,2}

^{*}Microbiologie, Adaptation et Pathogénie, Unité Mixte de Recherche 5240, Centre National de la Recherche Scientifique, Université Lyon 1, Bayer CropScience, F-69622 Villeurbanne, France, [†]Institut de Biologie des Plantes, Centre National de la Recherche Scientifique Unité Mixte de Recherche 8618, Université Paris-Sud, Saclay Plant Sciences, F-91405 Orsay Cedex, France, [‡]Mitochondries, Métaux et Stress Oxydatif, Institut Jacques Monod Unité Mixte de Recherche 7592, Centre National de la Recherche Scientifique, Université Paris Diderot, Sorbonne Paris Cité, F-75205 Paris Cedex 13, France, and [§]Université Pierre et Marie Curie, Paris 06, Unité de Formation et de Recherche 927 Sciences De la Vie, F-75005 Paris Cedex 05, France

ABSTRACT The *Saccharomyces cerevisiae* Aft1 and *Kluyveromyces lactis* KIAft are orthologous yeast transcription activators that regulate the expression of the same group of iron-uptake genes but bind to the different DNA sites: TGCACCC for Aft1 and PuCACCC for KIAft. To establish whether the DNA-binding mechanisms of Aft1 and KIAft have diverged during the evolution of the Aft-type transcription factor, we examined the function of a nonconserved region in their DNA-binding domains. A large part of this region is composed of a sequence predicted to be disordered in structure and potentially phosphorylated. We show with deletion mutant analyses that this sequence is essential for the binding of Aft1 to its DNA site and for the iron uptake and growth of *S. cerevisiae* under iron-limited conditions. We constructed hybrid proteins by exchanging the nonconserved regions of Aft1 and KIAft. We show that the Aft1 region is necessary and sufficient for KIAft to bind efficiently to the Aft1 DNA site in *S. cerevisiae* and to complement the iron-dependent phenotype of the *aft1Δaft2Δ* mutant. This demonstrates that the changes in the nonconserved region of the Aft-type DNA-binding domain have led to changes in the DNA-binding specificity and have major consequences for the regulation of iron homeostasis. The combination of bioinformatic and experimental analyses indicates that the sequence TGCACCC is the most probable ancestral Aft-type element. Our findings suggest that the changes in the nonconserved region of the DNA-binding domain are responsible for the evolution of the TGCACCC sequence toward PuCACCC in the *K. lactis* species.

IRON is an essential element but its accumulation can be toxic for the cell (Halliwell and Gutteridge 1984). Therefore, organisms have developed various regulated systems to adjust the iron uptake flux to the cell requirements (Blaiseau *et al.* 2010). In fungi, the expression of iron uptake genes is under the control of either of two opposite modes of transcriptional regulation. In most fungi, the iron-regulatory pathway is mediated by a Zn-finger GATA-type factor that represses the transcription of the iron-uptake genes under

high-iron conditions (Haas *et al.* 2008). In contrast, in the hemiascomycete yeasts *Saccharomyces cerevisiae* and *Kluyveromyces lactis*, the transcription of the same group of genes is activated by an Aft-type transcription factor under conditions of low iron availability (Yamaguchi-Iwai *et al.* 1995; Conde e Silva *et al.* 2009). An AFT gene has been identified in the genome of all hemiascomycetes studied except *Yarrowia lipolytica*. Some hemiascomycete genomes, including that of *S. cerevisiae*, have two copies of the AFT gene resulting from the whole genomic duplication (WGD) event (Wolfe and Shields 1997). We previously showed that the *S. cerevisiae* Aft1 and Aft2 paralogs have specialized iron-regulatory functions: Aft1 specifically activates the transcription of the genes involved in the iron uptake at the cell surface, whereas Aft2 activates the transcription of genes involved in vacuolar and mitochondrial transport (Courel *et al.* 2005; Dhaoui *et al.* 2011). Examination of the promoter sequences of Aft1- and Aft2-regulated

Copyright © 2014 by the Genetics Society of America

doi: 10.1534/genetics.113.157693

Manuscript received September 19, 2013; accepted for publication October 21, 2013; published Early Online October 23, 2013.

Supporting information is available online at <http://www.genetics.org/lookup/suppl/doi:10.1534/genetics.113.157693/-DC1>.

¹These authors contributed equally to this work.

²Corresponding author: Mitochondries, Métaux et Stress Oxydatif, UMR 7592 Institut Jacques Monod, CNRS, Université Paris Diderot, 15 rue Hélène Brion, 75205 Paris Cedex 13, France. E-mail: blaiseau.pierre-louis@ijm.univ-paris-diderot.fr

genes and DNA-binding analyses show that *Aft1* and *Aft2* recognize similar but distinct DNA sequences: TGCACCC and PuCACCC, respectively (Yamaguchi-Iwai *et al.* 1996; Rutherford *et al.* 2003; Courel *et al.* 2005). The investigation of iron regulation in the pre-WGD species *K. lactis* indicates that KLAft, the ortholog of *Aft1* and *Aft2*, activates the transcription of homologs of *Aft1*-regulated genes, but not of *Aft2*-regulated genes (Conde e Silva *et al.* 2009). This suggests that *Aft1* and KLAft have kept their ancestral iron-regulatory function, which is to control genes encoding the cell surface iron uptake components, whereas *Aft2* evolved to have a new Aft-type regulatory function emerging after the WGD involving the control of intracellular iron transport.

Although *Aft1* and KLAft regulate the same group of genes, there is evidence that their iron-regulated mechanisms of transcription are different. Heterologous complementation experiments indicate that KLAft does not totally suppress the *aft1Δaft2Δ* mutant phenotype (Conde e Silva *et al.* 2009). DNA-binding analyses with *Aft1* translated *in vitro* showed that it binds better to the TGCACCC sequence than to TACACCC and that a DNA duplex containing the GGCACCC sequence competes poorly with TGCACCC (Yamaguchi-Iwai *et al.* 1996). In contrast, KLAft binds to promoters containing various ACACCC sequences such as GACACCC, CACACCC, and TACACCC (Conde e Silva *et al.* 2009). Consistent with this, bioinformatic analyses identify PuCACCC and TGCACCC as the Aft-type consensus sequences in the promoters of KLAft-regulated and *Aft1*-regulated genes, respectively.

Several examples of the divergence of DNA-binding specificity between orthologous yeast transcription regulators with conserved function have been reported (Gasch *et al.* 2004; Kuo *et al.* 2010; Baker *et al.* 2011). However, only one study correlated the change of DNA-binding specificity with a particular change in the amino acid sequence of the DNA-binding domain (Kuo *et al.* 2010). Here, we investigated whether particular differences in the DNA-binding domains of *Aft1* and KLAft are associated with the differences in their DNA-binding sites. The DNA-binding domain of the Aft proteins has been identified as a WRKY-GCM1 zinc-finger domain (Babu *et al.* 2006). It belongs to the subgroup of the WRKY-GCM1 domains with insert because it contains an insert of 20–90 residues between two conserved cysteine residues that have been suggested to be involved in a zinc-finger domain (Babu *et al.* 2006; Conde e Silva *et al.* 2009). We examined the amino acid sequence of the insert and identified a variable sequence (VS) within the insert that is predicted to be disordered and highly phosphorylated. We show that *Aft1* VS (residues 152–203) is necessary for binding of *Aft1* to the DNA sequence TGCACCC. Deletion of *Aft1*(152–203) led to severe deficiency in both the transport of iron and growth in iron-depleted conditions. By ectopic expression of hybrid proteins, we show that the replacement of the KLAft insert by that of *Aft1* is necessary and sufficient for KLAft to bind the TGCACCC sequence and complement the iron-dependent *aft1Δaft2Δ* mutant phenotype. This

demonstrates that the divergence of the DNA-binding specificity between the orthologs *Aft1* and KLAft is strictly dependent on the nonconserved region inserted into their DNA-binding domains.

Materials and Methods

S. cerevisiae strains

The *S. cerevisiae* strains used in this study were BY4742 (*MATα*; *his3Δ1*; *leu2Δ0*; *lys2Δ0*; *ura3Δ0*) and SCMC01 (*MATα*; *his3Δ1*; *leu2Δ0*; *lys2Δ0*; *ura3Δ0*; *aft1::kanMX4*; *aft2::kanMX4*).

Plasmid constructions

The plasmid *pEG202-AFT1* containing the sequence encoding the DNA-binding domain of LexA has been described in Blaiseau *et al.* (2001). Plasmids *pCXJ22-AFT1* and *pCXJ22-KLAFT* have been described in Conde e Silva *et al.* (2009). The plasmids containing the *AFT1Δ* deletion mutants were constructed by using *pEG202-AFT1* as the PCR template with the QuikChange Mutagenesis Kit (Stratagene, La Jolla, CA) according to the manufacturer's instructions. For *pEG202-AFT1Δ*(152–203), *pEG202-AFT1Δ*(152–181), *pEG202-AFT1Δ*(182–203), and *pEG202-AFT1Δ*(172–203), the primers used were, respectively, 5'-GCTGCTAAAAGG GGAAGGAACAAGAAAAAAGATGTGTATCGAGG-3', 5'-GCTGCTAAAAGGGGAAGGAACCTTCTAATGCCACA GTAACCAATG-3', 5'-GATGACGAATTAGAATATGCGAGTA AGAAAAAAGATGTGTATCGAGG-3', 5'-CACGAAGACGA GAAATCCAAGAAGAAAAAAGATGTGTATCGAGG-3', and their complements. Plasmids *pEG202-KLAFT* and *pEG202-KLAFT^{AFT1}*(143–215) were generated by gap repair in yeast. For *pEG202-KLAFT*, the cells were cotransformed with *Bam*HI-linearized *pEG202* and a DNA fragment produced by PCR from *pCXJ22-KLAFT* as the template, with the following primers: 5'-GGCGGTTGGGGTTATTCGCAACGGCG ACTGGCTGGAATTCATGAAGCACGAGCTACAGACAC-3' and 5'-CCGGAATTAGCTTGGCTGCAGGTCGACTCGAGCGGC CGC TCATTGAAAAAGCCGTGAGGTTTC-3'. For *pEG202-KLAFT^{AFT1}*(143–215), cells were cotransformed with *Rsr*II-linearized *pEG202-KLAFT* and a DNA fragment produced by PCR from *pCXJ18-AFT1* (Conde e Silva *et al.* 2009) as the template, with the following primers: 5'-CATTGAAC GGTCGGATAGCATGAAGATCGTGTTCAAATGTAAAGCTGC TAAAAGGGGAAGGAAC-3' and 5'-TTCTCTTCAAAGAGAA CGTGGCTCGCACTCTGAAGGGACAGTTATTAAACCTCGA TACACATC-3'. To construct *pEG202-AFT1^{KLAFT}*(134–229), we used a two-step PCR protocol (Wach 1996) involving the following primers: 5'-GGCGGTTGGGGTTATTCGC AACGGCGACTGGCTGGAATTCATGGAAGGCTTCAATCCGG C-3', 5'-CCGGAATTAGCTTGGCTGCAGGTCGACTCGAGCG GCCGCTAATCTTCTGGCTTCACATACTTC-3', 5'-GTATTT CTATTCTTTTAAACAGCTTTACACTTGAAGACAACCTTTAA ATGCG-3', and 5'-ACGTGCCATCGGACCGTACAATTCAT GTCCGTTTAGAGTACGAGCTAC-3'.

The sequences underlined are homologous to sequences in pEG202, the sequences in italics are homologous to sequences in *KLAFT*, and the other sequences are homologous to sequences in *AFT1*. For the first and the second PCR steps, the DNA templates were pCXJ22-*AFT1* and pCXJ22-*KLAFT*, respectively. The final PCR product was first inserted into the vector pCR-2.1-TOPO (Invitrogen) and then transferred to between the *Bam*HI-*Xho*I sites of pEG202. To obtain pCXJ22-*KLAFT*^{*AFT1* (143–215)}, the *Age*I-*Xho*I fragment of pEG202-*KLAFT*^{*AFT1* (143–215)} was inserted between the *Age*I-*Xho*I sites of pCXJ22-*KLAFT*. Plasmid pEG202-*LTAFT* expressing LexA-*LTAFT* was constructed by integrating the open reading frame of *LTAFT* in frame with the DNA binding domain of LexA. The genomic DNA from the *Lachancea thermotolerans* strain CBS 6340 was used as template for PCR to amplify the open reading frame of *LTAFT* with the following oligonucleotides: 5'-GGCGTTGGGGTTATTGCGAACGGCGACTGGCTGGAATTCATGGATCAGTCAGCATTAGACG-3' and 5'-CCGGAATTAGCTTGGCTGCAGGTCGACTCGAGCGGCCGCTAGAAGTCATTACTATATCCATGAGG-3'. The amplified fragments were digested by *Eco*RI and *Not*I and ligated to the DNA-binding domain of LexA derived from the plasmid pEG202 digested with the appropriate restriction enzymes. Plasmid pSH18-34 contains the *GAL1*-lexop-lacZ reporter gene from which the upstream activation sequence *GAL1* had been deleted and replaced by four lexA operators (Hanes and Brent 1989). Plasmid pFC-W contains the -263/-234 upstream activation sequence of the *FET3* gene inserted into the *CYC1* promoter and fused to the *LacZ* gene (Yamaguchi-Iwai *et al.* 1996). All PCR-generated sequences were confirmed by DNA sequencing. All yeast transformations were performed by the lithium acetate method.

Media, growth conditions, and plate assays

Yeast strains were grown in rich medium (1% yeast extract, 2% peptone) containing either 2% glucose (YPD) or 2% glycerol and in iron- and copper-limiting yeast nitrogen base medium (Bio101 minus iron and copper) containing 0.17% yeast nitrogen base without iron or copper, 0.5% ammonium sulfate, 2% glucose, and the required amino acids and bases. For all experiments, the cells were pregrown at 30° in Bio101 minus iron and copper, supplemented with 1 μ M ferric ammonium sulfate. The cultures were then diluted to OD₆₀₀ = 0.3 and grown in Bio101 minus iron and copper except for ferrichrome uptake experiments for which the diluted cells were shifted for 6 hr to YPD medium with 200 μ M bathophenanthroline disulfonic acid (BPS). For plate assays, cells were suspended in water at 2×10^6 cells/ml, and serial 10-fold dilutions were plated onto YPD without BPS (+Fe) or with 200 μ M BPS (-Fe) or onto rich medium containing 2% glycerol and incubated at 30° for 3–4 days prior to imaging.

Ferrichrome uptake

Desferri-ferrichrome [purchased from Sigma (St. Louis)] was loaded with ⁵⁵Fe (30 mCi·mg⁻¹). Iron uptake was

assayed in microtitration plates (Lesuisse *et al.* 2001). Yeast cultures were harvested by filtration and incubated in 50 mM citrate buffer (pH 6.5) containing 5% glucose for 30 min at 30° with 1 μ M ⁵⁵Fe-ferrichrome (final concentration), collected with a cell harvester (Brandel), and washed on the filter with water.

β -Galactosidase assay

β -Galactosidase was assayed using *o*-nitrophenyl-D-galactopyranoside as described previously (Guarente 1983).

Electrophoretic mobility shift assays

Cultures were prepared as described above and the cell extracts were prepared as described previously (Conde e Silva *et al.* 2009). For mobility shift assays, the binding reaction mixtures (20 μ l) contained 25 mM HEPES, pH 7.6, 60 mM potassium chloride, 7.5% glycerol, 0.1 mM EDTA, 1 mM dithiothreitol, and 5 mM MgCl₂. Aliquots of 20–40 μ g of cell extract and 1 μ g of poly(dIdC)₂ were used. The DNA probe consisted of a 155-pb DNA fragment containing the -251 TGCACCC sequence of the *FET3* promoter. It was produced by PCR from pFC-W (Yamaguchi-Iwai *et al.* 1996) as the template, with the following primers: 5'-CAATGGAATAGGAACTTTG-3' and 5'-GTGTCAGCACTAAAGTTGC-3'. Probes were end labeled with [γ -³²P]dATP. Approximately 10,000 cpm of probe (1.7 ng) was used in each binding mixture. Samples were incubated for 30 min at room temperature. Competition assays were performed by adding double-stranded DNA fragments made with unlabeled oligonucleotides to the reaction mixtures and incubating for a further 30 min. To test the presence of the LexA fusion proteins in the formed complexes, we used anti-LexA monoclonal antibodies (Santa Cruz Biotechnology). To test the specificity of the complex formation with respect to the TGCACCC element, competition experiments were performed with 20 pmol of double-stranded oligonucleotides centered on the wild-type TGCACCC [wild type (WT), 5'-CTTCAAAAGTGCACCCATTTGCAGGTGCTC-3'] or mutated TGCAGGG (M1, 5'-CTTCAAAGTGCAGGGATTTGCAGGTGCTC-3'). Samples were loaded onto a 5% polyacrylamide gel in 0.25 $\times$ TBE (22 mM Tris, pH 8.3, 22 mM boric acid, 0.6 mM EDTA) that had been preelectrophoresed for 1 hr at 7.5 V/cm at 7°. Gels were run at 15 V/cm at 7° for 4 hr, dried, and autoradiographed for 16 hr with an intensifying screen.

Genome and protein sequences and detection of orthology and homology

Data for *S. cerevisiae* (strain S288C) were obtained from the *Saccharomyces* Genome Database (SGD) (www.yeastgenome.org); data for *Eremothecium cymbalariae*, *Kazachstania africana*, *K. naganishii*, *L. kluyveri*, *L. waltii*, *Naumovozyma castellii*, *S. bayanus*, *Tetrapisispora blattae*, *T. phaffii*, *Torulaspora delbrueckii*, *Vanderwaltozyma polysporus*, and *Zygosaccharomyces rouxii* were from <http://wolfe.gen.tcd.ie/ygob/>; data for *Debaryomyces hansenii*, *E. gossypii*, *K. lactis* ATCC 8585, *L. thermotolerans*, and *Y. lipolytica* were from <http://www>.

Figure 1 Variable sequences in the inserted region of the Aft-type DNA-binding domain. Numbers above the multiple-sequence alignment correspond to the position of the residues in Aft1 (i.e., Scer_Aft1). Charged amino acids are highlighted in red (positive) or blue (negative) and predicted phosphorylated sites are shown in green. The phylogenetic tree of the AFT sequences was adapted from the Yeast Genome Order Browser (YGOB) website information (species phylogeny and synteny) and Wang *et al.* (2009). On this tree, the whole-genome duplication (WGD) and the CTG clade are indicated. Branches leading to the genera *Vanderwaltozyma* and *Tetrapispora* are shown in gray because they have lost one AFT copy; they represent the lineage among post-WGD species most distant from *S. cerevisiae* such that it is difficult to infer the orthologs of the remaining copy. Hence, it was arbitrarily decided to consider these copies as AFT1-like genes because the variable sequence is present in the inserted region of their Aft-type DNA-binding domain. Species abbreviations: Scer, *Saccharomyces cerevisiae*; Sbay, *S. bayanus*; Cgla, *Candida glabrata*; Kafr, *Kazachstania Africana*; Knag, *K. naganishii*; Ncas, *Naumovozyma castellii*; Ndai, *N. dairenensis*; Tbla, *Tetrapispora blattae*; Tpha, *T. phaffii*; Vpol, *Vanderwaltozyma polyspora*; Zrou, *Zygosaccharomyces rouxii*; Tdel, *Torulaspora delbrueckii*; Lklu, *Lachancea Kluyveri*; Lthe, *L. thermotolerans*; Lwal, *L. waltii*; Klac, *Kluyveromyces lactis*; Egos, *Eremothecium gossypii*; Ecy, *E. cymbalariae*; Calb, *C. albicans*; Psti, *Pichia stipitis*; Dhan, *Debaryomyces hansenii*.

ebi.ac.uk/integr8; data for *Candida albicans* SC5314, Assembly 21 and *C. glabrata* CBS138 were from <http://www.candidagenome.org>; and data for *Pichia stipitis* were from <http://www.ncbi.nlm.nih.gov>.

Comparison of the gene expression data for *S. cerevisiae* and *K. lactis* revealed an ancestral core of genes the expression of which is clearly iron regulated and dependent on either Aft1 or KlAft (Conde e Silva *et al.* 2009). We used the corresponding genes in *S. cerevisiae* to query the databases and combined the results with the following approaches and databases to find their orthologs when possible or construct the whole gene family in the various hemiascomycetes: (i) the Yeast Gene Order Browser (YGOB) (Byrne and Wolfe 2005), which considers both sequence similarity and genomic context (synteny); (ii) comparisons of proteome pairs with InParanoid v4.1 (O'Brien *et al.* 2005), which clusters best-matching pairs and inparalogs; (iii) the database of ortholog groups, orthoDB (Waterhouse *et al.* 2011); and (iv) for orthologous relations that were too difficult to infer, for the siderophore transporter- and metalloredutase-encoding gene families, homologs were selected using

BLASTP with an arbitrary *E*-value cutoff of $1e-30$, using the *S. cerevisiae* proteins as queries.

Identification of iron regulatory elements

The gene sets obtained (see Supporting Information, File S1), composed of genes in each species homologous to Aft1/KlAft iron-regulated genes, have a strong probability of being regulated both by iron and by an Aft-type protein (if one exists). Hence, they were considered to be iron regulons. For each species, we analyzed the sequences between coordinates -1 and -800 , upstream from the start of the coding sequence (CDS) of the genes in the sets. The MEME program (Bailey and Elkan 1994) was used to find elements highly conserved in these sets of sequences. This program was used for words of between 6 and 20 nucleotides, with the *anr* model, allowing zero or multiple occurrences of an element per sequence. The statistical significance of an element identified by this method is based on its log-likelihood ratio, its width and number of occurrences, the background nucleotide frequencies, and the size of the training set. Here, the background model used (fourth-order Markov model)

was the letter frequencies in all upstream gene sequences from the same species as the set analyzed.

Protein sequence analysis

The MAFFT E-INS-i algorithm was used for multiple Aft-type protein sequence alignment with default settings (Katoh *et al.* 2005; Katoh and Toh 2008) and the alignment was refined by hand by using the graphical multiple-sequence alignment editor SeaView (Galtier *et al.* 1996). For each ungapped position on Aft1 in the alignment, the percentage of sequences sharing the Aft1 residue was taken as a measure of sequence conservation. Disopred2 (Ward *et al.* 2004) was used to search for intrinsically disordered regions within the protein sequences. Three different software tools were used to identify putative phosphorylation sites: KinasePhos (Huang *et al.* 2005), NetPhos 2.0 (Blom *et al.* 1999), and DISPHOS 1.3 (Iakoucheva *et al.* 2004). A site was considered to be a potential phosphorylation site if predicted by two or three of the independent predictors, as in Prevorsevsky *et al.* (2011). All prediction algorithms were run with complete proteins and default settings.

Results

Conservation and composition of the Aft DNA-binding domain

The DNA-binding domain of the Aft-type proteins contains a region of variable length between two cysteine residues that have been suggested to participate in a zinc-finger domain; this region had been named the insert by Babu *et al.* (2006). The insert sequences (corresponding to residues 143–215 in Aft1) of all known Afts proteins were aligned (Figure 1): most include a variable central region flanked by two regions rich in basic amino acids. However, the Aft2 orthologs and the Afts of the yeasts from the CTG clade (e.g., *C. albicans*) lacked the central region such that their inserts are smaller. The VS (corresponding to residues 152–203 in Aft1) was compared with conserved flanking sequences of the DNA-binding domains, named CS1 and CS2 (Figure 2). VS was predicted to be more disordered than CS1 and CS2 (Figure 2B; paired Mann–Whitney tests, P -value < 0.001). On average, $> 82\%$ of the residues in all VS but only 8% of those in CS1 and 5% in CS2 were predicted to be disordered. In Aft1, 56% of the VS region residues and 11% and 5% for CS1 and CS2, respectively, were predicted to be disordered. Consistent with observations that phosphorylation sites are frequent in disordered regions in proteins (Iakoucheva *et al.* 2004), VS overall were also predicted to be more phosphorylated (on average 12% of the residues) than CS1 (1%) and CS2 (4%) (Figure 2C; paired Mann–Whitney tests, P -value < 0.001). Fifteen percent of the residues in Aft1(VS) were predicted to be phosphorylated whereas only 6% of those in CS2 and none of those in CS1 were predicted to be phosphorylated in Aft1. Searches in databases of experimental mass spectra indicated that the Aft1 (173–203) tryptic peptide was found to be phosphorylated

Figure 2 VS are predicted to be disordered and enriched in potentially phosphorylated residues. (A) Sequence conservation profile of VS-containing Aft-like proteins compared to Aft1. The 18 VS-containing Aft-like proteins correspond to Aft1-like proteins and sequences from pre-WGD species except the CTG clade. Conservation compared to Aft1 is based on a MAFFT E-INS-i multiple-sequence alignment. The DNA-binding domain (residues 1–313) and the activation domain (413–572) of Aft1 as defined by Rutherford *et al.* (2001) and Yamaguchi-Iwai *et al.* (2002) are indicated below the profile. The variable sequence (VS: residues 152–203) and the two flanking conserved sequences (CS1, 96–151; and CS2, 204–313) are shown above the profile. (B) The VS of Aft-like proteins are expected to be intrinsically disordered. The average percentages of residues predicted to be disordered by DISOPRED2 (Ward *et al.* 2004) are shown with error bars indicating the standard deviation. (C) Compared to the flanking conserved sequences CS1 and CS2, the VS sequences are also enriched for potential phosphorylation sites as predicted by independent algorithms. The average percentages of amino acid predicted to be phosphorylated are shown with error bars indicating the standard deviation.

on the Ser181 and Ser183 residues (PhosphoPep Database). Thus, although VS showed little protein sequence conservation, they appeared to share at least two common properties: being highly disordered and containing a high density of phosphorylation sites.

Aft1(152–203) is required for the binding to the TGCACCC sequence

Conserved disordered regions with rapidly evolving amino acid sequences, such as VS, were recently defined as flexible regions strongly associated with signaling and regulation (Belay *et al.* 2011). To assess the involvement of Aft1(VS) in the regulatory function of Aft1, we performed deletion mutation analyses (Figure 3). We constructed a mutant deleted for the entire 152–203 region and mutants deleted for the 152–181, 182–203, and 172–203 subregions (namely $\Delta 152$ –203, $\Delta 152$ –181, $\Delta 182$ –203, and $\Delta 172$ –203, respectively, Figure 3A) and evaluated their *trans*-activation and DNA-binding functions. To test the *trans*-activation function, transcriptional fusions of the proteins to the bacterial LexA DNA-binding domain were constructed and expressed in wild-type cells carrying a LacZ reporter gene driven by four LexA operators (Figure 3B). Equivalent levels of β -galactosidase activity were detected in cells expressing the LexA-Aft1 fusion construct and those expressing the

Figure 3 Deletion of the variable 152–203 Aft1(VS) sequence affects the recognition of the TGCACCC sequence, ferrichrome uptake, and cell growth in iron-depleted conditions. (A) Structure of the Aft1-deleted regions. Shown is a plot of the percentages of identity along the Aft1(143–215) DNA-binding domain insert as described in the legend to Figure 2. The variable 152–203 Aft1(VS) sequence is shown as a shaded box. The structures and names of the four deleted regions are depicted below. (B and C) β-Galactosidase activity. LexA-Aft1 deleted proteins have a full capacity for *trans*-activation (B) but are affected for the recognition of the AFT1 DNA-binding site (C). LexA, LexA-Aft1, and LexA-Aft1 deleted derivatives were expressed in wild-type (B) and *aft1Δaft2Δ* mutant strains (C) transformed with either pSH18-34 containing four *lexA* operators fused to the *LacZ* gene (B) or pFC-W containing the activation sequence of the *FET3* gene fused to the *LacZ* gene (C). Appropriate transformants were pregrown at 30° in Bio101 minus iron and copper supplemented with 1 μM ferric ammonium sulfate. Cultures were then diluted to OD₆₀₀ = 0.3 and then grown exponentially to an OD₆₀₀ = 1 in Bio101 minus iron and copper. Results are means ± SE from experiments with three independent transformants. (D) Ferrichrome uptake. Cultures of *aft1Δaft2Δ* cells producing LexA-Aft1 and LexA-Aft1 deleted proteins were pregrown as in B and C. Cultures were then diluted to OD₆₀₀ = 0.3 and shifted to YPD medium with 200 μM of the iron chelator BPS for 6 hr. Iron uptake was measured as described in *Materials and Methods*. Results are means ± SE from four experiments. (E) Growth of *aft1Δaft2Δ* cells expressing LexA-Aft1 or the LexA-Aft1 deleted derivatives. Cells were pregrown as in B and C and plated onto YPD with 200 μM BPS (–Fe) or rich medium containing 2% glycerol.

LexA- deleted Aft1 constructs. This demonstrates that all the deleted mutants retained full *trans*-activation activity and suggests that the protein stability of Aft1 is not affected by the deletions. To evaluate the effect of the deletions on the DNA-binding function of Aft1, the LexA-deleted Aft1 constructs were expressed in *aft1Δaft2Δ* cells carrying a *LacZ* reporter gene downstream from the Aft1-binding site, TGCACCC. Thus, the β-galactosidase activity gave a measure of the activation of transcription at the Aft1-binding site. As previously described (Yamaguchi-Iwai *et al.* 1996; Courel *et al.* 2005), Aft1 exhibited a strong level of β-galactosidase activity (Figure 3C). β-Galactosidase activity was 5.5-fold lower with the 152–203 deletion and 2.2- and 1.5-fold lower with the 152–181 and 182–203 deletions, respectively. The construct with the deletion of the 172–203 region resulted in slightly higher β-galactosidase activity (1.2-fold higher). The deletion of the region 152–203 did not affect the level of Aft1 *trans*-activation *per se* (Figure 3B), but substantially decreased the activation of transcription at the Aft1-binding

site (Figure 3C); this indicates that the 152–203 region is required for the recognition of the TGCACCC sequence.

To confirm this result and investigate more directly the role of the 152–203 region for the binding to the TGCACCC sequence, we performed electrophoretic mobility shift assays with extracts from the *aft1Δaft2Δ* cells expressing LexA, LexA-Aft1, or LexA-deleted Aft1 constructs and a probe containing the TGCACCC Aft1-binding site of the *FET3* promoter. Western blotting with extract proteins used for mobility shift assays and anti-LexA antibodies confirm that the protein stability of the Aft1 protein is not affected by the deletions (see Figure S1). Extracts containing LexA-Aft1 fusion protein gave a high-mobility complex (Figure 4, lane 2) that is absent in extracts containing the sole LexA protein (Figure 4, lanes 1 and 12). Addition of anti-LexA antibodies overshifted this complex, but had no effect on other mobility complexes (Figure 4, lane 3). This demonstrated that LexA-Aft1 is a component of the higher-mobility complex formed with the *FET3* promoter sequence. Extracts containing

Figure 4 The Aft1(152–203) region is required for the binding to the TGCACCC sequence of the *FET3* promoter sequence. Gel-shift assays were carried out with extracts from *aft1Δaft2Δ* cells expressing LexA (C, lanes 1 and 12), LexA-Aft1 (lanes 2 and 3), and LexA-Aft1 deleted derivatives: LexA-Aft1(Δ152–203) (lanes 4 and 5), LexA-Aft1(Δ152–181) (lanes 6 and 7), LexA-Aft1(Δ182–203) (lanes 8 and 9), and LexA-Aft1(Δ172–203) (lanes 10 and 11). The radiolabeled probe used to perform the gel-shift assays corresponds to a DNA fragment containing the –251 TGCACCC Aft1-binding site of the *FET3* promoter. The arrowhead indicates the Aft1-containing complexes. Where indicated, 2 μ l of a monoclonal antibody raised against the LexA epitope was added (lanes 3, 5, 7, 9, 11, and 12).

LexA-Aft1(Δ152–181), LexA-Aft1(Δ182–203), or LexA-Aft1(Δ172–203) exhibit also a complex of high mobility that is overshifted with addition of anti-LexA antibodies (Figure 4, lanes 6–11). In contrast, extracts containing LexA-Aft1(Δ152–203) did not exhibit the higher-mobility complex (Figure 4, lanes 4 and 5). This showed that the formation of a LexA-Aft1-containing complex is abolished by the 152–203 deletion but not by the 152–181, 182–203, or 172–203 deletions. In accord with the results presented in Figure 3C, this demonstrated that the Aft1(152–203) region is required for the binding to the TGCACCC sequence of the *FET3* promoter sequence.

Next, we address the functional importance of the Aft1(152–203) region for iron homeostasis. Aft1 is the major transcriptional regulator of iron transport systems, so we first conducted iron uptake experiments with the siderophore ferrichrome (FCH). FCH transport is mediated by the Aft1-regulated genes *ARN1* and *SIT1* and both their promoter regions contain the Aft1-binding site TGCACCC (Yun *et al.* 2000a,b). We compared the rate of FCH-⁵⁵Fe uptake by *aft1Δaft2Δ* cells expressing LexA-Aft1 with that of cells expressing the LexA-deleted Aft1 constructs (Figure 3D). FCH accumulation was substantially lower (4.1-fold) in cells expressing LexA-Aft1(Δ152–203) than in cells expressing LexA-Aft1, consistent with a decrease of transcription activity at the Aft1-binding site (Figure 3C). FCH accumulation by cells expressing the other deletion constructs was strictly correlated with the corresponding levels of transcription

activity at the Aft1-binding site. The tight correlation between the recognition of the Aft1-binding site and FCH uptake activity (Figure 3, C and D) confirms the importance of Aft1 binding to the TGCACCC promoter sequence for the regulation of FCH uptake.

We compared the growth on agar plates of the *aft1Δaft2Δ* mutant expressing LexA-Aft1 or the LexA-deleted Aft1 constructs. The *aft1Δaft2Δ* mutant is unable to grow in iron-depleted conditions or on carbon sources like raffinose or glycerol that require functional respiratory activity for their catabolism (Blaiseau *et al.* 2001; Rutherford *et al.* 2001). Consistent with previous results (Conde e Silva *et al.* 2009), the *aft1Δaft2Δ* mutant phenotypes were complemented by the expression of LexA-Aft1 (Figure 3E). However, the *aft1Δaft2Δ* cells expressing LexA-Aft1(Δ152–203) grew poorly in iron-depleted conditions and were unable to grow with glycerol as the only carbon source. The *aft1Δaft2Δ* cells expressing either LexA-Aft1(Δ152–181) or LexA-Aft1(Δ182–203) showed slight but reproducible growth deficiency on glycerol, whereas the growth of those expressing LexA-Aft1(Δ172–203) was similar to that of those expressing LexA-Aft1. These results demonstrate that deletion of the 152–203 region of Aft1 reduced the complementation both of the FCH uptake and of the growth deficiency of *aft1Δaft2Δ* mutants in iron-depleted conditions.

The insert of Aft1 (residues 143–215) is necessary and sufficient for the binding of KLAft to the TGCACCC sequence

The DNA-binding mechanisms of Aft1 and KLAft may be different (Conde e Silva *et al.* 2009). The inserts in Aft1 and KLAft are not well conserved (Figure 1), and this divergence may be responsible for differences in DNA binding. We used swapping experiments to test this possibility. As the limits of the regions not conserved between Aft1 and KLAft are uncertain (Figure 1), we constructed hybrid proteins by exchanging the entire Aft1(143–215) and KLAft(134–229) inserts (Figure 5A). The *trans*-activation and DNA-binding functions of Aft1, KLAft, Aft1^{KLAft(134–229)}, and KLAft^{Aft1(143–215)} proteins were studied as described in Figure 3, B and C. The native and hybrid proteins exhibited similar *trans*-activation abilities (Figure 5B). In contrast, the levels of transcription activity at the TGCACCC sequence differed substantially between Aft1 and KLAft and also between the native and the hybrid proteins (Figure 5C). The level of β -galactosidase activity in cells expressing KLAft was 35-fold lower than that in those expressing Aft1, indicating that KLAft recognized the TGCACCC sequence very much more poorly than Aft1. The replacement of Aft1(143–215) by KLAft(134–229) resulted in reciprocal opposite effects. The β -galactosidase activity in cells expressing the Aft1^{KLAft(134–229)} hybrid was strongly decreased compared to that in those expressing Aft1. In contrast, the β -galactosidase activity in cells expressing the KLAft^{Aft1(143–215)} hybrid was strongly increased. These findings indicate that Aft1(143–215) is necessary and sufficient for KLAft binding to the TGCACCC sequence.

Figure 5 The Aft1(143–215) region is necessary and sufficient for KIAft to recognize the TGCACCC sequence and complement the growth deficiency of *aft1Δaft2Δ* cells in iron-depleted conditions. (A) Schematic representation of the native Aft1 and KIAft and hybrid KIAft^{Aft1(143–215)} and Aft1^{KIAft(134–229)} proteins in which the Aft1(143–215) and KIAft(134–229) DNA-binding domain inserted regions have been exchanged. (B and C) β-Galactosidase activity. LexA-Aft1, LexA-KIAft, KIAft^{Aft1(143–215)}, and Aft1^{KIAft(134–229)} proteins have a similar capacity for *trans*-activation (B) and Aft1(143–215) is necessary and sufficient for KIAft to recognize the TGCACCC sequence (C). Native and hybrid proteins were expressed in wild-type (B) and *aft1Δaft2Δ* mutant strains (C) transformed with either pSH18-34 containing four *lexA* operators fused to the *LacZ* gene (B) or pFC-W containing the activation sequence

of the *FET3* gene fused to the *LacZ* gene (C). Strains, *LacZ* repertory systems, and culture conditions were as described in the legend to Figure 3. Results are means ± SE from experiments with three independent transformants. (D) Growth of *aft1Δaft2Δ* cells expressing pCXJ22 (Control), pCXJ22-Aft1, pCXJ22-KIAft, and pCXJ22-KIAft^{Aft1(143–215)}. Cells were pregrown as in B and C and plated onto YPD without BPS (+Fe) or with 200 μM BPS (–Fe).

To confirm this important point, we performed electrophoretic mobility shift assays with extracts from the *aft1Δaft2Δ* cells expressing LexA, LexA-Aft1, LexA-KIAft, or LexA-KIAft^{Aft1(143–215)} and the probe containing the TGCACCC Aft1-binding site of the *FET3* promoter. Western blotting checked that the LexA-KIAft and LexA-KIAft^{Aft1(143–215)} proteins were well expressed, at the expected size, in *S. cerevisiae* (see Figure S1). As shown in Figure 6, protein extracts containing either LexA-Aft1 (lane 2) or LexA-KIAft^{Aft1(143–215)} (lane 8) showed a high-mobility complex that is overshifted with addition of anti-LexA antibodies (lanes 3 and 9). In contrast, extracts containing LexA-KIAft (Figure 6, lanes 6 and 7) showed no significant difference from those containing LexA (Figure 6, lanes 1 and 12). To test the specificity of the complex formed with LexA-Aft1 and LexA-KIAft^{Aft1(143–215)}, we performed competition assays with addition of unlabeled oligomers centered on wild-type TGCACCC (WT) or mutated TGCAGGG (M) sequences. The high-mobility complex formed with both LexA-Aft1 and LexA-KIAft^{Aft1(143–215)} was specifically displaced by the addition of excess wild-type but not mutated oligomers (Figure 6, lanes 4 and 5 and 10 and 11). Therefore, the binding of LexA-Aft1 and LexA-KIAft^{Aft1(143–215)} was specific for the TGCACCC sequence. Thus, the Aft1(143–215) region confers on KIAft the capacity to bind specifically to the TGCACCC sequence. We investigated iron homeostasis in cells expressing KIAft^{Aft1(143–215)}. We compared the growth of *aft1Δaft2Δ* cells transformed with monocopy plasmids carrying the *AFT1*, the *KIAFT*, or the *KIAFT*^{Aft1(143–215)} genes controlled by their own promoter (Figure 5D). As previously reported (Conde e Silva *et al.* 2009), the growth defect of the *aft1Δaft2Δ* mutant was not complemented by the plasmid carrying *KIAFT*. However, the plasmid carrying *KIAFT*^{Aft1(143–215)} restored growth to the *aft1Δaft2Δ* mutant,

and therefore KIAft was able to complement the *aft1Δaft2Δ* iron-dependent phenotype only if the insert of KIAft was replaced with that of Aft1. This demonstrates the functional importance of the Aft1 insert in the regulation of iron homeostasis in *S. cerevisiae*.

Figure 6 The insert of Aft1 (143–215) is necessary and sufficient for the binding of KIAft to the TGCACCC sequence. Gel-shift assays were carried out with extracts from *aft1Δaft2Δ* cells expressing LexA (C, lanes 1 and 12), LexA-Aft1 (lanes 2–5), LexA-KIAft (lanes 6 and 7), and LexA-KIAft^{Aft1(143–215)} (lanes 8–11). The radiolabeled probe used to perform the gel-shift assays corresponds to a DNA fragment containing the –251 TGCACCC Aft1-binding site of the *FET3* promoter. The arrowhead indicates the Aft1 and KIAft-Aft1(143–215)-containing complexes. Where indicated, 2 μl of an antibody raised against LexA was added (lanes 3, 7, 9, and 12). Competitive assays were performed with excess oligonucleotide (competitor) centered on the wild-type TGCACCC sequence (WT, lanes 4 and 10) and mutant TGCAGGG sequence (M, lanes 5 and 11).

Figure 7 Divergence of the *cis*-regulatory sequence between the iron regulons of different hemiascomycete yeasts. Motif logos were constructed using WebLogo (Crooks *et al.* 2004) from motif sequences obtained, as *e*-values, using MEME with gene upstream sequences from inferred iron regulons (see *Materials and Methods* and *File S1*). Upstream sequences in *K. lactis* have diverged toward a PuCACC motif. For each species, the number of genes analyzed is indicated to the left of the *e*-values. The phylogenetic relationships between species were acquired from Conde e Silva *et al.* (2009). The whole-genome duplication (WGD) and the CTG clade are indicated on the tree. The genus names of species are given in the legend to Figure 1, except for *Y. lipolytica*, which stands for *Yarrowia lipolytica*.

The TGCACCC sequence is widespread among promoter regions of genes of the iron regulon in hemiascomycetes

Previous computer analyses identified TGCACCC and PuCACC Aft-type sequences as consensus sequences in the promoters of iron-regulated genes in *S. cerevisiae* and *K. lactis*, respectively (Rutherford *et al.* 2003; Courel *et al.* 2005; Conde e Silva *et al.* 2009). To study Aft-type *cis*-regulatory sequences and their evolution, we extended our analyses to diverse species belonging to hemiascomycete clades and for which the genome sequences are available. Comparison of gene expression in *S. cerevisiae* and *K. lactis* allowed us to identify a group of Aft-target genes, the expression of which is both iron regulated and dependent on either Aft1 or KlAft. Homologs of this group of genes, referred to here as the iron regulon, were identified in several hemiascomycete species. We analyzed their promoter regions (see *Materials and Methods*), and the Aft-type element was found in all hemiascomycetes (Figure 7), except in *Y. lipolytica* and in species of the CTG clade (*C. albicans*,

D. hansenii, and *P. stipitis*). Unlike the other species analyzed, *Y. lipolytica* and the CTG clade species possess an ortholog of the iron-responsive GATA-type repressor (Conde e Silva *et al.* 2009) that most probably controls their iron-regulatory pathway, as in most fungi. This has been demonstrated in *C. albicans* and *P. pastoris* (Lan *et al.* 2004; Miele *et al.* 2007). The Aft-type sequence identified in other hemiascomycetes was clearly TGCACCC except in the phylogenetic branch leading to *E. gossypii* and *K. lactis*. *E. gossypii* had a small predicted iron regulon (nine genes) and possessed the genome with the highest GC content in the analysis. This led to the detection of an Aft-type sequence with a low confidence level for *E. gossypii*. As previously reported, (Conde e Silva *et al.* 2009), we found the shorter and less restrictive Aft-type PuCACC sequence in *K. lactis* (Figure 7). The results of this study revealed mainly that the TGCACCC sequence is widespread in the promoter regions of genes of the iron regulons in hemiascomycetes. This suggests that the TGCACCC sequence corresponds to the ancestral Aft-type element and that it has diverged

Figure 8 LtAft, the *Lachancea thermotolerans* ortholog of Aft1 and KIAft, binds to the TGCACCC sequence and complements the *aft1Δaft2Δ* iron-dependent phenotype. Gel-shift assays were carried out with extracts from *Δaft1Δaft2Δ* cells expressing LexA-LtAft. The radiolabeled probe used to perform the gel-shift assays corresponds to a DNA fragment containing the -251 TGCACCC Aft1-binding site of the *FET3* promoter. The arrow-head indicates the LtAft-containing complexes. Where indicated, $2 \mu\text{l}$ of an antibody raised against LexA was added (lane 2). Competitive assays were performed with excess oligonucleotide (competitor) centered on the wild-type TGCACCC sequence (WT, lane 3) and the mutant TGCAGGG sequence (M, lane 4). (B) Growth of *aft1Δaft2Δ* cells expressing LexA (Control), LexA-Aft1, LexA-KIAft, and LexA-LtAft. Cells were plated onto YPD without BPS (+Fe) or with $200 \mu\text{M}$ BPS (–Fe).

toward the PuCACCC sequence in a recent ancestor of *K. lactis*.

LtAft, the *L. thermotolerans* ortholog of Aft1 and KIAft, binds to the TGCACCC sequence and complements the *aft1Δaft2Δ* iron-dependent phenotype

To confirm the results of the bioinformatic analysis described above, we tested the ability of LtAft, the *L. thermotolerans* ortholog of Aft1 and KIAft, to bind the TGCACCC sequence. We choose *L. thermotolerans* because the *Lachancea* species have a common ancestor with *K. lactis* more recent than the common ancestor of *K. lactis* and *S. cerevisiae* and the *e*-values associated with the predicted TGCACCC sequence identified in the upstream sequences of the *L. thermotolerans* iron regulon were of high confidence (Figure 7). A LexA-LtAft fusion protein was constructed and we performed electrophoretic mobility shift assays with extracts from the *aft1Δaft2Δ* cells expressing LexA-LtAft and the probe containing the TGCACCC Aft1-binding site of the *FET3* promoter. Extracts containing LexA-LtAft gave a mobility complex of high intensity with the *FET3* promoter (Figure 8A, lane 1). The mobility complex formed with LexA-LtAft was overshifted with addition of anti-LexA antibodies (Figure 8A, lane 2) and specifically displaced by the addition of excess unlabeled oligomers centered on the wild-type TGCACCC (WT) sequence but not by those centered on the TGCAGGG (M) mutated sequences (Figure 8A, lanes 3 and 4). These results demonstrated that, in contrast to KIAft, LtAft was able to bind the TGCACCC Aft1-binding site in *S. cerevisiae*. We next compared the ability of KIAft and LtAft to complement the *aft1Δaft2Δ* iron-dependent phenotype. The deficiency of *aft1Δaft2Δ* cells to grow in iron-depleted conditions is not suppressed with a plasmid expressing LexA-KIAft whereas it is totally restored with a plasmid expressing LexA-LtAft (Figure 8B). Thus, LtAft was able to bind the TGCACCC sequence and to complement the iron-dependent phenotype of the *S. cerevisiae*

aft1Δaft2Δ cells. In accord with the bioinformatic data, these experimental results reinforce the hypothesis that the TGCACCC sequence is ancestral and suggest that it diverged toward the PuCACCC sequence after the divergence of the *Lachancea* yeasts from the phylogenetic branch leading to *K. lactis*.

Discussion

The Aft1 DNA-binding domain has been described as containing a region of variable length (residues 143–215) between two cysteine residues (Babu *et al.* 2006). Our deletion analyses show that the 152–203 sequence, contained in the insert region, is necessary for the DNA-binding function of Aft1. The DNA-binding capacity of Aft1 is also affected, but to a much lesser extent, by the deletion of N-terminal (152–181) and C-terminal (182–203) subsequences. We showed a strict correlation between the variation in the TGCACCC-binding capacity of deleted Aft1 variant proteins and the effect on both FCH uptake activity and the ability to grow in iron-depleted conditions (Figure 3, C–E). This tight causal link between the DNA-binding level of Aft1 and metabolic effects indicates that binding of Aft1 to TGCACCC is of crucial importance to the control of iron homeostasis in *S. cerevisiae*. This agrees with a recent study showing that the iron-dependent dissociation of Aft1 from the promoter is the critical step in the regulation of iron metabolism (Ueta *et al.* 2012). Thus, unlike what was previously suggested (Yamaguchi-Iwai *et al.* 2002; Ueta *et al.* 2003, 2007), the way Aft1 regulates iron homeostasis depends more on its capacity to bind its target DNA sequence than on its capacity to shift from the cytosol to the nucleus. Comparative sequence analysis of the DNA-binding domains of various Aft proteins showed that Aft1(152–203), which we named “VS”, is a rapidly evolving sequence, bordered by two conserved regions designated CS1 and CS2. VS was predicted to

be highly disordered and phosphorylated. Many such are involved in protein/protein and/or protein/nucleic acid interactions (Dunker *et al.* 2002). Thus, Aft1(152–203) is a good candidate region to interact with DNA or other (as yet unidentified) proteins. Paradoxically, although Aft1(152–203) is essential for Aft binding to the nucleotide sequence TGCACCC in *S. cerevisiae*, it is not conserved.

Previous work indicated that Aft1 in *S. cerevisiae* binds to the sequence TGCACCC whereas K1Aft in *K. lactis* binds to the sequence PuCACCC (Yamaguchi-Iwai *et al.* 1996; Conde e Silva *et al.* 2009). Here, we compared the ability of K1Aft and Aft1 to bind to the TGCACCC Aft1-binding site of the *FET3* promoter in the same *S. cerevisiae* cellular context. Mobility shift assays and transcription analyses in *aft1Δaft2Δ* mutant cells show that Aft1 but not K1Aft binds efficiently to the TGCACCC sequence in *S. cerevisiae* cells. Reciprocally, mobility shift assays with a probe containing the CACACCC K1Aft-binding site of the KLLA0E14652g *K. lactis* promoter show that K1Aft but not Aft1 binds efficiently to the CACACCC sequence in *S. cerevisiae* (see Figure S2). Taken together, these results indicate that Aft1 and K1Aft present different DNA-binding sites preference. This difference is presumably due to the large differences between the amino acid sequences of VS in K1Aft and Aft1 (Figure 1). To test this, we constructed a hybrid protein, K1Aft^{Aft1(143–215)}, which allowed us to demonstrate that the Aft1(143–215) sequence inserted into K1Aft was sufficient for binding to TGCACCC. Also, binding of the K1Aft^{Aft1(143–215)} hybrid to TGCACCC totally suppressed the iron-dependent growth phenotype of the *aft1Δaft2Δ* mutant cells. To our knowledge, this is the first example in yeast that links the divergence of a limited DNA-binding region of a transcription regulator to major physiological effects.

These results indicate that the insert region of the DNA-binding domains of Aft1 and K1Aft coevolved with their DNA-binding sites preference. We examined the iron-regulon promoters in other hemiascomycete species (Figure 7). Contradicting our previous prediction, this more extensive study strongly suggests that the widespread TGCACCC sequence corresponds to the ancestral Aft-binding site and that this binding site evolved toward PuCACCC in the *K. lactis* lineage. Comparison of the Aft1(143–215) and K1Aft(134–229) regions may reveal the differences in the insert amino acid sequence that are associated with the divergence of the DNA-binding specificity (Figure 1). K1Aft contains an additional segment of 35 residues that interrupts a sequence enriched in basic amino acids in the N-terminal part of the Aft1 VS. This part of VS is important for the binding of Aft1 to the DNA sequence TGCACCC (Figure 3A), so the additional residues in K1Aft may affect binding to this sequence. Consistent with this hypothesis, LtAft, the *L. thermotolerans* ortholog of Aft1 and K1Aft that does not possess the additional segment, was shown able to bind TGCACCC in *S. cerevisiae* cells (Figure 8A). As shown in Figure 1, Aft2 and its orthologs lack most of the VS. This large difference may explain why Aft2, unlike Aft1, does not exhibit a preference for the TGCACCC sequence (Courel *et al.* 2005). The ability to bind effectively

to TGCACCC might have been lost independently by K1Aft and Aft2, but in both cases because of the major changes in their VS sequences.

Acknowledgments

We thank Gilles Fischer for the gift of the *L. thermotolerans* strain CBS 6340 and Marc Lemaire for the gift of the plasmid pCXJ22. We thank Véronique Proux-Gillardeaux for helpful discussions. English text was edited by Alex Edelman & Associates.

Literature Cited

- Babu, M. M., L. M. Iyer, S. Balaji, and L. Aravind, 2006 The natural history of the WRKY-GCM1 zinc fingers and the relationship between transcription factors and transposons. *Nucleic Acids Res.* 34: 6505–6520.
- Bailey, T. L., and C. Elkan, 1994 Fitting a mixture model by expectation maximization to discover motifs in biopolymers. *Proc. Int. Conf. Intell. Syst. Mol. Biol.* 2: 28–36.
- Baker, C. R., B. B. Tuch, and A. D. Johnson, 2011 Extensive DNA-binding specificity divergence of a conserved transcription regulator. *Proc. Natl. Acad. Sci. USA* 108: 7493–7498.
- Bellay, J., S. Han, M. Michaut, T. Kim, M. Costanzo *et al.*, 2011 Bringing order to protein disorder through comparative genomics and genetic interactions. *Genome Biol.* 12: R14.
- Blaiseau, P. L., E. Lesuisse, and J. M. Camadro, 2001 Aft2p, a novel iron-regulated transcription activator that modulates, with Aft1p, intracellular iron use and resistance to oxidative stress in yeast. *J. Biol. Chem.* 276: 34221–34226.
- Blaiseau, P.-L., A. Seguin, J. M. Camadro, and E. Lesuisse, 2010 Iron uptake in yeasts, pp. 265–284 in *Iron Uptake and Homeostasis in Microorganisms*, edited by P. Cornelis, and S. C. Andrews. Caister Academic Press, Brussels/Reading, UK.
- Blom, N., S. Gammeltoft, and S. Brunak, 1999 Sequence and structure-based prediction of eukaryotic protein phosphorylation sites. *J. Mol. Biol.* 294: 1351–1362.
- Byrne, K. P., and K. H. Wolfe, 2005 The Yeast Gene Order Browser: combining curated homology and syntenic context reveals gene fate in polyploid species. *Genome Res.* 15: 1456–1461.
- Conde e Silva, N., I. R. Goncalves, M. Lemaire, E. Lesuisse, J. M. Camadro *et al.*, 2009 K1Aft, the *Kluyveromyces lactis* ortholog of Aft1 and Aft2, mediates activation of iron-responsive transcription through the PuCACCC Aft-type sequence. *Genetics* 183: 93–106.
- Courel, M., S. Lallet, J. M. Camadro, and P. L. Blaiseau, 2005 Direct activation of genes involved in intracellular iron use by the yeast iron-responsive transcription factor Aft2 without its paralog Aft1. *Mol. Cell. Biol.* 25: 6760–6771.
- Crooks, G. E., G. Hon, J. M. Chandonia, and S. E. Brenner, 2004 WebLogo: a sequence logo generator. *Genome Res.* 14: 1188–1190.
- Dhaoui, M., F. Auchere, P. L. Blaiseau, E. Lesuisse, A. Landoulsi *et al.*, 2011 Gex1 is a yeast glutathione exchanger that interferes with pH and redox homeostasis. *Mol. Biol. Cell* 22: 2054–2067.
- Dunker, A. K., C. J. Brown, J. D. Lawson, L. M. Iakoucheva, and Z. Obradovic, 2002 Intrinsic disorder and protein function. *Biochemistry* 41: 6573–6582.
- Galtier, N., M. Gouy, and C. Gautier, 1996 SEAVIEW and PHYLO-WIN: two graphic tools for sequence alignment and molecular phylogeny. *Comput. Appl. Biosci.* 12: 543–548.

- Gasch, A. P., A. M. Moses, D. Y. Chiang, H. B. Fraser, M. Berardini *et al.*, 2004 Conservation and evolution of cis-regulatory systems in ascomycete fungi. *PLoS Biol.* 2: e398.
- Guarente, L., 1983 Yeast promoters and lacZ fusions designed to study expression of cloned genes in yeast. *Methods Enzymol.* 101: 181–191.
- Haas, H., M. Eisendle, and B. G. Turgeon, 2008 Siderophores in fungal physiology and virulence. *Annu. Rev. Phytopathol.* 46: 149–187.
- Halliwell, B., and J. M. Gutteridge, 1984 Oxygen toxicity, oxygen radicals, transition metals and disease. *Biochem. J.* 219: 1–14.
- Hanes, S. D., and R. Brent, 1989 DNA specificity of the bicoid activator protein is determined by homeodomain recognition helix residue 9. *Cell* 57: 1275–1283.
- Huang, H. D., T. Y. Lee, S. W. Tzeng, and J. T. Horng, 2005 KinasePhos: a web tool for identifying protein kinase-specific phosphorylation sites. *Nucleic Acids Res.* 33: W226–W229.
- Iakoucheva, L. M., P. Radivojac, C. J. Brown, T. R. O'Connor, J. G. Sikes *et al.*, 2004 The importance of intrinsic disorder for protein phosphorylation. *Nucleic Acids Res.* 32: 1037–1049.
- Katoh, K., and H. Toh, 2008 Recent developments in the MAFFT multiple sequence alignment program. *Brief. Bioinform.* 9: 286–298.
- Katoh, K., K. Kuma, H. Toh, and T. Miyata, 2005 MAFFT version 5: improvement in accuracy of multiple sequence alignment. *Nucleic Acids Res.* 33: 511–518.
- Kuo, D., K. Licon, S. Bandyopadhyay, R. Chuang, C. Luo *et al.*, 2010 Coevolution within a transcriptional network by compensatory trans and cis mutations. *Genome Res.* 20: 1672–1678.
- Lan, C. Y., G. Rodarte, L. A. Murillo, T. Jones, R. W. Davis *et al.*, 2004 Regulatory networks affected by iron availability in *Candida albicans*. *Mol. Microbiol.* 53: 1451–1469.
- Lesuisse, E., P. L. Blaiseau, A. Dancis, and J. M. Camadro, 2001 Siderophore uptake and use by the yeast *Saccharomyces cerevisiae*. *Microbiology* 147: 289–298.
- Miele, R., D. Barra, and M. C. Bonaccorsi di Patti, 2007 A GATA-type transcription factor regulates expression of the high-affinity iron uptake system in the methylotrophic yeast *Pichia pastoris*. *Arch. Biochem. Biophys.* 465: 172–179.
- O'Brien, K. P., M. Remm, and E. L. Sonnhammer, 2005 Inparanoid: a comprehensive database of eukaryotic orthologs. *Nucleic Acids Res.* 33: D476–D480.
- Prevorovsky, M., S. R. Atkinson, M. Ptackova, J. R. McLean, K. Gould *et al.*, 2011 N-termini of fungal CSL transcription factors are disordered, enriched in regulatory motifs and inhibit DNA binding in fission yeast. *PLoS ONE* 6: e23650.
- Rutherford, J. C., S. Jaron, E. Ray, P. O. Brown, and D. R. Winge, 2001 A second iron-regulatory system in yeast independent of Aft1p. *Proc. Natl. Acad. Sci. USA* 98: 14322–14327.
- Rutherford, J. C., S. Jaron, and D. R. Winge, 2003 Aft1p and Aft2p mediate iron-responsive gene expression in yeast through related promoter elements. *J. Biol. Chem.* 278: 27636–27643.
- Ueta, R., A. Fukunaka, and Y. Yamaguchi-Iwai, 2003 Pse1p mediates the nuclear import of the iron-responsive transcription factor Aft1p in *Saccharomyces cerevisiae*. *J. Biol. Chem.* 278: 50120–50127.
- Ueta, R., N. Fujiwara, K. Iwai, and Y. Yamaguchi-Iwai, 2007 Mechanism underlying the iron-dependent nuclear export of the iron-responsive transcription factor Aft1p in *Saccharomyces cerevisiae*. *Mol. Biol. Cell* 18: 2980–2990.
- Ueta, R., N. Fujiwara, K. Iwai, and Y. Yamaguchi-Iwai, 2012 Iron-induced dissociation of the aft1p transcriptional regulator from target gene promoters is an initial event in iron-dependent gene suppression. *Mol. Cell. Biol.* 32: 4998–5008.
- Wach, A., 1996 PCR-synthesis of marker cassettes with long flanking homology regions for gene disruptions in *S. cerevisiae*. *Yeast* 12: 259–265.
- Wang, H., Z. Xu, L. Gao, and B. Hao, 2009 A fungal phylogeny based on 82 complete genomes using the composition vector method. *BMC Evol. Biol.* 9: 195.
- Ward, J. J., J. S. Sodhi, L. J. McGuffin, B. F. Buxton, and D. T. Jones, 2004 Prediction and functional analysis of native disorder in proteins from the three kingdoms of life. *J. Mol. Biol.* 337: 635–645.
- Waterhouse, R. M., E. M. Zdobnov, F. Tegenfeldt, J. Li, and E. V. Kriventseva, 2011 OrthoDB: the hierarchical catalog of eukaryotic orthologs in 2011. *Nucleic Acids Res.* 39: D283–D288.
- Wolfe, K. H., and D. C. Shields, 1997 Molecular evidence for an ancient duplication of the entire yeast genome. *Nature* 387: 708–713.
- Yamaguchi-Iwai, Y., A. Dancis, and R. D. Klausner, 1995 AFT1: a mediator of iron regulated transcriptional control in *Saccharomyces cerevisiae*. *EMBO J.* 14: 1231–1239.
- Yamaguchi-Iwai, Y., R. Stearman, A. Dancis, and R. D. Klausner, 1996 Iron-regulated DNA binding by the AFT1 protein controls the iron regulon in yeast. *EMBO J.* 15: 3377–3384.
- Yamaguchi-Iwai, Y., R. Ueta, A. Fukunaka, and R. Sasaki, 2002 Subcellular localization of Aft1 transcription factor responds to iron status in *Saccharomyces cerevisiae*. *J. Biol. Chem.* 277: 18914–18918.
- Yun, C. W., T. Ferea, J. Rashford, O. Ardon, P. O. Brown *et al.*, 2000a Desferrioxamine-mediated iron uptake in *Saccharomyces cerevisiae*. Evidence for two pathways of iron uptake. *J. Biol. Chem.* 275: 10709–10715.
- Yun, C. W., J. S. Tiedeman, R. E. Moore, and C. C. Philpott, 2000b Siderophore-iron uptake in *Saccharomyces cerevisiae*. Identification of ferrichrome and fusarinine transporters. *J. Biol. Chem.* 275: 16354–16359.

Communicating editor: M. Hampsey

GENETICS

Supporting Information

<http://www.genetics.org/lookup/suppl/doi:10.1534/genetics.113.157693/-/DC1>

The Basis for Evolution of DNA-Binding Specificity of the Aft1 Transcription Factor in Yeasts

**Isabelle R. Gonçalves, Natalia Conde e Silva, Cesar La Torre Garay, Emmanuel Lesuisse,
Jean Michel Camadro, and Pierre Louis Blaiseau**

Figure S1 Expression of the LexA-Aft proteins in *aft1Δaft2Δ* cells. 40μg of extract proteins used for the electrophoretic mobility shift assays and prepared as described in (CONDE E SILVA *et al.* 2009) were loaded on a 8% acrylamide gel. After transfert on PVDF membrane, the LexA fusion proteins were detected with anti-LexA monoclonal antibodies (Santa Cruz Biotechnology) diluted at 1:1,000.

Figure S2 KIAft but not Aft1 binds efficiently to the -531 CACACCC KIAft DNA-binding site of the KLLA0E14652g promoter. Gel-shift assays were carried out with extracts from *aft1Δaft2Δ* cells expressing LexA (C, lane 1), LexA-KIAft (lane 2) and LexA-Aft1 (lane 3). The arrow indicates the KIAft- and Aft1-containing complexes. The DNA probe corresponds to the sequence of positions -570 to -429 of KLLA0E14652g containing the -530 CACACCC KIAft DNA-binding site and was prepared as described in (Conde e Silva et al., 2009). The binding reaction mixtures (20 μ l) contained 25 mM HEPES, pH 7.6, 60 mM KCl, 7.5% glycerol, 0.1 mM EDTA, 1 mM dithiothreitol, 5 mM $MgCl_2$. Aliquots of 20 μ g of cell extract and 0,5 μ g of poly(dIdC)₂ were used. Approximately 10,000 cpm of probe (1.7 ng) was used in each binding mixture. Samples were incubated for 30 min at room temperature. Samples were loaded onto a 5% polyacrylamide gel as described in MATERIALS AND METHODS.

File S1

Inferred iron regulons in ascomycete species

Available for download as an Excel file at <http://www.genetics.org/lookup/suppl/doi:10.1534/genetics.113.157693/-/DC1>