Role of the Yap Family in the Transcriptional Response to Oxidative Stress in Yeasts

Christel Goudot¹,²,³, Frédéric Devaux⁴ and Gaëlle Lelandais¹,²,³

¹Dynamique des Structures et Interactions des Macromolécules Biologiques (DSIMB),
²Univ Paris Diderot, Sorbonne Paris Cité, Paris,
³INTS, 75015, Paris,
⁴Laboratoire de Génomique des Microorganismes, Univ Pierre et Marie Curie, Paris, France

1. Introduction

Oxidative stress can be defined as physiological changes that arise in a living organism, in reaction to an abnormal level of cytotoxic oxidants and free radicals in the environment. Because they have unpaired valence shell electrons, free radicals are very unstable and constantly seek to bond to other molecules in order to increase their stability. Within the cell, free radicals can therefore cause considerable damages in different components such as DNA (Storz et al., 1987), lipid membranes (Davies, 1985) or proteins (Smith et al., 1984). Free radicals are implicated in the ageing process (Wickens, 2001), in some autoimmune diseases (Blake et al., 1987) and in the development of cancer (Trush and Kensler, 1991). Notably, the generation of free radicals is a natural process in cell functioning and for instance, reactive oxygen species (ROS) are continuously produced as side products of aerobic metabolic pathways. To neutralize them, living systems have developed specific strategies. In normal conditions, equilibrium thus exists between the generation and the degradation of free radicals, whereas in case of oxidative stress conditions a persistent imbalance is observed.

In this chapter, we will focus on the role of the Yap proteins in the regulation of the transcriptional response to oxidative stress in yeasts. As other aerobically growing organisms, yeasts are constantly exposed to ROS molecules and have acquired sophisticated mechanisms to control modifications of its redox status due to impaired metabolism or to an excess of oxidative molecules in its environment. Yeasts thus represent an interesting model to understand how eukaryotes can cope with different levels of oxidative stress. In that respect, the species Saccharomyces cerevisiae has been extensively studied and in the literature several reviews present in details the antioxidant defence systems of this model organism (Costa and Moradas-Ferreira, 2001; Herrero et al., 2008; Ikner and Shiozaki, 2005; Jamieson, 1998; Lushchak, 2006, 2010; Moye-Rowley, 2002). Together these reviews draw one of the most complete pictures of the genomic strategy used by a cell to protect their components against ROS molecules. Yeasts possess both enzymatic and non-enzymatic defence strategies. Enzymatic system comprise enzymes including catalases, superoxide dismutases...
(SOD), glutathione reductases or glutathione peroxidases, which are able to remove partially reduced forms of molecular oxygen and/or to repair the cellular damages caused by oxidative stress. Non-enzymatic defence system involves small molecules that act as radical scavengers, i.e. they can be oxidized by ROS and thereby allow to neutralize cytotoxic oxidants. One of the best-known examples of a non-enzymatic defence system is glutathione (GSH), a tripeptide \( \gamma \)-L-glutamyl-L-cystinylglycine. In this context, to correctly produce the numerous enzymes and components required for its oxidative defence, the expression of the corresponding genes must be precisely synchronized. A major role is played by specific transcription factors (TFs), i.e. proteins that bridge sensors of cytotoxic oxidants and the transcriptional activation of particular sets of genes. In yeast \( S. ~cerevisiae \), the TFs Msn2p/Msn4p, Snk7p, Hsf1p or Yap1p are known to be involved in the response to oxidant-induced stress. The focus of this review will be on the Yap protein family that includes the protein Yap1p, i.e. the main transcriptional regulator of the genomic response to oxidative stress (Lelandais et al., 2008; Lucau-Danila et al., 2005; Rodrigues-Pousada et al., 2010). Note that other reviews provide complete descriptions of the other TF roles (Ikner and Shiozaki, 2005; Lushchak, 2010; Moye-Rowley, 2002).

2. Yap proteins and the response to oxidative stress

Yap proteins belong to the b-ZIP super family of TFs that is widely conserved from yeast to human (Rodrigues-Pousada et al., 2010). In yeast \( S. ~cerevisiae \), the role of Yap1p in oxidative stress response was initially suggested by the analysis of \( YAP1 \) mutant strains, in which increased sensitivity to cadmium and hydrogen peroxide (H\(_2\)O\(_2\)) was observed (Delaunay et al., 2000; Hirata et al., 1994; Kuge and Jones, 1994; Wu et al., 1993). Functional studies next demonstrated that Yap1p is able to activate the transcription of genes that encode key proteins in oxidative stress response such as \( TRX2 \) (thioredoxin) and \( GSH1 \) (\( \gamma \)-glutamylcysteine synthase) (Kuge and Jones, 1994), \( GSH2 \) (glutathione synthase) (Wu and Moye-Rowley, 1994), \( TRR1 \) (thioredoxin reductase) (Sugiyama et al., 2000), \( GPX2 \) (glutathione peroxidase) (Sugiyama et al., 2000), \( TSA1 \) (thioredoxin peroxidase) (Grant et al., 1996; Inoue et al., 1999) or \( AHP1 \) (alkylhydroperoxide reductase) (Lee et al., 1999a; Lee et al., 1999b) (see (Lushchak, 2010) for review). Others analyses found that Yap1p controlled expression of the ABC transporter gene \( YCF1 \) (Wemmie et al., 1994) and of the multidrug resistance (MDR) transporter genes \( FLR1 \) and \( ATR1 \) (Alarco et al., 1997; Coleman et al., 1999). The activation of the Yap1p protein in case of stress relates to a post-translational modification of the protein (Wemmie et al., 1997). In normal condition the protein Yap1p is constantly produced and exported from the nucleus to the cytoplasm, whereas in case of oxidant exposure the protein is recruited in the nucleus (Kuge et al., 1997), therefore allowing the transcriptional activation of its target genes. A key modulator of the sub cellular distribution of Yap1p is the nuclear export regulator Crm1p (Kuge et al., 1998; Yan et al., 1998) and it was demonstrated that the protein Gpx3p is involved in the activation of Yap1p in case of stress induced by H\(_2\)O\(_2\) (Delaunay et al., 2000). Notably, the response of Yap1p varies according to the nature of the oxidative stress (induced by H\(_2\)O\(_2\) or diamide for instance) (Wemmie et al., 1997) and it was proposed that Yap1p has two distinct molecular redox centers, one triggered by ROS (hydroperoxides and the superoxide anion) and the other triggered by chemicals with thiol reactivity (electrophiles and divalent heavy metal cations) (Azevedo et al., 2003). This last class of chemical Yap1p activators does not require the presence of the protein Gpx3p to initiate the oxidative stress response. In addition to
these post-translational controls, Yap1p is transcriptionally activated by strong and long-standing oxidative stresses, possibly through positive autoregulatory loop, as suggested by the observation that Yap1p binds to its own promoter (Salin et al., 2008). Analyses of promoter sequences of genes whose transcription is modulated by Yap1p allowed the identification of specific regulatory DNA motifs. In particular, different sequences were experimentally characterized with a clear preference for the TTACTAA motif (Nguyen et al., 2001). Yap1p binding sites are referred to as Yap Response Element (YRE). Recently, it was shown that the Yap1p protein covers a larger DNA fragment than strictly the TTA•TAA half sites, with a conserved adenine (and to a less extend a cytosine) located in 5' of the canonical YREs (Goudot et al., 2011).

Interestingly, functional homologous proteins of Yap1p were identified in the pathogenic yeast species Candida glabrata and Candida albicans. They are respectively named Cgap1p and Cap1p and are also involved in the response to oxidative stress (Chen et al., 2007; Lelandais et al., 2008; Znaidi et al., 2009). Infections caused by fungal pathogens have become major life-threatening diseases, especially in patients with defects in their immune system. The yeast species C. albicans and C. glabrata respectively rank as the first and the second causes of invasive infections, including systemic candidiasis and candidemia (Pfaller and Diekema, 2007). They are both human pathogens that can colonize numerous sites within their human host and are thus continuously exposed to rapid and drastic changes in their external milieu. Azole antifungals, especially fluconazole, were widely used to treat these fungal diseases, but rapidly an important number of clinical drug resistance cases was reported (White et al., 1998). Recently, it was proposed that drug resistance and the response to oxidative stress are interconnected processes in C. albicans (Znaidi et al., 2009). This connection exists through the activity of several TFs like Cap1p, Tac1p, Mrr1p or Upc2p for which (i) gain-of-function mutations were observed as being related to clinical azole resistance and (ii) genome-wide studies identified several of their target genes as being involved in the general response to oxidative stress (Liu et al., 2007; Morschhauser et al., 2007; Znaidi et al., 2009; Znaidi et al., 2008). Also, it is interesting to remind that, in S. cerevisiae, YAP1 was initially described as a gene involved in pleiotropic drug resistance (Wu et al., 1993). Moreover, as a defence strategy against fungal infections for the animal host consists in using oxidative killing carried out by macrophages, an essential feature for the virulence of pathogenic yeasts relates to their ability to tolerate ROS molecules (Abegg et al., 2010). In this context, understanding the role of Yap TFs in the genomic response to oxidant-induced stress in fungi provide valuable information to better understand fungal pathogenesis.

Our text is organized in the following manner. We will first present genome-wide strategies to decipher the genomic response related to the activity of Yap1p, Cgap1p and Cap1p TFs in case of oxidative stress. Data coming from high-throughput experimental technologies will be presented. Next, three different yeast species (the pathogen species C. glabrata and C. albicans and the model yeast S. cerevisiae) will be compared. As a large amount of genomic information is available for S. cerevisiae (genomic sequence, transcriptome data, gene functional annotations, TF DNA binding sequences, etc.), this species will serve as a reference to describe the genomic oxidative stress response in Candida species. Finally, we will discuss oxidative stress response in the light of the other Yap TFs. In S. cerevisiae, this family comprises eight members (Yap1p to Yap8p) that are TFs carrying both overlapping
and distinct biological functions (Fernandes et al., 1997; Rodrigues-Pousada et al., 2010). In yeasts *C. glabrata* and *C. albicans*, we will present a description of the Yap families, comprising respectively seven and four members. We will discuss how modifications in the number of paralogous TFs that belong to the Yap family in each species imply similarities and differences in the regulatory control of their target genes (genes being transferred from one factor to another).

3. Integration of multiple data sources for the characterisation of AP-1 transcriptional modules

3.1 General principle

In a recent study, we proposed a strategy to identify the sets of genes for which transcription was activated by the regulatory proteins Yap1p (in *S. cerevisiae*), Cgap1p (in *C. glabrata*) and Cap1p (in *C. albicans*), in response to a specific stress induced by benomyl (Goudot et al., 2011). These sets of genes are referred to as AP-1 transcriptional modules (TMs). In each yeast species, benomyl is known to activate an oxidative stress, for which the genomic response of the cell is primarily dependent on the TFs Yap1p (Lucau-Danila et al., 2005), Cgap1p (Lelandais et al., 2008) and Cap1p (Znaidi et al., 2009). To summarize, our strategy consisted in considering three different layers of information based on the analysis of genome-wide datasets. First we used expression patterns of genes to identify those that are significantly upregulated in response to benomyl induced-stress (see below Step 1). Second we analyzed the transcriptome alterations in yeast strains deleted for the genes coding the yeast AP-1 TFs (respectively Yap1p, Cgap1p and Cap1p) (see Step 2), and third we searched for genomic locations of the TF binding DNA sequences using ChIP-chip experiments (see Step 3). A global overview of the computational framework is presented Figure 1. Note that each experimental dataset was carefully chosen in order to ensure both intra and inter-species comparisons of the obtained results, *i.e.* comparable benomyl concentrations in each species and similar time point measurements for transcriptome analyses.

3.2 Step 1: Genome-wide expression data to measure the transcriptional response induced by benomyl

In a previous study (Lelandais et al., 2008), we carried out microarray analyses of the transcriptome responses of yeasts *S. cerevisiae* and *C. glabrata*, following identical treatments with the antifungal agent benomyl (Gupta et al., 2004) (20 μg/ml benomyl concentration and time measurements at 2, 4, 10, 20 and 40 minutes). More recently, Znaidi et al. (2009) performed similar microarray experiments in *C. albicans*, analyzing the transcriptome modifications after 30 minutes of benomyl treatment with a 30 μg/ml benomyl concentration. As each dataset was originally examined using different bioinformatics methodologies, we collected the initial raw data from the GEO database (Barrett et al., 2009) and applied in each species the same procedure for identifying genes whose transcription was significantly modified after benomyl addition. Besides data pre-processing, we used a combination of three different algorithms: SAM (Tusher et al., 2001), LIMMA (Wettenhall and Smyth, 2004) and SMVar (Jaffrezic et al., 2007). These algorithms were chosen because they are representative of different variance modelling strategies in gene expression data (Jeanmougin et al., 2010). In total, 786 genes were identified as being significantly up regulated in *S. cerevisiae*, 327 genes in *C. glabrata* and 337 genes in *C. albicans* (Figure 1, step 1).
3.3 Step 2: Mutant analyses to measure the deletion impact of genes coding for yeast AP-1 transcription factors

Previous studies shown that even if Yap1p (in *S. cerevisiae*), Cgap1p (in *C. glabrata*) and Cap1p (in *C. albicans*) are important coordinators of the early transcriptional response to benomyl stress (Lelandais et al., 2008; Znaidi et al., 2009), other TFs are also involved in the system (Lelandais et al., 2008). The lists of genes identified in Step 1 therefore comprised in one hand, target genes for Yap1p, Cgap1p and Cap1p TFs, but also on the other hand, target genes for additional regulators (for instance Msn2p/Msn4p (Lelandais et al., 2008)). To specifically highlight the yeast AP-1 responsive genes, we analyzed in each species transcriptome data comparing gene expression between wild-type strains and strains deleted for genes coding yeast AP-1 TFs: ΔYAP1 (in *S. cerevisiae*), ΔCgAP1 (in *C. glabrata*) and ΔCAP1 (in *C. albicans*). The yeast AP-1 TF target genes are genes whose expression is significantly altered in the deleted strains. Mutant datasets were therefore collected from the studies of Lucau-Danila et al. (2005) (ΔYAP1 analyses with a 20 μg/ml benomyl concentration), Lelandais et al. (2008) (ΔCgAP1 analyses with a 20 μg/ml benomyl concentration) and Znaidi et al. (2009) (ΔCAP1 analyses with a 30 μg/ml benomyl concentration). Besides data pre-processing, we used the combination of the three algorithms SAM, LIMMA and SMVar (Goudot et al., 2011). As a result, 33 genes were identified as Yap1p-dependent genes in *S. cerevisiae*, 134 genes as Cgap1p-dependent genes in *C. glabrata* and 168 genes as Cap1p-dependent genes in *C. albicans* (Figure 1, step 2).

3.4 Step 3: ChIP-chip experiments to identify yeast AP-1 transcription factor binding sites *in vivo*

As the data sources analyzed in Step 1 and 2 provided only indirect information concerning the transcriptional regulation between yeast AP-1 TFs and their target genes, the goal for this additional step was to discern dependencies between the gene expression patterns observed in Step 1 and 2 and the physical interactions revealed by genome-wide location profiling experiments. We thus analyzed ChIP-chip experiments performed for TFs Yap1p, Cgap1p and Cap1p. Note that experiments in benomyl condition were available only in *S. cerevisiae* for Yap1p (20 μg/ml benomyl concentration) (Salin et al., 2008). Other experiments in *C. glabrata* and *C. albicans* were performed in non-stress induced conditions (respectively (Kuo et al., 2010) and (Znaidi et al., 2009)). Besides data pre-processing, we identified promoter sequences that were significantly bond by the yeast AP-1 TFs by a combination of the results obtained with the SAM, LIMMA and SMVar methods, and those obtained using the ChIPmix algorithm (Martin-Magniette et al., 2008). Unlike SAM, LIMMA and SMVar methods that work on log ratio, ChIPmix has the originality to directly analyze the signals of IP (DNA fragments cross-linked to TF protein) and INPUT (genomic DNA) by modelling the distribution of the IP signal conditional to the INPUT signal (Martin-Magniette et al., 2008). Therefore, 260 genes were found to be bond by Yap1p, 327 genes by Cgap1p and finally 373 genes by Cap1p (Figure 1, step 3).

3.5 Final step: Data integration

Results obtained in Step 1, 2 and 3 were finally combined as described in Figure 1. Genes selected in “Step 1 and Step 2”, or in “Step 1 and Step 3” were conserved in the final AP-1 TMs. In *S. cerevisiae* the Yap1p TM therefore comprised 67 genes, in *C. glabrata* the Cgap1p
TM comprised 98 genes, and finally in \textit{C. albicans} the Cap1p TM comprised 130 genes. Complete list of genes in each TM together with their corresponding functional annotations can be found in (Goudot et al., 2011). Note that the number of genes in each AP-1 TM is underestimated compared to information available in database like YEASTRACT (Abdulrehman et al., 2010) (in which more than 400 genes are annotated as potential targets for Yap1p). However, one must keep in mind that the TMs described here have the particularity (i) to be focused on the AP-1 responsive genes in benomyl stress-induced conditions (genes regulated by AP-1 TFs in other conditions are not considered), and (ii) to include only genes for which at least two types of experimental evidences were available for interactions with Yap1p, Cgap1p or Cap1p. All together, we integrated experimental results arising from more than 80 individual microarray experiments applying four different bioinformatics methodologies. The predictive strength of the strategy is based on the combined constraints that arise from the use of multiple biological and bioinformatics data sources.

![Fig. 1. General strategy to characterize the yeast AP-1 transcriptional modules (TMs).](image)

Characterizing a TM consists in identifying all genes whose transcription is modulated by a particular transcription factor (TF). Applied to the analysis of AP-1 TMs in yeasts during benomyl induced-stress, the general strategy shown here proceeds in several successive steps, using datasets which focus on different levels of transcriptional regulation: (Step 1) identification of differentially expressed genes during benomyl stress, (Step 2) identification of genes whose benomyl induction is affected by the deletion of the gene encoding the AP-1 TFs, and (Step 3) genome-wide location of the AP-1 TFs, as determined by ChIP-chip analyses (see the main text). Genes selected in “Step 1 and Step 2”, or in “Step 1 and Step 3” were conserved in the final AP-1 TMs. In \textit{S. cerevisiae} the Yap1 TM therefore comprised 67 genes, in \textit{C. glabrata} the Cgap1p TM comprised 98 genes, and finally in \textit{C. albicans} the Cap1p TM comprised 130 genes.

4. Comparative analysis of yeast AP-1 transcriptional modules based on sequence orthology

One of the main objective for genomics studies is to transfer functional annotations from well-studied organisms to the newly sequence species. For that, the most widely used
Role of the Yap Family in the Transcriptional Response to Oxidative Stress in Yeasts

approach for function prediction is based on orthology assignments. Orthology defines the relationship between genes in different species that originate from a single gene in the last ancestor of these species (Fitch, 2000; Sonnhammer and Koonin, 2002). Therefore, orthologous genes are most likely to share the same function and should, in principle, have conserved their expression patterns together with their regulatory control by TFs. To test this hypothesis, the three AP-1 TMs defined separately in S. cerevisiae, C. glabrata, and C. albicans represent valuable information. Indeed, as they were characterized using in each yeast species independent experimental datasets (see previous Section), they represent an interesting reference model. We applied the INPARANOID algorithm (O’Brien et al., 2005) comparing all protein sequences of the three yeast species (Goudot et al., 2011). Orthologous links were thus inferred for 80% of genes between S. cerevisiae and C. glabrata, 61% of genes between S. cerevisiae and C. albicans and 63% of genes between C. glabrata and C. albicans. These results were coherent with the phylogeny of the yeast species analyzed here, i.e. C. glabrata being more closely related to S. cerevisiae than C. albicans is. Then, we determined whether orthologous genes were present in each of the three AP-1 TMs. Strikingly we could only distinguish 11 orthologous genes between the S. cerevisiae and C. glabrata AP-1 TMs (16%), 7 between the S. cerevisiae and C. albicans AP-1 TMs (10%) and 14 between the C. glabrata and the C. albicans AP-1 TMs (14%) (Goudot et al., 2011). Considering the global amount of orthologous genes between the three species (more than 60%), these values were surprisingly low and demonstrated that in yeasts, the functioning of Yap1p, Cgap1p and Cap1p TFs during the transcriptional response to benomyl stress has been significantly rewired. Gene duplication and multigenic protein families are parameters that can explain changes in the yeast AP-1 TMs. For instance, we observed that OYE paralogous genes, which encode NADPH oxydoreductases involved in sterol metabolism and oxidative stress response, were present in AP-1 TMs in each species, but with a different number of copies. In S. cerevisiae, only two OYE genes (OYE2 and its paralogue OYE3) belong to the Yap1p TM, whereas 3 and 4 OYE paralogues were identified respectively in Cgap1p and Cap1p TMs (in C. glabrata and C. albicans) (see (Goudot et al., 2011) for a detailed description of these genes). In the three yeasts, the general function mediated by the OYE genes is therefore conserved, but because of several duplication events, direct orthologous relationships between genes were lost. Additionally, AP-1 proteins belong to Yap families that comprised 3 to 8 paralogous genes in the Hemiascomycetes. In the model yeast S. cerevisiae, the Yap family is composed of eight proteins (Yap1p to Yap8p) that are TFs carrying both overlapping and distinct biological functions (Rodrigues-Pousada et al., 2010), and recognize similar DNA consensus (Tan et al., 2008). Structural features of Yap proteins in S. cerevisiae are presented in Figure 2. Considering that these factors can interact functionally, they certainly cross-influenced the evolution of their respective TM. To clarify this point, we searched for potential roles in the response to oxidative stress for other members of the Yap families, using genomic sequence analyses and functional information stored in the databases.

5. Characterisation of the Yap families in Candida species

In contrast to S. cerevisiae, in which an important number of functional information concerning Yap proteins is available (see (Rodrigues-Pousada et al., 2010) for review), very few studies have yet been carried out in Candida species to analyze the redundant Yap families (Chen et al., 2007; Singh et al., 2011). Most information is currently based on sequence similarity with Yap genes of S. cerevisiae. We searched for all potential Yap
proteins in complete genomes of *C. glabrata* and *C. albicans* yeast species, applying a methodology derived from the strategy used in *S. cerevisiae* to identify the eight Yap family members (Fernandes et al., 1997). Starting from the protein sequences of Cgap1p and Cap1p in *Candida* species (referred to as “query sequences”), our approach consists in comparing these AP-1 sequences with all other protein sequences available in *Candida* genomes (referred to as “target sequences”). For each comparison between query and target sequences, a score is calculated (see Equation 1). This score (between 0 and 1) as the advantage to take into account (i) the amino acid similarity observed after a global alignment between the query and the target sequences, (ii) the amino acid similarity observed after a local alignment and (iii) the amino acid similarity observed by restricting the alignment to the DNA binding region of the queried protein sequences (Cgap1p or Cap1p). The higher the score, the more the probability is for the target sequence to belong to the Yap families in one of the *Candida* species. Note that we could observe that a score value higher than 0.4 represents a “highly confident” homology relationship, a score value between 0.3 and 0.4 a “confident” homology relationship and a score below 0.3 a “poorly confident” homology relationship. Global and local alignments were performed using respectively the Needleman-Wunsch and the Smith-Waterman algorithms implemented in EMBOSS (Lamprecht et al., 2011). After careful manual inspection of the obtained results, we identified 7 members in the *C. glabrata* Yap family and 4 members in *C. albicans* (Table 1). Note that this lower number of Yap proteins in *C. albicans* is connected to the whole genome duplication that arose in the common history of *S. cerevisiae* and *C. glabrata*, but not in the *C. albicans* ancestor.
<table>
<thead>
<tr>
<th>Systematic name (C. glabrata)</th>
<th>Standard name</th>
<th>Description (Génolevures database)</th>
</tr>
</thead>
<tbody>
<tr>
<td>CAGL0H04631g</td>
<td>CgAP1</td>
<td>Similar to uniprot</td>
</tr>
<tr>
<td>CAGL0F03069g</td>
<td>-</td>
<td>Some similarities with uniprot</td>
</tr>
<tr>
<td>CAGL0K02585g</td>
<td>-</td>
<td>Some similarities with uniprot</td>
</tr>
<tr>
<td>CAGL0M10087g</td>
<td>-</td>
<td>Some similarities with uniprot</td>
</tr>
<tr>
<td>CAGL0M08800g</td>
<td>-</td>
<td>Weakly similar to uniprot</td>
</tr>
<tr>
<td>CAGL0K08756g</td>
<td>-</td>
<td>Weakly similar to uniprot</td>
</tr>
<tr>
<td>CAGL0M08800g</td>
<td>-</td>
<td>Weakly similar to uniprot</td>
</tr>
<tr>
<td>CAGL0F01265g</td>
<td>-</td>
<td>Some similarities with uniprot</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Systematic name (C. albicans)</th>
<th>Standard name</th>
<th>Description (CGD database)</th>
</tr>
</thead>
<tbody>
<tr>
<td>orf19.1623</td>
<td>CAP1</td>
<td>Transcription factor, AP-1 bZIP family; role in oxidative stress response and resistance, multidrug resistance; oxidative stress regulates nuclear localization; partially complements S. cerevisiae yap1 mutation; human neutrophil-ind</td>
</tr>
<tr>
<td>orf19.3193</td>
<td>FCR3</td>
<td>Transcriptional regulator of the bZIP family; partially functionally complements the fluconazole sensitivity of an S. cerevisiae pdr1 pdr3 double mutant; probable ortholog of S. cerevisiae Yap3p; Hap43p-induced</td>
</tr>
<tr>
<td>orf19.681</td>
<td>HAP43</td>
<td>CCAAT-binding factor-dependent transcriptional repressor required for low iron response; similar to bZIP transcription factor AP-1; HAP4L-bZIP bipartite domain; gene negatively regulated by Sfu1p; ciclopirox olamine induced</td>
</tr>
<tr>
<td>orf19.861</td>
<td>CAP4</td>
<td>Predicted transcriptional regulator with bZip domain; possibly an essential gene, disruptants not obtained by UAU1 method</td>
</tr>
</tbody>
</table>

Table 1. **Characterisation of the Yap families in Candida species.** Systematic protein sequence comparisons were performed to identify Yap family respectively in C. glabrata and C. albicans yeast species (see the main text). After manual inspection of the obtained results, 7 genes were identified as members of the Yap family members in C. glabrata genome and 4 in the C. albicans genome. Systematic and standard gene names are presented here, with the associated descriptions found in the Génolevures http://www.genolevures.org/ (for C. glabrata) or Candida Genome Database (CGD) http://www.candidagenome.org/ (for C. albicans).
Equation 1. Score used to identify Yap family members in Candida genomes. Sq and St respectively means “Query sequence” and “Target sequence”. \( NW_{\text{global}} \) corresponds to the score value returned by the Needleman-Wunsch algorithm after a global alignment of the query and target sequences. \( SW_{\text{local}} \) corresponds to the score value returned by the Smith-Waterman algorithm after a local alignment of the query and target sequences. Finally, \( SW_{\text{DNAbinding}} \) corresponds to the score value returned by the Smith-Waterman algorithm restricting the alignment to the DNA binding region.

In a second step, we performed cross-species comparisons of the Yap sequences identified in Candida genomes with those described in the model yeast \( S. \ ceriseiae \), i.e. proteins Yap1p to Yap8p. The highest values of the comparative score based on protein sequence alignments (see previous section) are presented in Table 2. As expected, we could observe important score values between Yap1p (in \( S. \ ceriseiae \)) and Cgap1p (in \( C. \ glabrata \)) (score value = 0.48),

<table>
<thead>
<tr>
<th>Yap family in ( S. \ ceriseiae )</th>
<th>Yap family in ( C. \ glabrata )</th>
<th>Score value (between protein sequences)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yap1p</td>
<td>Cgap1p</td>
<td>0.4767</td>
</tr>
<tr>
<td>Yap2p</td>
<td>CAGL0F03069g</td>
<td>0.4298</td>
</tr>
<tr>
<td>Yap3p</td>
<td>CAGL0K02585g</td>
<td>0.4788</td>
</tr>
<tr>
<td></td>
<td>CAGL0M10087g</td>
<td>0.4309</td>
</tr>
<tr>
<td>Yap4p</td>
<td>CAGL0M08800g</td>
<td>0.3408</td>
</tr>
<tr>
<td>Yap5p</td>
<td>CAGL0K08756g</td>
<td>0.3976</td>
</tr>
<tr>
<td>Yap6p</td>
<td>CAGL0M08800g</td>
<td>0.3352</td>
</tr>
<tr>
<td>Yap7p</td>
<td>CAGL0F01265g</td>
<td>0.3536</td>
</tr>
<tr>
<td>Yap8p</td>
<td>CAGL0H04631g</td>
<td>0.2558</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Yap family in ( S. \ ceriseiae )</th>
<th>Yap family in ( C. \ albicans )</th>
<th>Score value (between protein sequences)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yap1p</td>
<td>Cap1p</td>
<td>0.4158</td>
</tr>
<tr>
<td>Yap2p</td>
<td></td>
<td>0.3713</td>
</tr>
<tr>
<td>Yap3p</td>
<td></td>
<td>0.3957</td>
</tr>
<tr>
<td>Yap4p</td>
<td>Fcr3p</td>
<td>0.2499</td>
</tr>
<tr>
<td>Yap5p</td>
<td></td>
<td>0.2757</td>
</tr>
<tr>
<td>Yap6p</td>
<td></td>
<td>0.2317</td>
</tr>
<tr>
<td>Yap7p</td>
<td>Hap43p</td>
<td>0.2970</td>
</tr>
<tr>
<td>Yap8p</td>
<td></td>
<td>0.2680</td>
</tr>
<tr>
<td>Cap4p</td>
<td></td>
<td>&lt; 0.25</td>
</tr>
</tbody>
</table>

Table 2. Cross-species comparisons of the Yap families in Candida species with the Yap family in \( S. \ ceriseiae \). Protein sequences of each member of the Yap families described in Candida species were compared with the protein sequences of each member of the Yap family in \( S. \ ceriseiae \). Highest scores values are shown here and allow to assign homologous relationships between proteins. A score value higher than 0.4 represents a “highly confident” homology relationship, a score value between 0.3 and 0.4 a “confident” homology relationship and a score below 0.3 a “poorly confident” homology relationship.
and Yap1p and Cap1p (in *C. albicans*) (score value = 0.41). For Yap2p protein, we found a homologous protein in *C. glabrata* named CAGL0F03069g (score value = 0.43). This result is in agreement with the annotation stored in the Génolevures database (Sherman et al., 2006). In *C. albicans*, the best homologous sequence was Cap1p (score value = 0.37). Interestingly in *S. cerevisiae* Yap2p protein is the Yap member that shares highest sequence similarity and functional redundancy with Yap1p (Figure 2). Considering that the whole genome duplication arose only in *S. cerevisiae* and *C. glabrata* genomes, the protein Cap1p certainly retains functional properties in *C. albicans* that are split between Yap1p and Yap2p in *S. cerevisiae*, and Cgap1p and CAGL0F03069g in *C. glabrata*. In the case of the protein Yap3p, we could identify two proteins with important scores in *C. glabrata*: CAGL0K02585g (score value = 0.48) and CAGL0M10087g (score value = 0.43). On the other hand Yap4p and Yap6p proteins in *S. cerevisiae* appeared to have a unique homologous protein CAGL0M08800g in *C. glabrata* (scores values are respectively 0.34 and 0.33). Finally, Yap5p and Yap7p proteins exhibited unique homologous proteins in *C. glabrata*, which are respectively CAGL0K08756g (score value = 0.39) and CAGL0F01265g (score value = 0.35). In *C. albicans*, the protein Fcr3p (orf19.3193) shared an important score value with Yap3p (score value = 0.39) and lower values with Yap4p (score value = 0.25), Yap5p (score value = 0.27) and Yap6p (score value = 0.24). Yap7p and Yap8p proteins also had a single homologous protein named Hap43p (orf19.681, score values = 0.30 and 0.27). Finally, Yap8p protein in *S. cerevisiae* appeared to have the lowest score values with the identified Yap proteins in both *Candida* species (lower than 0.3) meaning that the homology relationships are “poorly confident”. Such an observation suggest an important evolution of Yap8p protein in yeast species with a potential functional transfer of the Yap8p regulated genes to other Yap factors in *Candida* species. No confident homologous protein for Cap4p (orf19.861, in *C. albicans*) could also be identified in *S. cerevisiae*. Again, this indicates that in yeast species, significant rewiring occurred in the Yap families during evolution.

6. Potential roles in oxidative stress response for others members of the Yap family

Yeast species possess very flexible and complex gene expression programs to respond quickly and efficiently to several environmental challenges. The complexity of the underlying transcriptional networks is only become to light. Since the pioneer study of Harbison et al. (2004), an important challenge consists in understanding combinatorial regulations (Bhardwaj et al., 2010a; Bhardwaj et al., 2010b), *i.e.* how TFs partner with each other to control the expression of common target genes. This allows a limited number of regulators to govern many genes and obtain more plastic and specific functions in the cell. The observation that multiple homologous links exist between the Yap proteins identified in the three different species (see Table 2), together with the observation that only 10% of genes share orthologous relationships between AP-1 TMs (see Section 3) raise important questions concerning the inter- and intra-species functional redundancies that exist between Yap proteins. In the model yeast *S. cerevisiae*, several functional studies were performed in order to characterize the role of the different Yap family members (see (Rodrigues-Pousada et al., 2010) for review). It has been shown that Yap2p (also names Cad1p) plays a role in the response to toxic compounds such as cadmium (Azevedo et al., 2007). Yap4p and Yap6p were identified as being involved in the osmotic stress response (Posas et al., 2000). Yap5p has been shown to play a role in the metabolism and storage of iron and Yap8p, the last Yap
family member, has a central role in the detoxification of the cell and more specifically in arsenic stress response (Ilina et al., 2008). Only the roles of TFs Yap3p and Yap7p remain to be established.

In order to consider the potential roles of the different Yap family members in the specific response to oxidative stress, we inferred a transcriptional regulatory network based on information stored in databases. Our approach consisted, in a first step, to select all the genes in *S. cerevisiae*, annotated in the Gene Ontology database (2010) as being involved in oxidative stress (functional category GO:0006979). 88 genes were thus collected and were searched in a second step, for potential regulatory interactions with Yap TFs. For that, we used information deposited in the YEASTRACT database (Abdulrehman et al., 2011). YEASTRACT was developed to support the analysis of transcription regulatory associations in *S. cerevisiae* and stores regulatory associations described in the literature. Among the 88 genes selected in the first step, more than 50% (41 genes) were annotated as target genes for one of the Yap TFs (only using YEASTRACT direct evidences). Interactions between Yap TFs and their target genes are represented Figure 3A. As expected, the transcription factor Yap1p appears to have the predominant role in the transcriptional control of genes associated with oxidative stress. Interactions with 24 analyzed genes (27%) were identified. Interestingly our results reveal that Yap4p, Yap7p and Yap2p potentially exert significant transcriptional controls on a significant number of oxidative responsive genes. They are the regulators of respectively 14 genes (16%), 13 genes (15%) and 11 genes (12.5%). Yap5p, Yap6p and Yap3p present only a small number of target genes in the oxidative stress response GO category (respectively 8, 6 and 2 genes). Therefore, the Yap transcriptional network involved in response to oxidative stress appeared to be larger than this generally stated in the literature (Moye-Rowley, 2002). Despite its incontestable and crucial role, Yap1p is certainly not the unique Yap regulator necessary to obtain an optimal oxidative response. It belongs to a redundant network, with a majority of genes being potentially regulated by several Yap transcription factors (only 17 genes are target genes of a single Yap protein).

Our observations suggest a potential important role for Yap2p, Yap4p and Yap7p proteins in the transcriptional control of oxidative stress response. Interestingly, Yap2p and Yap4p share common target genes with Yap1p. These genes are known to play a key role in the response of oxidative stress. For example, the gene TRX2 is a highly conserved oxydoreductase, required to maintain the redox homeostasis of the cell. In *S. cerevisiae*, thioredoxin TRX2 and thioredoxin reductase TRR1 were shown to be necessary to protect the cell against oxidative stress induced by the presence of ROS compounds in condition H$_2$O$_2$ (Boisnard et al., 2009) or exposure to dithiothreitol (DTT) (Garrido and Grant, 2002). This agrees with the review of Rodrigues-Pousada *et al.* (2010), in which Yap2p and Yap4p were described as involved in several types of stress responses, including oxidative stress (Cohen et al., 2002; Nevitt et al., 2004). Yap5p and Yap6p proteins also seem to be involved in the oxidative stress response, but in a more general context. Indeed, whereas Yap1p, Yap2p, Yap4p (and probably Yap7p) are involved in the core response to oxidative stress by activating target genes (catalases, reductases, etc.) specialized in response to the presence ROS compounds in the cell or its environment, Yap5p and Yap6p rather activate the transcription of genes involved in cell protection against damages caused by oxidative stress (for instance *GRX4*, *GRX5*, *CTT1*), in the biosynthetic pathway of ergosterol (for instance *MCR1*), in the regulation of apoptosis (for instance *MCA1*) or in the maintain of membrane...
Role of the Yap Family in the Transcriptional Response to Oxidative Stress in Yeasts

Fig. 3. Potential transcriptional networks involved in yeast oxidative stress response. (A) Using information stored in the GO (2010) and YEASTRACT (Abdulrehman et al., 2011) databases, we inferred a Yap transcriptional network in yeast S. cerevisiae (see the main text). Regulatory interactions are color-coded according to their dependant YAP TF: purple for Yap1p, pink for Yap2p, red for Yap3p, beige for Yap4p, brown for Yap5p, yellow for Yap6p, orange for Yap7p. No interaction with Yap8p TF was identified. (B) Representation of the transcriptional networks obtained in Candida species using orthologous relationships with S. cerevisiae genes and regulatory interactions described in (Goudot et al., 2011). Genes for which no orthologue was found are crossed out with red crosses. These networks are incomplete compared to this obtained in S. cerevisiae and highlight the need for more experimental data in C. glabrata and C. albicans species.

organization (for instance HSP12). Interestingly, potential roles of Yap7p in response to stress in general, has not been defined (Rodrigues-Pousada et al., 2010). Even if additional
experimental validations are required to confirm the role of Yap7p in the stress response in *S. cerevisiae*, this review gives interesting information and represents a good starting point.

To go further, we tried to extend our representation of the Yap transcriptional network to *C. glabrata* and *C. albicans* species. In a first step, we searched in each species for a list of genes involved in oxidative stress response. GO annotations were unfortunately incomplete and compelled us to restrict our search to orthologous *S. cerevisiae* genes. Only 32 genes were thus selected in *C. glabrata* genome and 31 in *C. albicans* genome. They are represented in Figure 3B. Concerning the regulatory interactions, literature information was just available for Cgap1p (in *S. cerevisiae*) and Cap1p (in *C. albicans*) TFs (Goudot et al., 2011). Therefore only 7 and 5 interactions were identified in *C. glabrata* and *C. albicans* respectively. It is clear that the obtained networks in *Candida* species are largely incomplete compared to this of *S. cerevisiae*. The classical approach that consists in transferring functional annotations from well-studied organisms (like *S. cerevisiae*) to newly sequence species (like *Candida* species) cannot be properly used in the case of a regulatory network that involved highly redundant TFs. In a recent study, we evaluated that such a strategy would have led (in case of AP-1 TMs, see Section 2) to a rate of false positive and false negatives predictions higher than 70% (Goudot et al., 2011). Fortunately, the recent development of costless and efficient multispecies transcriptomic platforms should lead to a rapid accumulation of new experimental datasets obtained directly in *Candida* species. It will allow to reproduced the analysis performed with AP-1 TFs to other factors and hence will greatly enhance our comprehension of the role of the Yap family in the response to oxidative stress in yeasts.

7. Conclusion

Comparative functional analyses have been made possible by the accumulation of large-scale gene expression datasets for a large number of organisms, due directly to the exponential increase in the number of species for which whole genome sequences are available (Liolios et al., 2010). To increase the accuracy of investigations in the evolution of transcriptional networks, we would like in an ideal case, compare species with different lifestyle and physiological properties. In this respect, yeasts are ideal organisms for comparative functional genomic studies, since they have evolved in niches with constantly varying nutrient availability and growth conditions. Also, the sequencing of genomes of a dozen of different species open new possibility for studying the adaptation of transcriptional networks to environmental stresses (see (Lelandais and Devaux, 2010; Lelandais et al., 2011) for reviews). Analysis of the genomic events that underlie the response to oxidative stress demonstrated that although the gene expression patterns characterizing the response are relatively well conserved between yeast species (Lelandais et al., 2008; Znaidi et al., 2009), part of the underlying regulatory networks differed. In particular, the roles of the oxidative stress response TFs Yap1p (in *S. cerevisiae*), Cgap1p (in *C. glabrata*) and Cap1p (in *C. albicans*) appeared to have diverged (Goudot et al., 2011; Lelandais et al., 2008). In *C. glabrata*, the preferred DNA binding sites of Cgap1p protein is different from the *S. cerevisiae* and *C. albicans* canonical YRE sequence (Goudot et al., 2011; Kuo et al., 2010). The functioning of Yap1p TM is therefore more similar to the functioning of Cap1p TM than to the Cgap1p TM. Considering that *C. glabrata* is phylogenetically more
related to *S. cerevisiae* than *C. albicans* is (Dujon, 2010), this observation illustrates that evolution of regulatory modules may be different from the phylogeny established from comparative genomics. Our review of Yap families suggests that the associated TMs have been importantly shuffled during evolution. Constrains exists to maintain key cellular functions as oxidative defence, but the genomic strategies to obtain these function evolve more rapidly than originally expected. Deciphering the regulatory interactions, as well as the dynamics of these processes, represent important challenges to understand the genomic control of stress response, a highly conserved process in eukaryotes.

8. References


Role of the Yap Family in the Transcriptional Response to Oxidative Stress in Yeasts


Since the discovery of free radicals in biological systems researchers have been highly interested in their interaction with biological molecules. Denoted in 1980, and due to fruitful results and ideas, oxidative stress is now appreciated by both basic and applied scientists as an enhanced steady state level of reactive oxygen species with wide range of biological effects. This book covers a wide range of aspects and issues related to the field of oxidative stress. The association between generation and elimination of reactive species and effects of oxidative stress are also addressed, as well as summaries of recent works on the signaling role of reactive species in eukaryotic organisms. The readers will gain an overview of our current understanding of homeostasis of reactive species and cellular processes they are involved in, as well as useful resources for further reading.

How to reference

In order to correctly reference this scholarly work, feel free to copy and paste the following: