Transcriptional control of the quorum sensing response in yeast†

Arthur Wuster* and M. Madan Babu*

Received 7th July 2009, Accepted 10th August 2009
First published as an Advance Article on the web 1st September 2009
DOI: 10.1039/b913579k

Quorum sensing is a process of intercellular communication. It allows individual cells to assess population density and to co-ordinate behaviour by secreting and sensing communication molecules. In the yeast *Saccharomyces cerevisiae*, the communication molecules are the aromatic alcohols tryptophol and phenylethanol, and quorum sensing regulates the transition between the solitary yeast form and the filamentous form. Though it is known that addition of these communication molecules to yeast cultures causes large changes in gene expression, how these changes are orchestrated and whether this system is conserved in related fungal species is still unknown. In this work, by employing an integrated computational approach that makes use of large-scale genomics datasets, such as ChIP-ChIP and expression analysis upon deletion and over-expression of transcriptional factors, we predict CAT8 and MIG1 as key transcriptional regulators that control the differential expression of the genes affected by aromatic alcohol communication. In addition, through a comparative genomic analysis involving 31 fungal species, we show that the *S. cerevisiae* quorum sensing system is a recent evolutionary innovation and that the genes which are differentially expressed upon treatment with these molecules are distributed across the genome in a highly non-random manner. The identified transcription factors will aid in further unravelling the molecular mechanisms of *S. cerevisiae* quorum sensing and may facilitate the engineering of regulatory circuits for applications such as the expression of heterologous proteins via aromatic alcohols.

Introduction

In order to assess population density, many unicellular organisms secrete specific small molecules into their environment. Assuming that all cells of a growing population secrete these communication molecules, their concentration rises with increasing cell density. By sensing the concentration of the communication molecules, individual cells can gauge local cell density and change their behaviour accordingly. This process is referred to as quorum sensing.1–3 Different quorum sensing systems in bacteria and unicellular eukaryotes vary in details such as the chemical identity of the communication molecules or the behaviour regulated by the process. However, all quorum sensing systems have in common that they need a molecular machinery to synthesise, sense, and respond to the secreted communication molecules.

Apart from bacteria, quorum sensing has also been reported in fungi. Although in fungi the chemical identity of the communication molecules varies, in all reported cases the major phenotype reported is the transition between the filamentous growth form and the solitary yeast form.4–7 In the model organism *Saccharomyces cerevisiae*, the aromatic alcohols phenylethanol and tryptophol, derived from the amino acids phenylalanine and tryptophan, serve as communication molecules in low nitrogen conditions.7–8 Although the molecular machinery for synthesising phenylethanol and tryptophol is known, it is not yet clear how *S. cerevisiae* senses and transduces the signal. What we do know is that phenylethanol and tryptophol synergistically affect the up-regulation of *FLO11* via the cAMP-dependent PKA subunit Tpk2p and the transcription factor Flo8p7 (Fig. 1). Flo11p, the product of *FLO11*, is a glycosylphosphatidylinositol (GPI) anchored cell surface flocculin protein and is essential for filamentous growth.9–11 *S. cerevisiae* strains with deletions of either *TPK2* or *FLO8* do not form filaments in response to the aromatic alcohols.7

The production of phenylethanol and tryptophol is cell-density dependent. *AR09* and *AR010* are two genes required for their synthesis. The expression of these aromatic aminotransferases is induced by tryptophol via the Aro80p transcription factor. Because tryptophol induces the enzymes required for its own synthesis, this results in a positive feedback loop. Therefore, cells at high population density produce more aromatic alcohols per cell than cells at low population density. The tryptophol and phenylethanol synthesis pathways also overlap with the nitrogen sensing pathway. Ammonia represses filamentous growth as well as the expression of the above mentioned *AR09* and *AR010* genes. However, elements of the mitogen activated protein kinase-protein kinase A (MAPK-PKA) pathway, which is required in *S. cerevisiae* to transmit information about nitrogen shortage, do not seem to be affected by aromatic alcohol communication molecules.7

In *S. cerevisiae*, aromatic alcohols affect the transcript abundance of hundreds of different genes.7 Despite this strong effect, *FLO11* is the only differentially regulated gene for which the specific transcription factor that links aromatic
importance of S. cerevisiae the aromatic alcohols of this study. Applied to examining the effects of small molecules other than genome sequences, and comparative genomics data (Fig. 2).

present knowledge about the function of quorum sensing in related species. Here we address these questions by integrating the cell and the evolutionary conservation of its components in aromatic alcohols with altered expression is known. Considering the importance of S. cerevisiae as a model organism, it is important to gain a better understanding of the interaction of aromatic alcohol response with the rest of the transcriptional network of the cell and the evolutionary conservation of its components in related species. Here we address these questions by integrating present knowledge about the function of quorum sensing in S. cerevisiae with transcription regulatory networks, fungal genome sequences, and comparative genomics data (Fig. 2). We expect that the approach we present here can also be applied to examining the effects of small molecules other than the aromatic alcohols of this study.

Fig. 1 Utlrately, very few genes have been implicated to be part of the aromatic alcohol quorum sensing pathway in yeast. Aromatic aminotransferases (Aro8, Aro9 and Aro10), pyruvate decarboxylases (Pdc1, Pdc5, and Pdc6) and alcohol dehydrogenases (Adh) synthesise the aromatic alcohols tryptophol (TrpOH), phenylethanol (PheOH) and tyrosol (TyrOH) from the amino acids Trp, Phe and Tyr, respectively. Aro9 and Aro10 are regulated by the transcription factor Aro80, whose activity in turn is regulated by TrpOH, thereby establishing a positive feedback loop. TrpOH and PheOH also change the expression level of around 200 other target genes, but the precise pathway by which these changes are affected was previously unknown.

alcohols with altered expression is known. Considering the importance of S. cerevisiae as a model organism, it is important to gain a better understanding of the interaction of aromatic alcohol response with the rest of the transcriptional network of the cell and the evolutionary conservation of its components in related species. Here we address these questions by integrating present knowledge about the function of quorum sensing in S. cerevisiae with transcription regulatory networks, fungal genome sequences, and comparative genomics data (Fig. 2).

We expect that the approach we present here can also be applied to examining the effects of small molecules other than the aromatic alcohols of this study.

Materials and methods

Gene expression data

Identification of key transcription factors

In order to identify the transcription factors (TFs) that regulate the genes that are differentially expressed by the addition of aromatic alcohols, we integrated previously published gene expression data with transcriptional regulatory interactions that were reconstructed using three experimental methods that exploit different principles. One of them is based on ChIP-ChIP data, which identifies direct in vivo DNA binding events for TFs. The other two are genetic methods that identify differentially expressed genes upon over-expression or deletion of TFs using microarrays. We refer to the networks obtained in this way as the CC (for ChIP–ChIP; 144 TFs regulating 4347 targets via 12230 regulatory interactions), GRD (for Genetic Reconstruction upon gene Deletion; 157 TFs regulating 1992 targets via 5589 regulatory interactions) and GROE (for Genetic Reconstruction upon Over-Expression; 55 TFs regulating 1951 targets via 5802 regulatory interactions) network. In the CC network, nodes represent TFs or target genes (TGs) and edges represent direct binding of the TF in the promoter region of the TG. Likewise, in the GRD and GROE networks, nodes represent TFs or TGs and a TF is linked to a TG if it is differentially expressed upon deletion or over-expression of the TF, respectively. For each of the three networks independently, we calculated the score $d_f$, for each transcription factor $f$, where

$$d_f = \log_2 \left( \frac{\text{diffex}}{a_{\text{all}}} \right) \cdot \left( \frac{\text{diffex}}{a_{\text{all}}} \right)$$

$\text{diffex}$ is the number of differentially regulated genes controlled by $f$ according to the specific network, $\text{all}$ is the number of all genes regulated by $f$ in the same network. $d_f$ is the number of all differentially expressed genes upon treatment, and $a_{\text{all}}$ is the number of all genes in the network. In other words, this value corresponds to how much genes controlled by the TF $f$ in a particular network are overrepresented in the differentially expressed genes upon treatment. In this way, transcription factors that are likely to be responsible for regulating the genes that are differentially expressed upon the addition of the aromatic alcohols can be identified. In our analysis, we chose TFs that were more than two-fold overrepresented ($d_f > 1$) as
those which are likely to regulate the differentially expressed genes. We also calculated the statistical significance by estimating a \( p \)-value that each TF controls a higher proportion of the differentially expressed genes than what would be expected using a hypergeometric distribution as a null model (Supplementary Table S1). The Perl script and the datasets for this calculation are available from the authors on request.

**Genomic distribution of differentially expressed genes**

The chromosomal location of genes was obtained from http://www.yeastgenome.org/ for *S. cerevisiae* and from http://www.candidagenome.org/ for *C. albicans*. In order to assess whether an input set of genes is distributed randomly across the chromosomes or whether there is a preference for certain chromosomes, we performed two different types of statistical test. The first one is a \( \chi^2 \) test that has as a null hypothesis that the differentially expressed genes are distributed across the 16 yeast chromosomes in an unbiased manner. The second test computes the expected distribution of genes on the 16 chromosomes by repeatedly (1000 times) and randomly picking the same number of genes as the set of differentially expressed genes. This is compared to the observed distribution of the differentially expressed genes to obtain statistical significance for chromosomal preference (\( p \)-values and \( Z \)-scores). In order to assess whether the distribution of genes within chromosomes is random or clustered, we calculated the proximity index (PI) for the differentially expressed set of genes. The PI is defined as the ratio of proximal genes to the set of differentially expressed genes. Two genes are defined to be proximal if they are separated by less than 10 genes on the chromosome. To assess statistical significance, the PI value was calculated from random sets of genes (1000 in this case) and compared to the observed value.

**Orthologue detection**

In order to assess whether the genes comprising and regulated by aromatic alcohol communication are well-conserved within fungi, we identified their orthologues across 31 different fungal genomes. The complete genome sequences were downloaded from http://fungal.genome.duke.edu/ and the orthologues were identified using OrthoMCL with the default parameters. The OrthoMCL output was filtered and visualised using an in-house Perl script (Fig. 5). The Perl scripts and the sequence data used are available from the authors on request.

**Epigenetic modifications**

In order to determine whether genes that are differentially expressed upon the addition of aromatic alcohols tryptophol (TrpOH), phenylethanol (PheOH) and tyrosol (TyrOH), or a combination of all three (3OH), but not upon entry into stationary phase (A) By integrating this data with three transcription regulatory networks obtained in different ways (GRD: changes in gene expression upon transcription factor deletion, GROE: changes in gene expression upon transcription factor overexpression, CC: ChIP-ChIP), we determined the transcription factors that are most likely to be involved in the differential regulation of these genes. (B) By examining the evolutionary conservation of a subset of the differentially expressed genes, plus other genes that have previously been implicated in the quorum sensing response, we found that many of the key components of the quorum sensing system are only conserved in the genus *Saccharomyces*. (C) By examining the chromosomal location of the differentially expressed genes, we found that they tend to be distributed in a non-random fashion. (D) In order to determine whether genes that are differentially expressed upon the addition of aromatic alcohols have different histone modifications than genes that are not, we compared the histone modifications of those genes.
Results and discussion

Aromatic alcohols cause the differential expression of hundreds of genes by two-fold or more. The genes that are differentially expressed upon addition of the three different aromatic alcohols tryptophol, phenylethanol, tyrosol, or a combination of all three, which we refer to as 3OH, overlap significantly (Supplementary Fig. S2†). Phenylethanol has the strongest effect and causes the differential expression of 412 genes. Tryptophol causes the differential expression of 264 genes, tyrosol of 251 genes and 3OH of 246 genes (Table 1). 71 genes are differentially regulated by all four of these treatments (Supplementary Fig. S2†). Though tyrosol has not been implicated as a quorum sensing molecule, its effect on transcription with 251 differentially expressed genes is almost as strong as that of tryptophol with 264 differentially expressed genes and comparable to the 412 genes that are differentially expressed upon treatment with phenylethanol. The overlap between the differentially expressed genes between the two treatments is 141. The overlap between genes differentially expressed upon treatment with tyrosol and treatment with phenylethanol is 146 (Supplementary Fig. S2†).

Attributing differential expression to transcription factors

Filamentation and the cellular response to aromatic alcohol communication molecules involves a number of cellular changes, including the altered subcellular location of signalling proteins and changes in the transcript levels of select genes. We hypothesised that differential expression of genes upon addition of aromatic alcohols is due to the differential activity of only a few transcription factors. In order to establish their identity, we developed a method that integrates transcription regulatory networks with gene expression data. We independently integrated three different transcription regulatory networks with the set of differentially expressed genes. For each transcription factor, we calculated the value of for genes differentially expressed upon treatment (see Materials and methods). It should be noted that the transcription regulatory networks were not created under quorum sensing conditions and that chromatin remodelling under different conditions may have affected the topology of the networks by blocking or unblocking transcription factor access to gene regulatory regions. Despite these effects, our analysis clearly identifies key transcription factors that are likely to have a role in this process. Our results suggest that the transcription factors Msn2p, Mig1p, Rgm1p, Sip4p, and Cat8p regulate genes that are differentially expressed upon phenylethanol treatment according to more than one transcriptional network. Only Cat8p and Mig1p regulate the differentially expressed genes according to all three methods, which makes it likely that these genes are important for aromatic alcohol communication. Cat8p has previously been reported to be involved in the regulation of gluconeogenic genes and most of the enzymes of the glyoxylate cycle. Like Cat8p, Sip4p is also involved in the regulation of enzymes involved in gluconeogenesis. Mig1p is also involved in the regulation of enzymes that participate in the response to glucose repression by influencing activators of respiration and has also been shown to regulate CAT8. Because filamentous growth is promoted by quorum sensing, and because filamentation may aid in survival in nutrient limiting conditions by increasing access to nutrients, these findings are consistent. It has also been hypothesised that pleiotropic drug resistance (PDR) transporters are involved in quorum sensing in S. cerevisiae by exporting communication molecules. Our findings support this hypothesis as the transcription factor Pdr1p has a value of greater than one for genes differentially regulated upon the addition of tryptophol and tyrosol (Fig. 3). As Pdr1p also regulates the expression of the PDR transporter, this could mean that in an additional feedback loop quorum sensing molecules activate Pdr1p, which activates expression of PDR5, which may in turn lead to the export of further quorum sensing molecules. This is also supported by

Chen and Fink have demonstrated that the transcription factor Flo8p has differential activity in the presence of phenylethanol and tryptophol. Nevertheless, FLO8 is not differentially expressed after aromatic alcohol treatment. Neither does Flo8p control a higher proportion of differentially regulated genes than expected according to any transcriptional network upon phenylethanol treatment (Fig. 3 and Supplementary Table S1†). Upon tryptophol treatment, Flo8p regulates a higher proportion of differentially regulated genes only according to the GRD network (see Material and methods). Taken together, these results show that Flo8p is not solely responsible for the large changes in transcript levels observed upon aromatic alcohol treatment. The same applies to Aro80p, which is the only other transcription factor that has previously been reported to differentially regulate genes after the addition of an aromatic alcohol. Therefore, other transcription factors must be responsible for the regulation of the differentially expressed genes.

Table 1 In order to identify genes that are differentially expressed by aromatic alcohol treatment, but that are not also differentially expressed during entry into stationary phase, we removed all stationary phase-specific genes from the dataset (column 2). The number of C. albicans genes whose orthologues are differentially expressed in S. cerevisiae is given in the last column.

<table>
<thead>
<tr>
<th>Treatment</th>
<th>All differentially expressed genes</th>
<th>Genes not also differentially expressed upon entry into stationary phase</th>
<th>Proportional decrease (%)</th>
<th>C. albicans orthologues of differentially expressed S. cerevisiae genes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Phenylethanol</td>
<td>412</td>
<td>147</td>
<td>64</td>
<td>125</td>
</tr>
<tr>
<td>Tryptophol</td>
<td>264</td>
<td>114</td>
<td>57</td>
<td>63</td>
</tr>
<tr>
<td>Tyrosol</td>
<td>251</td>
<td>118</td>
<td>53</td>
<td>62</td>
</tr>
<tr>
<td>3OH (combination)</td>
<td>246</td>
<td>99</td>
<td>60</td>
<td>61</td>
</tr>
</tbody>
</table>
In order to identify genes that are differentially expressed upon entry into the stationary phase, microarray data from two different previously published datasets was obtained.\textsuperscript{12,13} We defined a gene as being differentially expressed upon entry into stationary phase if it was differentially expressed in any experiment involving entry into stationary phase in the dataset by Gasch \textit{et al.},\textsuperscript{12} as well as being differentially expressed in any experiment involving entry into stationary phase in the dataset by Andalis \textit{et al.}.\textsuperscript{13} Between 53\% and 64\% of the genes whose transcript levels are altered in response to aromatic alcohols also have altered transcript levels upon entry into stationary phase (Table 1). This shows that the two responses overlap but are distinct.

**Genomic organization and nucleosome modification of the differentially expressed genes**

Next, we investigated whether the genes that are differentially expressed upon the addition of aromatic alcohol communication molecules are distributed across the chromosomes randomly. Non-random distribution of these genes may allow quicker access of transcription factors to accessible regions of chromatin whenever this is required.\textsuperscript{21} To assess whether genes differentially expressed after aromatic alcohol addition are distributed across chromosomes in a specific way, we created random sets of genes and compared their chromosomal locations to the actually observed locations. We found that genes which are differentially expressed when aromatic alcohols are added are not distributed amongst chromosomes randomly. Genes that are differentially expressed upon the addition of phenylethanol tend to reside on chromosome 15 more often than expected by chance (Fig. 4). This also holds when genes that are differentially expressed upon entry into stationary phase are removed as described in the methods. No highly significant effect was found for tyrosol, which is in agreement with the finding that only tryptophol and phenylethanol are likely to be quorum sensing molecules.\textsuperscript{7}

We also found that genes that are differentially expressed in response to phenylethanol are clustered within chromosomes. 48.3\% of genes that are differentially regulated in response to phenylethanol are located within 10 genes of another gene\textsuperscript{20,21} regulated by this aromatic alcohol (\(P\text{I} = 0.483, p = 0.03\); Fig. 4). This drops to 26.1\% when filtering removing genes that are also differentially regulated upon entry into stationary phase (\(P\text{I} = 0.261, p = 0.02\)). This suggests that genes regulated by the communication molecule phenylethanol are not randomly distributed, but show preference for certain yeast chromosomes and regions on those chromosomes.

Altogether, this implies that genes that are differentially expressed upon the addition of the aromatic alcohol quorum sensing molecules tryptophol and phenylethanol are preferentially located in specific chromosomal territories. In order to investigate this further, we examined the nucleosome modifications of those differentially expressed genes by integrating our data with a large-scale histone modification dataset.\textsuperscript{24} Our results show that the histone modifications H3K4Me (Methylation of Lysine 4 on histone H3) and H4Ac (Acetylation of histone H4) are significantly (\(p < 0.01\)) less prevalent in genes that are differentially expressed upon the addition of tryptophol (Supplementary Fig. S3f).

**Quorum sensing and entry into stationary phase**

In order to identify genes that are differentially expressed upon entry into the stationary phase, microarray data from two different previously published datasets was obtained.\textsuperscript{12,13} We defined a gene as being differentially expressed upon entry into stationary phase if it was differentially expressed in any experiment involving entry into stationary phase in the dataset by Gasch \textit{et al.},\textsuperscript{12} as well as being differentially expressed in any experiment involving entry into stationary phase in the dataset by Andalis \textit{et al.}.\textsuperscript{13} Between 53\% and 64\% of the genes whose transcript levels are altered in response to aromatic alcohols also have altered transcript levels upon entry into stationary phase (Table 1). This shows that the two responses overlap but are distinct.

**Conservation of quorum sensing components in fungal genomes**

To investigate if the quorum sensing system is conserved in other fungal species, we identified the orthologues of key \textit{S. cerevisiae} genes involved in quorum sensing in 31 fungal species (Fig. 5 and Table 1). A similar method has been used previously to identify instances of the \textit{agr} quorum sensing system in bacteria.\textsuperscript{30} The results show the genes involved in aromatic alcohol synthesis are present in most yeast species. This might be because they occupy key positions in the amino acid metabolic network.\textsuperscript{31} The transcription factor Aro80p, which has been implicated in positive feedback between aromatic alcohols and the expression of their synthesis proteins, seems to be conserved only within the genera \textit{Saccharomyces} and \textit{Candida}. Most of the other transcription factors and signal transducers regulated by aromatic alcohols are conserved only in the genus \textit{Saccharomyces} (Fig. 5).

From our analysis of the chromosomal location of the differentially expressed genes, we observed that these genes are distributed in a non-random way across the 16 different chromosomes in \textit{S. cerevisiae}. This could be an effect of the
differentially expressed genes being regulated by the same system. In order to assess whether the orthologues of differentially expressed genes are also distributed in a non-random manner in other species, we investigated the chromosomal preference and clustering of those genes in *Candida albicans*. We found that the *C. albicans* orthologues are not clustered. When also taking into account experimental data, which shows that filamentation of *C. albicans* is not stimulated by tryptophol or phenylethanol, this could suggest that these genes do not act in the same pathway. Together with our finding that many of the genes that are involved in quorum sensing pathways in *S. cerevisiae* are not present in *C. albicans* and other more distantly related fungi, we suggest that it is unlikely that quorum sensing by aromatic alcohols is conserved in the same form outside the genus *Saccharomyces*.

Taken together, our results show that although quorum sensing using aromatic alcohols is a relatively recent evolutionary innovation in *S. cerevisiae*, and that the genes that are differentially expressed are not randomly distributed. Instead they show a clear chromosomal preference and tend to be clustered on the chromosomal arms. Through the integrated analysis of large-scale genomic datasets, we propose several
transcription factors previously not known to participate in quorum sensing, such as Cat8p and Mig1p, to be involved in the process. Our observation that the key elements in the circuit are conserved only in the genus *Saccharomyces* but not in other genomes suggests that the system is unlikely to operate in other species in the same form.

Fig. 5 The pattern of presence and absence of *S. cerevisiae* quorum sensing genes in other fungal species supports the notion that aromatic alcohol quorum sensing is a recent evolutionary innovation. The columns of the matrix refer to fully sequenced fungal genomes, and the rows refer to genes involved in quorum sensing. If a gene is present in a genome, the corresponding square is white, and if it is absent, it is black. The genes are grouped according to how they are involved in quorum sensing (aromatic alcohol synthesis, known components of quorum sensing signal transduction pathway, transcription factors that regulate genes differentially expressed by phenylethanol and tryptophol, transcription factors regulated by phenylethanol and tryptophol). The tree at the bottom of the figure shows how the fungal species are related to each other according to Fitzpatrick *et al.*

**Acknowledgements**

AW and MMB thank Roberto Zanchi and Nitish Mittal for critically reading the manuscript and the UK Medical Research Council for funding. MMB thanks Darwin College and Schlumberger Ltd. for funding.
References