bioRxiv preprint doi: https://doi.org/10.1101/033167; this version posted November 27, 2015. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

Analysis of a rare variant of mitotic gene *TAO3* reveals its meiotic interactors

Saumya Gupta[*], Aparna Radhakrishnan[*], Rachana Nitin[*], Pandu Raharja-Liu[†], Gen Lin[‡], Lars M. Steinmetz[‡,§,**], Julien Gagneur[†], Himanshu Sinha[*,1]

[*]Department of Biological Sciences, Tata Institute of Fundamental Research, Mumbai 400005, India

[†]Gene Center, Ludwig-Maximilians-Universität, 81377 Munich, Germany

[‡]European Molecular Biology Laboratory, Genome Biology Unit, 69117 Heidelberg, Germany

[§]Department of Genetics, Stanford University School of Medicine, Stanford, California 94305

[**]Stanford Genome Technology Center, Stanford University, Palo Alto, California 94304

[1]Corresponding author: Department of Biological Sciences, Tata Institute of Fundamental Research, Homi Bhabha Road, Colaba, Mumbai 400005, India. Email: hsinha@tifr.res.in

Running title:

Analyzing functional role of a rare variant

Keywords:

Rare variant, *TAO3*, meiosis, transcriptome profiling, allelic variant

bioRxiv preprint doi: https://doi.org/10.1101/033167; this version posted November 27, 2015. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

ABSTRACT

Genome-wide association studies have successfully identified thousands of common variants associated with complex traits and diseases. If only common variants are considered, a significant proportion of heritability in common diseases remains unexplained. One of the sources considered to explain this missing heritability are the rare genetic variants. Studying the consequences of rare genetic alterations offers additional opportunity for predicting molecular mediators underlying pathway deregulation. Here, we characterized the functional role of a rare variant having a significant contribution in sporulation efficiency variation. This causal variant is present in the coding sequence of *TAO3*, encoding a putative scaffolding protein conserved from yeast to humans and component of RAM network in yeast involved in mitosis. We observed that the role of *TAO3* allele on meiosis is independent of *ACE2*, a transcription factor regulated by RAM network during mitosis. By expressing *TAO3* causal allele conditionally during sporulation and quantitatively measuring cell cycle progression, we determined its role within the first 6 hours in meiosis, which coincides with cells entering into meiotic cell division. Time-resolved genome-wide gene expression analysis identified genes showing early and increasing expression trend during sporulation to be target genes of *UME6*, a crucial meiotic regulator. Genes regulating *UME6*, and other target genes expressed early, specifically, in the presence of causal *TAO3*, were chosen as candidate genes. Allele-specific functional validations identified *ERT1* (regulator of switch from fermentation to respiration) and *PIP2*, (regulator involved in beta-oxidation of fatty acids) as the mediating genes associated with *TAO3* causal allele and responsible for sporulation efficiency variation. Our study uncovered interesting link between *TAO3* and regulators of metabolic cues that modulate the switch between multiple developmental phenotypes, *viz.* mitosis to meiosis. Although a small proportion in a population might contain rare variants, identification of novel regulators of sporulation from our study highlighted the significance of studying rare genetic variants to obtain novel insights into the phenotype and disease biology.

bioRxiv preprint doi: https://doi.org/10.1101/033167; this version posted November 27, 2015. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

INTRODUCTION

The genetic architecture of complex traits is not completely understood. The 'common disease, common variant' rationale of genome-wide association studies (GWAS) is being challenged owing to the limited proportion of disease heritability explained by common variants, giving rise to the 'missing heritability' problem in quantitative genetics (Manolio *et al.* 2009; Zuk *et al.* 2014). Not considering the effect of rare variants has been suggested as one of the potential contributor of this 'hidden' instead of 'missing' heritability (Saint Pierre and Génin 2014). This view has been substantiated by the identification of multiple rare variants that confer considerable risk in diseases such as autism, schizophrenia and epilepsy (Stankiewicz and Lupski 2010). Since the initial genomic technologies were more adept at identifying common variants, sequencing methods and rare variant association studies are now being increasingly used to identify rare variants (Cirulli and Goldstein 2010; Zuk *et al.* 2014). Once identified, characterizing the functional role of rare variants associated with complex diseases has the potential for revealing new biology.

Common or rare, if genetic variants are in the core regulatory pathways underlying the trait, understanding the biological mechanism of phenotypic modulation is more direct. However, many genetic variants are identified in the ancillary biological processes. These processes are not part of the core regulatory pathways underlying the trait, but could impinge on the phenotype by being associated with the core pathways. For instance, gene identified by GWAS for face shape variation is a gap junction protein, which indirectly affects the mechanical load taken up by face bones (Liu *et al.* 2012). Ancillary pathways, hence, provide a good hub for accumulation of genetic variation. Their location in the regulatory network confers an advantage in regulating the phenotype while the core pathway genes could be too important to be varied, and hence probably under a negative selection pressure. This could be one of the reasons why crucial face forming gene of *Shh* has not been identified for facial shape development (Hallgrimsson *et al.* 2014). Therefore, characterization of variants, not part of the core biological processes associated with the phenotype is challenging.

Here, we characterized the functional role of a rare variant in the novel mitotic gene, *TAO3*, in yeast sporulation efficiency variation. We studied genome-wide gene

bioRxiv preprint doi: https://doi.org/10.1101/033167; this version posted November 27, 2015. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

expression dynamics in presence of both the alleles (common and causative rare alleles) of *TAO3 (G4477C)* during sporulation, and identified the core meiotic pathways and the ancillary pathways associated to sporulation showing differential expression. Our approach showed regulatory pathways involved in nutrient metabolism impinging on the core sporulation pathway, contributing to trait variation.

MATERIALS AND METHODS

Yeast strains and media

The yeast strains were grown in standard conditions at 30˚C in YPD (1% Yeast extract, 2% Bacto peptone, 2% dextrose). Whole-genome resequencing of the strain YAD331 (Deutschbauer and Davis 2005) with S288c strain as the reference strain, identified two additional polymorphisms (Supporting Figure S1, Supporting Table S1). Three consecutive backcrosses were performed between the haploid obtained from this strain and the haploid reference strain, to remove the secondary polymorphisms. The sole genetic difference between the reference S288c strain and the backcrossed allele replacement strain was at *TAO3(G4477C)* position, which was confirmed by performing PCR-based sequencing 650bp up and downstream around the two secondary polymorphisms and the *TAO3* polymorphism region. This backcrossed strain was diplodized to make it homozygous at *TAO3(4477C)* position and was termed as "T strain" in this study. The diploid parental strain S288c was termed as "S strain" in the study. All gene deletions in the study were made in the haploids of T and S strains, except the ones made in SK1 strain (Supporting Table S2). Deletions were performed and verified as described previously (Goldstein and McCusker 1999; Gietz and Woods 2002). All the experiments in this study were performed using the diplodized parent strains and their diploid derivatives. For replacing the endogenous *TAO3* promoter (-150 to -1bp upstream start site) in T strain, with a tetracycline-responsive promoter, a *tetO*$_7$-based promoter substitution cassette containing *kanMX4* was amplified from the plasmid pCM225 (Bellí *et al.* 1998b). Diploid T strain with this *tetO*$_7$-based cassette is termed P_{Tet}-*TAO3(4477C)* strain. The primers for sequencing, deletions and their confirmations are listed in Supporting Table S3.

bioRxiv preprint doi: https://doi.org/10.1101/033167; this version posted November 27, 2015. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

126 ***Phenotyping***

127 Sporulation efficiency estimation at 48h, progression through meiotic landmark
128 events Meiosis I (MI) and Meiosis II (MII) and its quantitation was done as described
129 previously (Gupta *et al.* 2015). For quantitation of meiotic landmarks in T strain,
130 parametric curves assuming delayed and 1st order kinetics were fitted to the DAPI-
131 stained meiotic progression time course data and fitting uncertainties were estimated
132 by bootstrapping (Supporting File S1). Cell cycle progression data for S288c and SK1
133 strains was taken from Gupta *et al.* (2015) (Figure 1D-E). Conditional expression of
134 *TAO3(4477C)* was performed by constructing P_{Tet}-*TAO3* strain (details in Supporting
135 File S1), which was responsive to tetracycline-analogue doxycycline (Bellí *et al.*
136 1998a; b). Addition of 2µg/ml doxycycline was utilized to decrease the activity of the
137 gene, unless any other concentration was stated. For each strain, minimum three
138 biological replicates were used and the experiment was carried minimum two times.
139 Approximately 300 cells were counted per replicate. Fold difference was calculated as
140 the ratio of mean sporulation efficiencies of the two strains A and B when the
141 sporulation efficiency of A is greater than of B.
142
143 ***Statistical test for calculating sporulation efficiency***
144 For comparing sporulation efficiency, two statistical tests were used: the pair test and
145 the interaction test. The pair test tests the null hypothesis that two given strains (S and
146 T) have the same sporulation efficiency.
147
148 The number $y_{i,k}$ of sporulated cells (4-nuclei count) among the total number of cells
149 $n_{i,k}$ of strain i in replicate experiment k was modeled with a quasi-binomial
150 generalized linear model using the *logit* link function and subject to a common log-
151 odd ratio β_i between replicates, *i.e.*:
152

153 $$\log\left(\frac{\mu_{i,k}}{n_{i,k}-\mu_{i,k}}\right)=\beta_i \text{ for all } k,$$

154
155 where, $\mu_{i,k}=E\left(y_{i,k}\right)$.
156

bioRxiv preprint doi: https://doi.org/10.1101/033167; this version posted November 27, 2015. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

The pair test tests the null hypothesis of equality of log odd-ratios for two strains *i* and *j*, *i.e.* $H_0 : \beta_i = \beta_j$.

In the case of S and T strains, the interaction test tests the null hypothesis that the effect of mutation A is independent of the effect of mutation B, taking the T strain as reference background. This test, thus, compares four strains: mutation A only, mutation B only, both A and B and neither A nor B (T strain). Here, the strain S was considered as a T strain mutated for *TAO3(4477)*. For every interaction test, we considered the dataset of the four strains of interest and fitted a quasi-binomial generalized linear model using the *logit* link function and subject to:

$$\log\left(\frac{\mu_{i,k}}{n_{i,k}-\mu_{i,k}}\right) = \beta_0 + \beta_A A_i + \beta_B B_i + \beta_{A,B} A_i B_i \text{ for all } k,$$

where, A_i and B_i are indicator variables of the mutations A and B in strain *i* respectively. The interaction test tested the null hypothesis that the odd ratio of sporulation in the double mutant equals the product of the odd ratios of each mutation, *i.e.* $H_0 : \beta_{A,B} = 0$.

Both the pair test and the interaction test were implemented in the statistical language R with the function *glm()* assuming a constant variance function fitted by maximizing the quasi-likelihood and using the t-test on tested parameters (Gupta *et al.* 2015).

Whole genome gene-expression profiling

Sporulating yeast cell collection at 0h, 30m, 45m, 1h10m, 1h40m, 2h30m, 3h50m, 5h40m and 8h30m (logarithmic time-series), RNA isolation and cDNA preparation were performed as described (Xu *et al.* 2009). Samples were hybridized to *S. cerevisiae* yeast tiling array (Affymetrix, Cat# 520055). Arrays at each time point for both the strains were normalized together using *vsn* normalization method (Huber *et al.* 2002).

bioRxiv preprint doi: https://doi.org/10.1101/033167; this version posted November 27, 2015. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

187 ***Whole genome gene-expression analysis***

188 Within each strain, the $\log_2$ expression values obtained were smoothed using *locfit* at
189 optimized bandwidth parameter $h = 1.2$ (Supporting Figure S2), base transformed for
190 each transcript by subtracting the expression value at each time point from the
191 baseline value at time point t = 0h (t_0) (Supporting Table S4. This $\log_2$ fold change
192 value with respect to t_0 is described as "expression" throughout the manuscript. For
193 identifying the genes showing temporal differential expression between T and S
194 strains (Supporting Table S5, method implemented in EDGE software was used,
195 which calculated statistically significant changes in expression between T and S strain
196 over time (Storey *et al.* 2005). The differentially expressed genes were clustered
197 according to their temporal expression patterns using time abstraction clustering
198 algorithm implemented in the TimeClust software (Magni *et al.* 2008) (see Supporting
199 File S1). Four major clusters were identified in each strain: Cluster I (Early trend),
200 Cluster II (Increasing trend), Cluster III (Late trend), Cluster IV (Repressing trend)
201 (Supporting Table S6). The transcription factors regulating a cluster of genes were
202 extracted using the YEASTRACT database (Teixeira *et al.* 2014). Only those
203 transcription factors were considered as candidate genes whose target genes were
204 significantly enriched in the corresponding cluster ($P \leq 0.05$, odds ratio ≥ 1.5).
205 YEASTRACT database was also used to obtain the regulation matrix of yeast, for
206 identifying target genes of regulators in this study such as *UME6*. Target genes for
207 *ACE2* were obtained from Nelson *et al.* (2003). Significantly enriched Gene Ontology
208 terms by biological process (Bonferroni corrected $P < 0.05$, Table 1) were obtained
209 from SGD Yeastmine (Balakrishnan *et al.* 2012).
210
211 ***Data availability***
212 The array data for T strain reported has been deposited in ArrayExpress
213 (http://www.ebi.ac.uk/arrayexpress/) with accession number E-MTAB-3889. The
214 whole genome sequence data for T strain has been deposited in the European
215 Nucleotide Archive (http://www.ebi.ac.uk/ena/) with the accession number
216 PRJEB8698. The rest of the data is available as Supporting Information. Array data
217 and whole genome sequence data for S strain were downloaded from Gupta *et al.*
218 (2015).
219

bioRxiv preprint doi: https://doi.org/10.1101/033167; this version posted November 27, 2015. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

RESULTS

Role of TAO3 causative allele in sporulation efficiency variation

The causative allele in *TAO3(4477C)*, identified in high sporulating SK1 strain (Deutschbauer and Davis 2005), was not present in other high sporulating natural and laboratory yeast strains in the SGRP collection (Figure 1A). This indicated that *TAO3(4477C)* was a rare variant for high sporulation efficiency. Substitution of this causative single nucleotide polymorphism *(G4477C)* in 7,131bp long *TAO3*, in the low sporulating S288c strain, was sufficient to increase its sporulation efficiency by three-fold ($P = 1.8 \times 10^{-10}$, pair test in Methods, Figure 1B) compared to the S288c strain (*TAO3(4477G)*), reconfirming a significant role of *TAO3(4477C)* SNP in sporulation efficiency variation (Deutschbauer and Davis 2005). The increased sporulation efficiency in T strain saturated within 48h and did not increase beyond three-folds even when kept in sporulation medium for a week (Figure 1C). Quantitative analysis of cell cycle progression in T strain predicted that the difference between T and S strain occurred already during their entry into meiosis, since both, the time to initiate meiosis and rate of transition from G_1/G_0 into MI, were significantly different between the two strains (Figures 1D-E, Supporting Figure S3). Studying the progression of meiotic phases also showed that the T strain initiated meiosis within 12h (Figures 1D-E). To further resolve this 12h temporal phase when *TAO3(4477C)* could have an effect on the phenotype, this allele was conditionally expressed during sporulation (see Methods). The sporulation efficiency of the tetracycline-responsive T strain (P_{Tet}-*TAO3(4477C)* strain) in the presence of 2µg/ml doxycycline throughout the sporulation period was equivalent to the S strain (Figure 1F). Shorter periods of inactivation of *TAO3* activity (in presence of doxycycline) showed that a 3h inhibition indeed reduced sporulation efficiency but this was significantly different ($P = 0.02$) from 48h inhibition, although, a 6h inhibition was equivalent to 48h inhibition. This showed that *TAO3(4477C)* allele affected sporulation efficiency within the first 6h in sporulation.

Role of TAO3 in meiosis is distinct from its role during mitosis

A part of the RAM (Regulation of Ace2p activity and cellular Morphogenesis) signaling network, Tao3 is required for activation and localization of an NDR protein kinase, Cbk1, another essential component of RAM network (Du and Novick 2002; Hergovich *et al.* 2006). This signaling network (consisting of Cbk1, Hym1, Kic1,

bioRxiv preprint doi: https://doi.org/10.1101/033167; this version posted November 27, 2015. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

Mob2, Tao3) contributes to various important aspects of mitotic cellular growth, particularly in cell separation by regulating transcription factor Ace2, in polarized growth (Nelson *et al.* 2003) and in cellular progression, through a Ace2-independent pathway (Bogomolnaya *et al.* 2006). Ace2 is a transcription factor that peaks early in mitosis and is involved in G_1/S transition (Spellman *et al.* 1998). In S288c-background strains, the RAM network genes are essential and hence the affect of their deletion in sporulation efficiency has not been previously studied (Deutschbauer *et al.* 2002). Here, we tested if this upstream mitotic function of RAM network was also involved in meiotic regulation.

To determine the unique molecular effects of the causative *TAO3* allele during sporulation, we compared genome-wide expression between T and S strains (see Methods). Genome-wide gene expression was measured for both these strains from 0h to 8h30m in sporulation medium (see Methods). Over one thousand gene transcripts (1,122 including non-coding SUTs, Supporting Table S5) were identified to be statistically significant in gene expression as a function of time between the two strains (FDR cut-off 10%). However, in this set, none of the RAM network genes showed significant differential expression in the presence of *TAO3(4477C)*. A few of *ACE2*-regulated genes showed differential expression in the presence of *TAO3(4477C)* (11 genes labeled green in Figure 2A), however, deletion of *ace2*Δ had no affect on the sporulation efficiency of T strain (Figure 2B). Furthermore, even in the high sporulating parent SK1, *ace2*Δ, which showed a lesser clumping phenotype during growth in rich medium, had no effect on sporulation efficiency (Figure 2B). Since previously Ace2-independent effect of RAM network on cellular polarization have been observed (Nelson *et al.* 2003), therefore, RAM network could still be involved in meiosis independently of Ace2. Since by reducing *TAO3(4477C)* activity only during the growth phase in rich medium (in glucose) had no effect the sporulation efficiency (Figure 1F), hence we concluded that the causative allele of *TAO3* might affect sporulation efficiency by interacting with a set of genes unique from the gene set involved during mitosis.

While studying the expression profiles of RAM network genes, we also observed a higher expression of *TAO3* in the presence of the causative polymorphism in the T strain, compared to the S strain ($P = 0.004$, Figure 2C). By altering *TAO3* expression,

bioRxiv preprint doi: https://doi.org/10.1101/033167; this version posted November 27, 2015. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

using different concentrations of doxycycline, we found that higher than native expression of *TAO3* increased sporulation efficiency (no doxycycline added, $P < 1 \times 10^{-4}$) and concordantly, a decrease in expression (increased doxycycline concentration) reduced sporulation efficiency (Figure 2D). These results suggested an interesting possibility of how the coding polymorphism in the transcriptional activator *TAO3* might be involved in sporulation efficiency variation by affecting its gene expression.

Global gene expression variation during sporulation in presence of causative TAO3 allele

While majority of regulation of sporulation occurs during both early and middle stages of meiosis (Neiman 2011), even late genes are known to affect sporulation efficiency in yeast (Deutschbauer *et al.* 2002). Between T and S strains, all of these classes of sporulation genes, *i.e.*, early (*IME2*, *HOP1*, *DMC1*), middle and late (*NDT80*, *SPR3*, *SPS1*) showed differential expression (Figure 3A, Supporting Table S5). Most of the known sporulation genes (*IME2, DMC1, NDT80, SMK1, SWM1, RAD6, REC8*), including various regulators of meiosis (*IME1, IME2, NDT80*) were enriched ($P = 5.5 \times 10^{-12}$) in the cluster showing an increasing expression (Cluster II) during sporulation in T strain (Figure 3B). Approximately 50% of genes in this cluster (including *IME1*, *IME2*, *DMC1*, *ECM11*, *NDT80*) showed a similar trend in S strain also (Supporting Figure S4, Supporting Table S6). Our cell cycle progression results showed that lesser number of cells in S strain entered meiosis compared to T strain (Figure 1C). Higher expression of meiotic and meiosis-associated genes in T strain also reflected more cells entering meiosis. These results indicated that the early effects of *TAO3* allele on sporulation were important for regulation of sporulation. The early expressing Cluster I genes of T strain belonged to biological processes regulating entry in sporulation, such as carbohydrate metabolic process, ion transport, mitochondrion organization and cellular respiration (Table 1). Sparse overlap (7%, Supporting Figure S4) was observed in genes of Cluster I between T and S strains, with many meiosis-associated pathway genes (carbohydrate metabolic process and mitochondrial organization) showing repression in S (Table 1, Supporting Figure S5). We next studied the genes regulating sporulation genes in Cluster II and of those getting expressed earlier than these sporulation genes, *viz.* early genes in Cluster I (Supporting Tables S7, S8).

bioRxiv preprint doi: https://doi.org/10.1101/033167; this version posted November 27, 2015. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

Identifying candidate genes mediating the allele specific effects of TAO3 during sporulation using the temporal gene expression data

Regulators of genes upregulated either early or increasingly uniquely in T strain as time progresses in sporulation were enriched in nutrient metabolism and chromatin modification - the biological processes that are important for initiation of meiosis (Neiman 2011). *UME6*, a transcriptional repressor whose degradation is required for meiotic progression (Kassir *et al.* 2003) and is known to cause sporulation defects in both S288c and SK1 strains, was also identified. This core sporulation gene is also associated with various other important biological processes initiating meiosis (Lardenois *et al.* 2015) and was identified as a regulator for most of the biological processes occurring early in the presence of *TAO3(4477C)* (Table 2). From the list of regulators (Supporting Tables S7, S8), we identified those regulators that were upstream to *UME6* also (Figure 4A, Supporting Table S9). These sporulation-associated regulators were selected as candidate causal mediating genes. They included *ERT1*, regulator of carbon source utilization (Turcotte *et al.* 2009) involved in the switch from fermentation to respiration in glucose-limiting conditions (Gasmi *et al.* 2014); *OAF1,* regulator of lipid metabolism forming a protein complex with *PIP2* (Karpichev and Small 1998); and *DAL81,* regulator of nitrogen degradation pathway (Marzluf 1997). Interestingly, similar to *UME6*, *OAF1* target genes were repressed (Cluster IV) in S strain (Supporting Table S10). Earlier work in S and SK1 strains has shown up regulation of *ERT1*, *PIP2* and *DAL81* in SK1 strain during sporulation (Primig *et al.* 2000), though their deletion in S strain did not effect sporulation efficiency (Deutschbauer *et al.* 2002).

A few other sporulation-associated regulators (Supporting Tables S7, S8) that were not upstream *UME6* were also considered as candidate genes to be further investigated. These were: *GAT3*, a transcription factor involved in spore wall assembly (Lin *et al.* 2013); *RSC2*, required for the expression of mid-late sporulation genes (Bungard *et al.* 2004); *GAT1* and *PHO4*, regulators of nitrogen and phosphorus metabolism, respectively. These sporulation genes are not known to have an effect on sporulation efficiency in S288c strain (Deutschbauer *et al.* 2002). Another putative causal mediating gene whose target genes were enriched in the Cluster II uniquely in *TAO3(4477C)* allele was *XBP1*, a transcriptional repressor induced in stress and

bioRxiv preprint doi: https://doi.org/10.1101/033167; this version posted November 27, 2015. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

starvation, involved in chromatin modification and G_1 to S cell cycle progression. In SK1 strain, *XBP1* was highly induced during meiosis and its deletion reduced sporulation efficiency (Mai and Breeden 2000), whereas, *xbp1*Δ in S strain showed no effect on its sporulation efficiency (Deutschbauer *et al.* 2002). Interestingly, *XBP1* regulates *VAC8*, a vacuolar membrane protein required for efficient sporulation (Neiman 2005; Tang *et al.* 2006) that is also known to interact with *TAO3* by yeast-two hybrid studies (Tang *et al.* 2006).

In order to functionally validate a few of the candidate genes identified from the above analysis, we used a genetic model described previously (Gupta *et al.* 2015). According to this model, if a gene has no effect on sporulation efficiency, its deletion would not affect on the phenotype in T strain. If a gene had an independent role in sporulation, reduction in sporulation efficiency by its deletion would be independent of the *TAO3* background, and we would observe an additive effect irrespective of the background. Any significant deviation from this expectation would imply dependence on the genotype, such as the case of epistasis where deleting the gene in S background would not lead to decreased sporulation efficiency. While no difference was observed for *GAT1* (Supporting Figure S6), single deletions of *ERT1*, *PIP2* and *GAT3* reduced the mean sporulation efficiency in the T strain significantly, by about 1.5-fold ($P = 2.1 \times 10^{-12}$, $P = 6.1 \times 10^{-13}$, $P = 9.6 \times 10^{-10}$ respectively, pair test in Methods, Figure 4B). Significant interaction terms (see Methods) were obtained between the genetic background (S and T) and *ert1*Δ and *pip2*Δ ($P = 2.3 \times 10^{-4}$, $P = 0.04$), but not for *gat3*Δ. This showed that the effect of *ert1*Δ and *pip2*Δ on sporulation efficiency was specific to *TAO3(4477C)*. These analyses concluded that the meiotic role of allelic variant of *TAO3* was complex and was effecting sporulation efficiency by interacting with regulatory pathways involved in carbon metabolism (*ERT1*) and membrane component proteins involved in beta-oxidation (*OAF1-PIP2*).

DISCUSSION

Strong effects on phenotypic variation have been observed as a consequence of rare coding variants (Cohen *et al.* 2004; 2005). *TAO3(4477C)*, a rare variant specific to the high sporulating laboratory yeast strain of SK1 isolated from soil, is not present in the other yeast strains in the SGRP collection (Liti *et al.* 2009). Here, we showed the applicability of temporal genome-wide transcriptome profiling approach (Gupta *et al.*

bioRxiv preprint doi: https://doi.org/10.1101/033167; this version posted November 27, 2015. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

2015) to understand more about regulation of yeast sporulation efficiency variation for the rare causal genetic variant *TAO3*.

Tao3, conserved from yeast to humans (Hergovich *et al.* 2006) has been functionally annotated for mitotic cell division (Du and Novick 2002; Nelson *et al.* 2003), until it got mapped for sporulation efficiency variation (Deutschbauer and Davis 2005). The SK1 allele of *TAO3* has also been identified to affect growth variation in glycerol and high temperature (Wilkening *et al.* 2014). Tao3 localizes to polarized bud sites during mitosis (Nelson *et al.* 2003). Here, we observed genetic interaction of *TAO3(4477C)* with membrane-associated *ERT1*. Thus, it would be interesting to ascertain the localization of *TAO3* variants during sporulation. Similar to other scaffolding proteins like *Fry* (*Drosophila*) and *SAX-2* (*C. elegans*), Tao3 has multiple conserved Armadillo-like repeats (Hergovich *et al.* 2006) and the causal sporulation variant is present in one of these domains. Since overexpression of *TAO3(4477G)* is known to negatively affect the number of cells transitioning from G_1 to S phase during mitosis (Bogomolnaya *et al.* 2006), a higher expression of the causative *TAO3* allele specifically in meiosis raised the interesting possibility of this expression level variation to be involved in G_1 to S phase transition in meiosis.

In complex diseases, combinations of molecular perturbations, which may vary in different patients, may be dysregulating similar components of a cellular system. Using a reductionist view of disease biology, many drug candidates have been successfully identified by studying the core pathways associated to a disease (Schadt *et al.* 2009). However, variation in the ancillary molecular players impinging on the core pathways and residing in the genetic bottleneck could also contribute variation in the phenotype (Lorenz and Cohen 2014). Genetic variation in mitosis-associated genes of *MKT1* and *TAO3* are examples of the non-canonical pathways associated with sporulation efficiency variation. Moreover, the novel metabolic and mitochondrial regulators identified for *MKT1* (Gupta *et al.* 2015) and in this work for *TAO3* as the mediating pathways causing variation in sporulation efficiency, are not part of the core sporulation regulatory framework. Their modulation by genetic variants provides important implications for understanding the rewiring of cellular networks underlying phenotypic variation. In this study, *ERT1*, regulator of carbon catabolite activation of transcription involved in the shift between fermentation to

bioRxiv preprint doi: https://doi.org/10.1101/033167; this version posted November 27, 2015. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

424 respiratory growth (Gasmi *et al.* 2014) and *OAF1-PIP2,* regulators of lipid
425 metabolism necessary for the expression of sporulation specific genes *SPS18* and
426 *SPS19* (Gurvitz *et al.* 2009), were identified as mediating pathways contributing to
427 efficiency in meiosis. The molecular pathways in which these regulators were
428 involved in, emphasized the role of pleiotropic metabolic pathways that act as a
429 switch between multiple developmental phenotypes (Granek *et al.* 2011), to also have
430 important functional roles in variation in these phenotypes. Since, Tao3 is a
431 scaffolding protein involved in mitosis, and our findings link this protein in meiosis,
432 this puts forth an interesting hypothesis for its dual role of in recruiting different
433 molecular players in the two developmental phenotypes. Identification of these
434 ancillary pathways contributing to variation in meiosis suggested that a systems view
435 is necessary to understand the underlying complex biology of phenotypes including
436 various diseases. This systems-based study provides support to the emerging view of
437 disease as a complex network of interconnected pathways (Schadt *et al.* 2009), which
438 could be used to refine the clinical translation of GWAS results to understand disease
439 mechanism.
440
441

bioRxiv preprint doi: https://doi.org/10.1101/033167; this version posted November 27, 2015. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

TABLES

Table 1. Comparison of functional GO categories of differentially expressed genes in T strain clusters with S strain. See Supporting Table S6 for full list of genes in each cluster.

Cluster	**Functional GO category**	**Genes**
Early in T strain (Cluster I)	Carbohydrate metabolic process	*DOG1*, *YPI1*
	Ion transport	*AVT4*, *DAL5*
	Mitochondrion organization	*PPE1*, *UPS3*
	Cellular respiration	*COX5B*
Early in T strain (Cluster I) repressed in S strain (Cluster IV)	Carbohydrate metabolic process	*ALG6*, *DEP1*, *DOG1*, *TPS3*, *YPI1*
	Mitochondrial organization	*ATG33*, *COX20*, *PPE1*, *UPS3*

bioRxiv preprint doi: https://doi.org/10.1101/033167; this version posted November 27, 2015. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

Table 2. Functional GO classification of regulators of differentially expressed genes showing early and increasing expression only in T strain. See Supporting Tables S7, S8 for full list of genes.

Functional GO category	**Regulators**	***P* value**
Carbon metabolism	*ERT1*, *OAF1*, *PIP2*, *MIG1*, *MIG2*	1.9×10^{-6}
Nitrogen catabolite regulation	*DAL81*, *DAL82*, *GAT1*, *UME6*	1.7×10^{-5}
Chromatin modification	*ISW1*, *PHO2*, *PHO4*, *UME6*, *OAF1*, *XBP1*, *SIF2*, *RSC2*	1.4×10^{-5}

bioRxiv preprint doi: https://doi.org/10.1101/033167; this version posted November 27, 2015. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

ACKNOWLEDGEMENTS

We thank Manu Tekkedil for help with whole genome sequencing sample preparation. This research was supported by Tata Institute of Fundamental Research intramural funds and Department of Biotechnology grant BT/PR14842/BRB/10/881/2010 (H.S.); Bavarian Research Center for Molecular Biosystems and Bundesministerium für Bildung und Forschung through the Juniorverbund in der Systemmedizin "mitOmics" grant FKZ 01ZX1405A (J.G.); the National Institutes of Health, Deutsche Forschungsgemeinschaft and a European Research Council Advanced Investigator Grant (L.M.S.). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

bioRxiv preprint doi: https://doi.org/10.1101/033167; this version posted November 27, 2015. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

FIGURE LEGENDS

Figure 1. Role of *TAO3* in sporulation efficiency

(A) Comparison of genomic sequence of *TAO3* (4,441-4,500) across the SGRP collection (Liti *et al.* 2009). The 4,477th position of *TAO3* consists of the sporulation causative variant, where identical nucleotides are indicated by the same color. Identity indicates the percentage match between the nucleotides in the shown region of the gene. The strains are ordered according to their mean sporulation efficiency (Tomar *et al.* 2013): high (60-100%), intermediate (10-60%), low (0-10%) and ND (not determined).

(B) Bar plots represents the mean sporulation efficiency, after 48h, of the SK1, T and S strains. The sporulation efficiency data is indicated as circles.

(C) Line graphs represent mean sporulation efficiency of S, T and SK1 strains measured till saturation, *i.e.*, till sporulation efficiency did not vary for 3 consecutive days.

(D) Percentage of 1-, 2- and 4-nuclei states of the T strain (y-axis) versus time in sporulation medium (x-axis). 1-nucleus stage is indicated as red circles (G_0/G_1 phase), 2-nuclei state as yellow circles (completion of MI phase) and blue circles is 4-nuclei stage (completion of MII phase).

(E) Bootstrap distribution of the time to initiate meiosis and rate of transition from G_1/G_0 into MI, estimated from time courses in (D). See Methods for details.

(F) Conditional expression of *TAO3(4477C)* during sporulation in P_{Tet}-*TAO3(4477C)* strain (depicted as P_{Tet}). Y-axis is mean sporulation efficiency in 48h. No doxycycline in growth (YPD) or spo (YPA + sporulation) medium is depicted as "-" condition on x-axis and addition of doxycycline is depicted as "+" in that condition. "+3h" condition in Spo implies doxycycline was throughout in the growth medium and in the sporulation medium till 3h, after which cells were sporulated in the absence of doxycycline. "+6h" condition implies doxycycline was throughout in the growth medium and in the sporulation medium till 6h, after which cells were sporulated in the absence of doxycycline. One-way ANOVA with Bonferroni's multiple comparison tests was performed to test significance.

bioRxiv preprint doi: https://doi.org/10.1101/033167; this version posted November 27, 2015. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

Figure 2. Role of *TAO3* in meiosis is distinct from its role during mitosis

(A) Heatmap showing gene expression of RAM network genes and Ace2-regulated genes in T and S strains. Gene names in green show differential expression (data in Supporting Tables S4, S5).

(B) Bar plots represent mean sporulation efficiency, after 48h, of the SK1 and T wild type (wt) and *ace2*Δ deletion strains. Pair and interaction tests (described in Methods) were performed to test significance.

(C) Expression profile ($\log_2$ fold change t_0) of *TAO3* is given in the y-axis for T (purple) and S strains (red) and the x-axis denotes the time in sporulation medium. (data in Supporting Tables S4, S5)

(D) Bar plots represent mean sporulation efficiency of P_{Tet}-*TAO3(4477C)* strain in absence and presence of different concentrations of doxycycline (x-axis), compared to S strain. One-way ANOVA with Bonferroni's multiple comparison tests was performed to test significance.

Figure 3. Global gene expression variation in presence of causative *TAO3* allele

(A) Temporal heat map of meiotic genes in T and S strains. The gene names shown in green are differentially expressed in the presence of *TAO3(4477C)*.

(B) The expression profile ($\log_2$ fold change t_0) for the meiotic landmark genes is given in the y-axis and the x-axis denotes the time in sporulation medium. Red line represents the expression profile of the respective gene in S strain and blue line is the same in T strain.

(C) Heat map of the T and S strains showing differentially expressed gene across time within each cluster. Each row represents a single gene and columns are time points of each strain (for gene list in each cluster see Supporting Table S6). The order of genes in the two strains is based on the clustering of the T strain. Functional GO categories of genes in each cluster are shown on left. The boxplots shown on right represent the average expression profile of each cluster in the T and S strain. The number of genes in each cluster in a strain is indicated in brackets.

Figure 4. Identifying candidate genes mediating the allele specific effects of *TAO3* during sporulation using the temporal gene expression data

(A) Regulatory network of candidate genes predicted to mediate the effects of *TAO3(4477C)* in sporulation. The candidate mediating genes are shown as bigger

bioRxiv preprint doi: https://doi.org/10.1101/033167; this version posted November 27, 2015. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

536 nodes (large circles), with their target genes (small circles) connected to them as
537 straight lines. The box contains the protein network interactions of the candidate
538 genes with core sporulation gene *UME6*, obtained from YEASTRACT (see Methods).
539 Colors inside the nodes were calculated as an average of the first six time points in
540 sporulation (early phase). For complete list of interacting genes and their expression
541 values, see Supporting Tables S9 and S4, respectively.
542 (B) Bar plots represent mean sporulation efficiency, after 48h, of the T and S wild
543 type (wt) and *ert1*Δ, *pip2*Δ and *gat3*Δ strains. Pair and interaction tests (see Methods)
544 were performed to test significance.
545
546

bioRxiv preprint doi: https://doi.org/10.1101/033167; this version posted November 27, 2015. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

SUPPORTING INFORMATION

Supporting Files

File S1. Detailed methods

Supporting Tables

Table S1. Whole genome resequencing results for the *TAO3* allele replacement strain

Table S2. Strains names

Table S3. Primer names

Table S4. Smoothed expression data, base transformed with respect to t_0 for T and S strains

Table S5. Differentially expressed genes between T and S strains, with their *P* and *Q* values calculated using EDGE

Table S6. Genes in each cluster using TimeClust

Table S7. Transcription factors regulating unique early (Cluster I) genes of the T strain

Table S8. Transcription factors regulating unique increasing (Cluster II) genes of the T strain

Table S9. Differentially expressed target genes of regulators of candidate genes mediating the affect of *TAO3*

Table S10. Transcription factors regulating unique repressing (Cluster IV) genes of the S strain

Supporting Figures

Figure S1. Whole genome resequencing of *TAO3* allele replacement strain (YAD331, (Deutschbauer and Davis 2005) in comparison to S288c reference strain

Figure S2. Smoothing of normalized temporal data using *locfit*

Figure S3. Mathematical modeling to identify stage of meiosis affected by *TAO3* causal allele

Figure S4. Comparison of genes showing early (Cluster I) and increasing trend (Cluster II) between T and S strains

Figure S5. Genes showing early expression in T strain show expression at later time points or repressed in S strain

Figure S6. Sporulation efficiency of *gat1*Δ in T strain

bioRxiv preprint doi: https://doi.org/10.1101/033167; this version posted November 27, 2015. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

582

583 **REFERENCES**

584 Balakrishnan R., Park J., Karra K., Hitz B. C., Binkley G., Hong E. L., Sullivan J.,
585 Micklem G., Cherry J. M., 2012 YeastMine—an integrated data warehouse for
586 Saccharomyces cerevisiae data as a multipurpose tool-kit. Database 2012:
587 bar062–bar062.
588 Bellí G., Garí E., Aldea M., Herrero E., 1998a Functional analysis of yeast essential
589 genes using a promoter-substitution cassette and the tetracycline-regulatable dual
590 expression system. Yeast 14: 1127–1138.
591 Bellí G., Garí E., Piedrafita L., Aldea M., Herrero E., 1998b An activator/repressor
592 dual system allows tight tetracycline-regulated gene expression in budding yeast.
593 Nucleic Acids Research 26: 942–947.
594 Bogomolnaya L. M., Pathak R., Guo J., Polymenis M., 2006 Roles of the RAM
595 signaling network in cell cycle progression in Saccharomyces cerevisiae. Curr
596 Genet 49: 384–392.
597 Bungard D., Reed M., Winter E., 2004 RSC1 and RSC2 are required for expression
598 of mid-late sporulation-specific genes in Saccharomyces cerevisiae. Eukaryotic
599 Cell 3: 910–918.
600 Cirulli E. T., Goldstein D. B., 2010 Uncovering the roles of rare variants in common
601 disease through whole-genome sequencing. Nat Rev Genet 11: 415–425.
602 Cohen J. C., Kiss R. S., Pertsemlidis A., Marcel Y. L., 2004 Multiple rare alleles
603 contribute to low plasma levels of HDL cholesterol. Science 305: 869–872.
604 Cohen J., Pertsemlidis A., Kotowski I. K., Graham R., Garcia C. K., Hobbs H. H.,
605 2005 Low LDL cholesterol in individuals of African descent resulting from
606 frequent nonsense mutations in PCSK9. Nat. Genet. 37: 161–165.
607 Deutschbauer A. M., Davis R. W., 2005 Quantitative trait loci mapped to single-
608 nucleotide resolution in yeast. Nat. Genet. 37: 1333–1340.
609 Deutschbauer A. M., Williams R. M., Chu A. M., Davis R. W., 2002 Parallel
610 phenotypic analysis of sporulation and postgermination growth in Saccharomyces
611 cerevisiae. Proceedings of the National Academy of Sciences 99: 15530–15535.
612 Du L.-L., Novick P., 2002 Pag1p, a novel protein associated with protein kinase
613 Cbk1p, is required for cell morphogenesis and proliferation in Saccharomyces
614 cerevisiae. Mol. Biol. Cell 13: 503–514.
615 Gasmi N., Jacques P.-E., Klimova N., Guo X., Ricciardi A., Robert F., Turcotte B.,
616 2014 The switch from fermentation to respiration in Saccharomyces cerevisiae
617 is regulated by the Ert1 transcriptional activator/repressor. Genetics 198: 547–
618 560.
619 Gietz R., Woods R. A., 2002 Transformation of yeast by lithium acetate/single-
620 stranded carrier DNA/polyethylene glycol method. Meth. Enzymol. 350: 87–96.
621 Goldstein A. L., McCusker J. H., 1999 Three new dominant drug resistance cassettes
622 for gene disruption in Saccharomyces cerevisiae. Yeast 15: 1541–1553.
623 Granek J. A., Kayıkçı Ö., Magwene P. M., 2011 Pleiotropic signaling pathways
624 orchestrate yeast development. Current Opinion in Microbiology 14: 1–6.
625 Gupta S., Radhakrishnan A., Raharja-Liu P., Lin G., Steinmetz L. M., Gagneur J.,
626 Sinha H., 2015 Temporal expression profiling identifies pathways mediating
627 effect of causal variant on phenotype. PLoS Genet 11: e1005195.
628 Gurvitz A., Suomi F., Rottensteiner H., Hiltunen J. K., Dawes I. W., 2009 Avoiding
629 unscheduled transcription in shared promoters: Saccharomyces cerevisiae Sum1p
630 represses the divergent gene pair SPS18-SPS19 through a midsporulation element

bioRxiv preprint doi: https://doi.org/10.1101/033167; this version posted November 27, 2015. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

631 (MSE). FEMS Yeast Res 9: 821–831.
632 Hallgrimsson B., Mio W., Marcucio R. S., Spritz R., 2014 Let's Face It—Complex
633 Traits Are Just Not That Simple. PLoS Genet 10: e1004724.
634 Hergovich A., Stegert M. R., Schmitz D., 2006 NDR kinases regulate essential cell
635 processes from yeast to humans. Nature Reviews Molecular Cell Biology 7: 253–
636 264.
637 Huber W., Heydebreck von A., Sültmann H., Poustka A., Vingron M., 2002
638 Variance stabilization applied to microarray data calibration and to the
639 quantification of differential expression. Bioinformatics 18 Suppl 1: S96–104.
640 Karpichev I. V., Small G. M., 1998 Global Regulatory Functions of Oaf1p and
641 Pip2p (Oaf2p), Transcription Factors That Regulate Genes Encoding Peroxisomal
642 Proteins in Saccharomyces cerevisiae. Molecular and Cellular Biology 18: 6560–
643 6570.
644 Kassir Y., Adir N., Boger-Nadjar E., Raviv N. G., Rubin-Bejerano I., Sagee S.,
645 Shenhar G., 2003 Transcriptional regulation of meiosis in budding yeast. Int.
646 Rev. Cytol. 224: 111–171.
647 Lardenois A., Becker E., Walther T., Law M. J., Xie B., Demougin P., Strich R.,
648 Primig M., 2015 Global alterations of the transcriptional landscape during yeast
649 growth and development in the absence of Ume6-dependent chromatin
650 modification. Mol. Genet. Genomics 290: 2031–2046.
651 Lin C. P.-C., Kim C., Smith S. O., Neiman A. M., 2013 A Highly Redundant Gene
652 Network Controls Assembly of the Outer Spore Wall in S. cerevisiae. PLoS
653 Genet 9: e1003700.
654 Liti G., Carter D. M., Moses A. M., Warringer J., Parts L., James S. A., Davey R. P.,
655 Roberts I. N., Burt A., Koufopanou V., Tsai I. J., Bergman C. M., Bensasson D.,
656 O'Kelly M. J. T., van Oudenaarden A., Barton D. B. H., Bailes E., Nguyen A. N.,
657 Jones M., Quail M. A., Goodhead I., Sims S., Smith F., Blomberg A., Durbin R.,
658 Louis E. J., 2009 Population genomics of domestic and wild yeasts. Mol Syst
659 Biol 458: 337–341.
660 Liu F., van der Lijn F., Schurmann C., Zhu G., Chakravarty M. M., Hysi P. G.,
661 Wollstein A., Lao O., de Bruijne M., Ikram M. A., van der Lugt A., Rivadeneira
662 F., Uitterlinden A. G., Hofman A., Niessen W. J., Homuth G., de Zubicaray G.,
663 McMahon K. L., Thompson P. M., Daboul A., Puls R., Hegenscheid K., Bevan
664 L., Pausova Z., Medland S. E., Montgomery G. W., Wright M. J., Wicking C.,
665 Boehringer S., Spector T. D., Paus T., Martin N. G., Biffar R., Kayser M., 2012
666 A Genome-Wide Association Study Identifies Five Loci Influencing Facial
667 Morphology in Europeans. PLoS Genet 8: e1002932.
668 Lorenz K., Cohen B. A., 2014 Causal Variation in Yeast Sporulation Tends to
669 Reside in a Pathway Bottleneck. PLoS Genet 10: e1004634.
670 Magni P., Ferrazzi F., Sacchi L., Bellazzi R., 2008 TimeClust: a clustering tool for
671 gene expression time series. Bioinformatics 24: 430–432.
672 Mai B., Breeden L., 2000 CLN1 and its repression by Xbp1 are important for
673 efficient sporulation in budding yeast. Molecular and Cellular Biology 20: 478–
674 487.
675 Manolio T. A., Collins F. S., Cox N. J., Goldstein D. B., Hindorff L. A., Hunter D. J.,
676 McCarthy M. I., Ramos E. M., Cardon L. R., Chakravarti A., Cho J. H.,
677 Guttmacher A. E., Kong A., Kruglyak L., Mardis E., Rotimi C. N., Slatkin M.,
678 Valle D., Whittemore A. S., Boehnke M., Clark A. G., Eichler E. E., Gibson G.,
679 Haines J. L., Mackay T. F. C., McCarroll S. A., Visscher P. M., 2009 Finding
680 the missing heritability of complex diseases. Nature 461: 747–753.

bioRxiv preprint doi: https://doi.org/10.1101/033167; this version posted November 27, 2015. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

Marzluf G. A., 1997 Genetic regulation of nitrogen metabolism in the fungi. Microbiology and Molecular Biology Reviews 61: 17–32.

Neiman A. M., 2005 Ascospore formation in the yeast Saccharomyces cerevisiae. Microbiology and Molecular Biology Reviews 69: 565–584.

Neiman A. M., 2011 Sporulation in the Budding Yeast Saccharomyces cerevisiae. Genetics 189: 737–765.

Nelson B., Kurischko C., Horecka J., Mody M., Nair P., Pratt L., Zougman A., McBroom L. D. B., Hughes T. R., Boone C., Luca F. C., 2003 RAM: a conserved signaling network that regulates Ace2p transcriptional activity and polarized morphogenesis. Mol. Biol. Cell 14: 3782–3803.

Primig M., Williams R. M., Winzeler E. A., Tevzadze G. G., Conway A. R., Hwang S. Y., Davis R. W., Esposito R. E., 2000 The core meiotic transcriptome in budding yeasts. Nat. Genet. 26: 415–423.

Saint Pierre A., Génin E., 2014 How important are rare variants in common disease? Briefings in Functional Genomics 13: 353–361.

Schadt E. E., Friend S. H., Shaywitz D. A., 2009 A network view of disease and compound screening. Nat Rev Drug Discov 8: 286–295.

Spellman P. T., Sherlock G., Zhang M. Q., Iyer V. R., Anders K., Eisen M. B., Brown P. O., Botstein D., Futcher B., 1998 Comprehensive identification of cell cycle-regulated genes of the yeast Saccharomyces cerevisiae by microarray hybridization. Mol. Biol. Cell 9: 3273–3297.

Stankiewicz P., Lupski J. R., 2010 Structural variation in the human genome and its role in disease. Annu. Rev. Med. 61: 437–455.

Storey J. D., Xiao W., Leek J. T., Tompkins R. G., Davis R. W., 2005 Significance analysis of time course microarray experiments. Proceedings of the National Academy of Sciences 102: 12837–12842.

Tang F., Peng Y., Nau J. J., Kauffman E. J., Weisman L. S., 2006 Vac8p, an armadillo repeat protein, coordinates vacuole inheritance with multiple vacuolar processes. Traffic 7: 1368–1377.

Teixeira M. C., Monteiro P. T., Guerreiro J. F., Gonçalves J. P., Mira N. P., Santos dos S. C., Cabrito T. R., Palma M., Costa C., Francisco A. P., Madeira S. C., Oliveira A. L., Freitas A. T., Sá-Correia I., 2014 The YEASTRACT database: an upgraded information system for the analysis of gene and genomic transcription regulation in Saccharomyces cerevisiae. Nucleic Acids Research 42: D161–6.

Tomar P., Bhatia A., Ramdas S., Diao L., Bhanot G., Sinha H., 2013 Sporulation Genes Associated with Sporulation Efficiency in Natural Isolates of Yeast. PLoS ONE 8: e69765.

Turcotte B., Liang X. B., Robert F., Soontorngun N., 2009 Transcriptional regulation of nonfermentable carbon utilization in budding yeast. FEMS Yeast Res 10: 2–13.

Wilkening S., Lin G., Fritsch E. S., Tekkedil M. M., Anders S., Kuehn R., Nguyen M., Aiyar R. S., Proctor M., Sakhanenko N. A., Galas D. J., Gagneur J., Deutschbauer A., Steinmetz L. M., 2014 An evaluation of high-throughput approaches to QTL mapping in Saccharomyces cerevisiae. Genetics 196: 853–865.

Xu Z., Wei W., Gagneur J., Perocchi F., Clauder-Münster S., Camblong J., Guffanti E., Stutz F., Huber W., Steinmetz L. M., 2009 Bidirectional promoters generate pervasive transcription in yeast. Nature 457: 1033–1037.

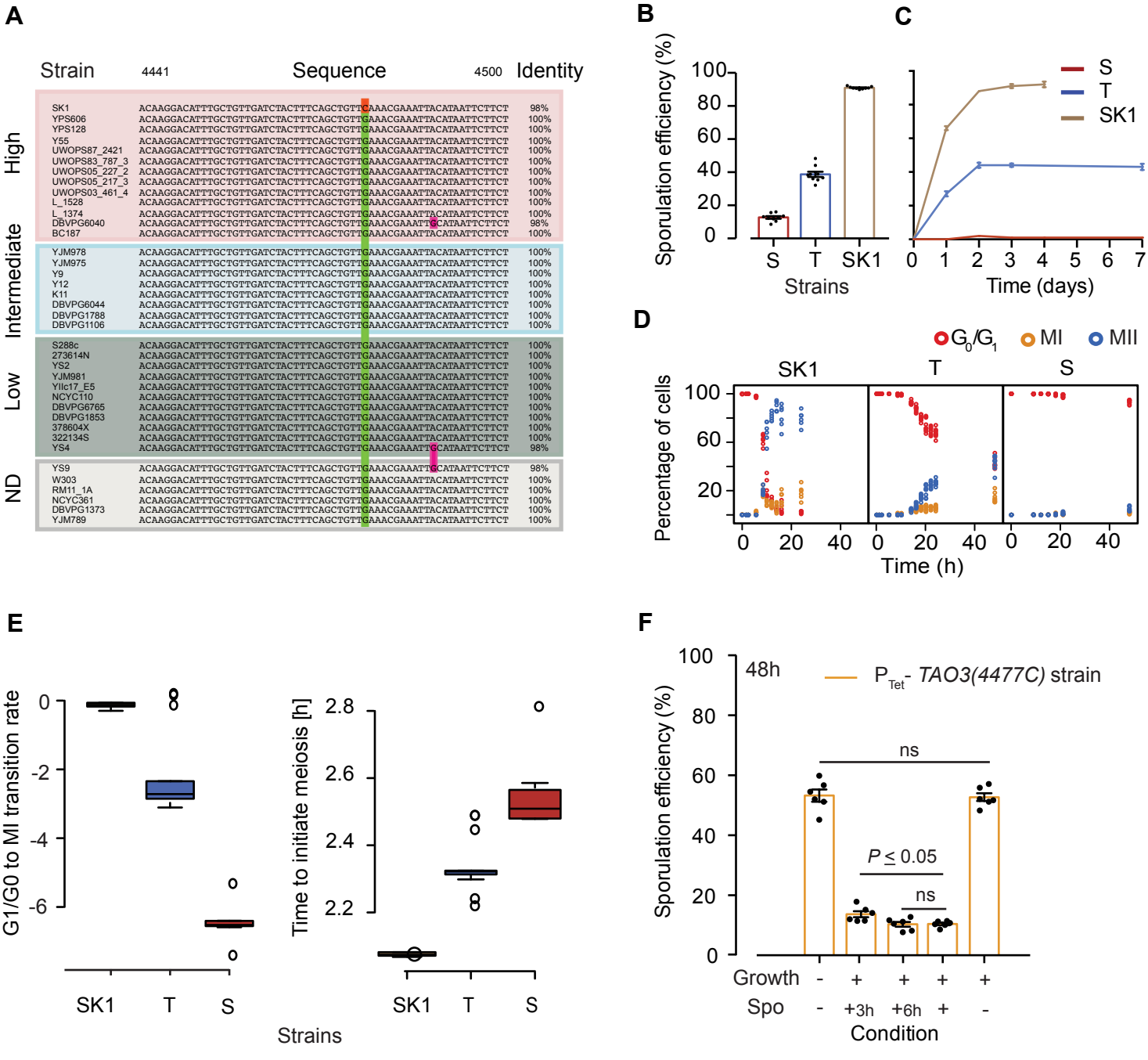
Zuk O., Schaffner S. F., Samocha K., Do R., Hechter E., Kathiresan S., Daly M. J., Neale B. M., Sunyaev S. R., Lander E. S., 2014 Searching for missing

bioRxiv preprint doi: https://doi.org/10.1101/033167; this version posted November 27, 2015. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

731 heritability: designing rare variant association studies. Proc. Natl. Acad. Sci.
732 U.S.A. 111: E455–64.

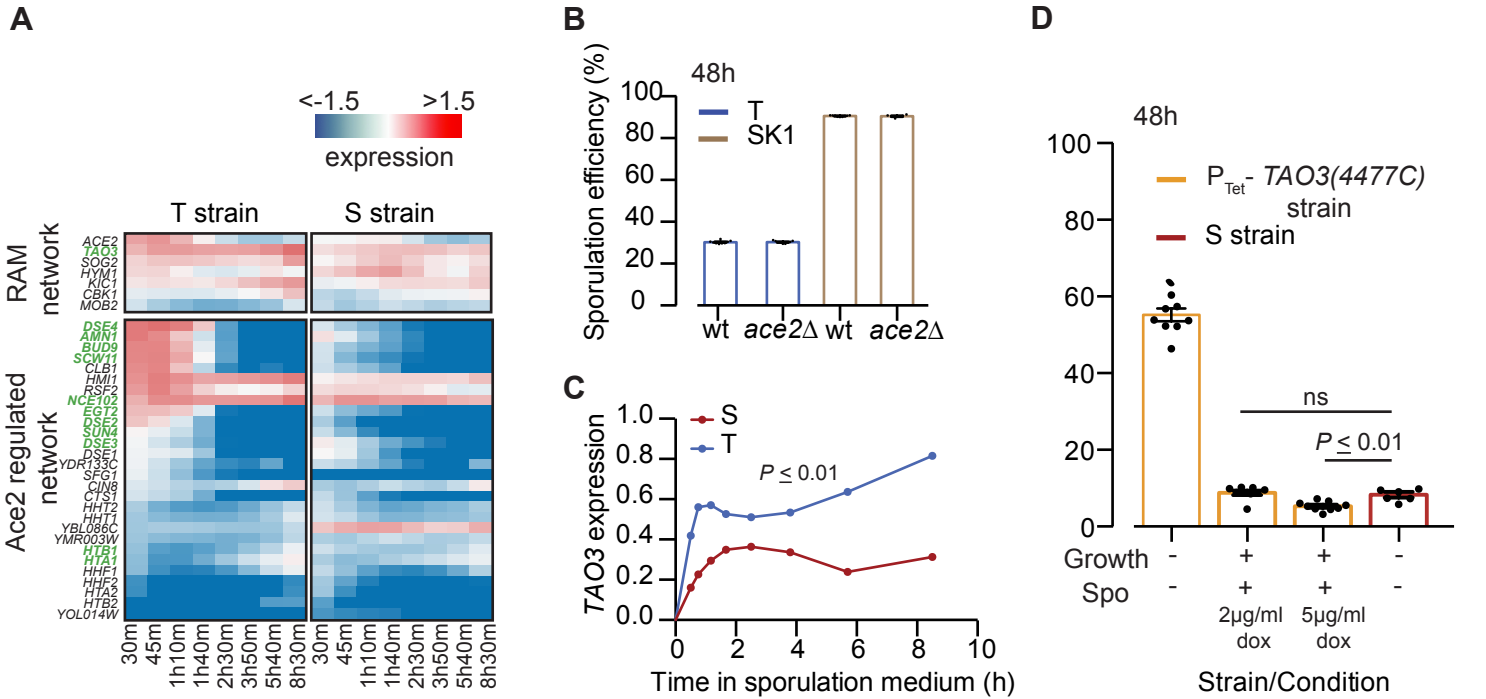
bioRxiv preprint doi: https://doi.org/10.1101/033167; this version posted November 27, 2015. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

Gupta_*et al*_Figure 1



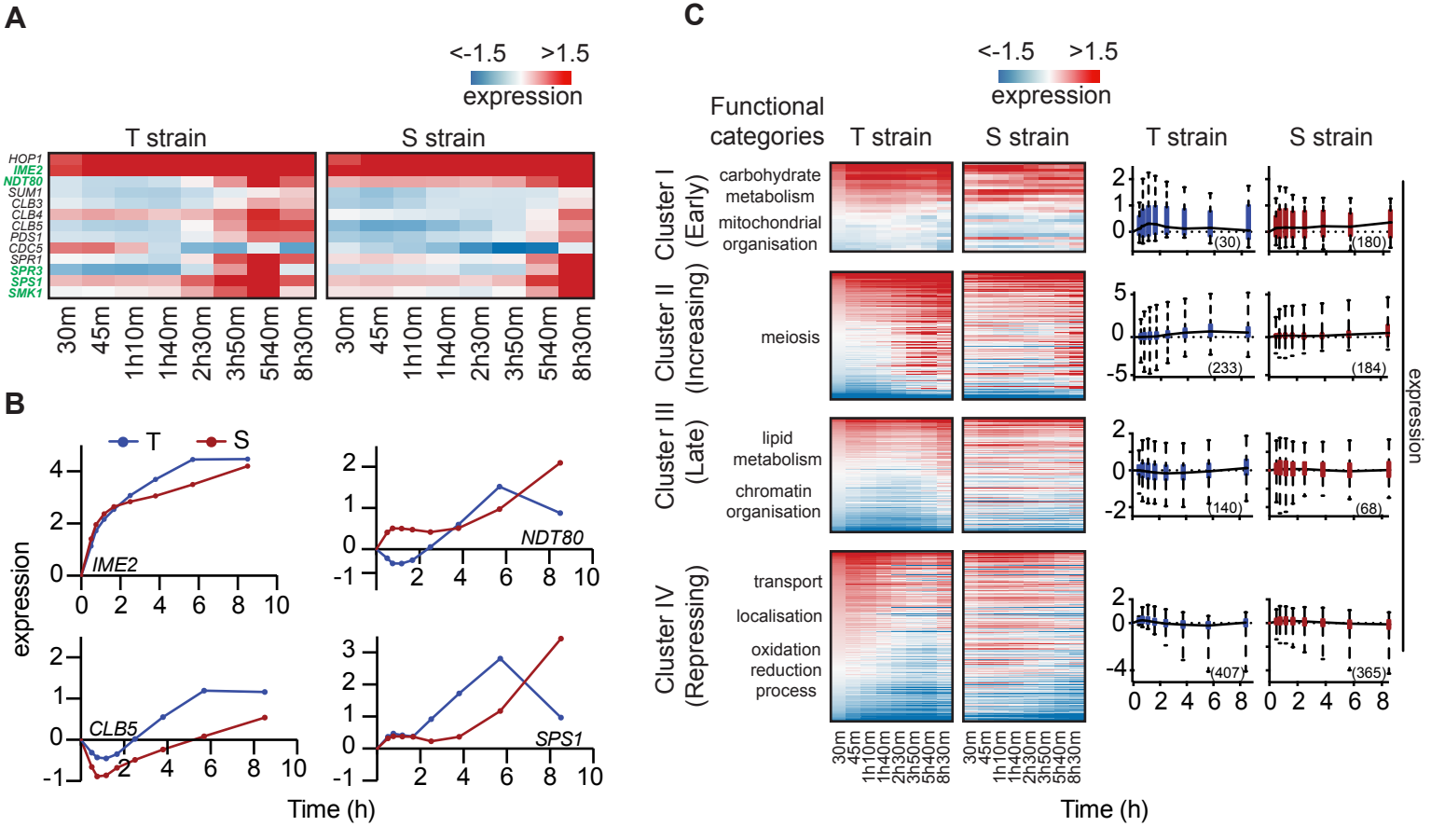
bioRxiv preprint doi: https://doi.org/10.1101/033167; this version posted November 27, 2015. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

Gupta_*et_al*_Figure 2



bioRxiv preprint doi: https://doi.org/10.1101/033167; this version posted November 27, 2015. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

Gupta_*et al*_Figure 3



bioRxiv preprint doi: https://doi.org/10.1101/033167; this version posted November 27, 2015. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

Gupta_*et_al*_Figure 4