Uncovering divergence in gene expression regulation in the adaptation of yeast to nitrogen scarcity

Carlos A. Villarroel[a,b,c], Macarena Bastías[d], Paulo Canessa[a,d], Francisco A. Cubillos[a,b]

[a] ANID - Programa Iniciativa Científica Milenio - Instituto Milenio de Biología Integrativa (iBio), Santiago, Chile.

[b] Universidad de Santiago de Chile, Facultad de Química y Biología, Departamento de Biología, Santiago, Chile.

[c] Laboratorio Interacciones Insecto-Planta, Instituto de Ciencias Biológicas, Universidad de Talca, 1141 Talca, Chile

[d] Centro de Biotecnología Vegetal, Facultad de Ciencias de la Vida, Universidad Andres Bello, Santiago, RM 837-0146, Chile

Running head: Allele-specific gene expression regulation in yeast

#Address correspondence to Francisco A. Cubillos francisco.cubillos.r@usach.cl

Abstract word count = 288

Main text word count = 4855

Abstract

Saccharomyces cerevisiae rewires its transcriptional output to survive stressful environments, such as nitrogen scarcity under fermentative conditions. Although divergence in nitrogen metabolism has been described among natural yeast populations, the impact of regulatory genetic variants modulating gene expression and nitrogen consumption remains to be investigated. Here, we employed an F1 hybrid from two contrasting *S. cerevisiae* strains, providing a controlled genetic environment to map *cis* factors involved in the divergence of gene expression regulation in response to nitrogen scarcity. We used a dual approach to obtain genome-wide allele-specific profiles of chromatin accessibility, transcription factor binding, and gene expression through ATAC-seq and RNA-seq. We observed large variability in allele-specific expression and accessibility between the two genetic backgrounds, with a third of these differences specific to a deficient nitrogen environment. Furthermore, we discovered events of allelic bias in gene expression correlating with allelic bias in transcription factor binding solely under nitrogen scarcity, where the majority of these transcription factors orchestrates the Nitrogen Catabolite Repression regulatory pathway and demonstrates a *cis* x environment-specific response. Our approach allowed us to find *cis* variants modulating gene expression, chromatin accessibility and allelic differences in transcription factor binding in response to low nitrogen culture conditions.

Importance

Historically, coding variants were prioritized when searching for causal mechanisms driving adaptation of natural populations to stressful environments. However, the recent focus on non-coding variants demonstrated their ubiquitous role in adaptation. Here we performed genome-wide regulatory variation profiles between two divergent yeast strains when facing nitrogen nutritional stress. The open chromatin availability of several regulatory regions changes in response to nitrogen scarcity. Importantly we describe regulatory events that deviate between strains. Our results demonstrate a widespread variation in gene expression regulation between naturally occurring populations in response to stressful environments.

Keywords

Yeast, allele-specific, ATAC-seq, nitrogen, regulatory divergence, wine

Introduction

Uncovering the molecular configurations that underlie gene expression divergence in adaptation to stressful environments constitutes a relevant genetic quest. The yeast *Saccharomyces cerevisiae* provides an excellent genetic model to investigate the link between the regulatory divergence of sequences and environmental fluctuations (1). Yeast cells undergo extensive reprogramming of their gene expression profiles to withstand different environmental stresses; of these, the transcriptional response to nitrogen scarcity has been comprehensively described (2-5). Yeast fitness strongly depends on the availability of preferred nitrogen sources, and changes in such sources trigger the immediate transcriptional rewiring of nitrogen metabolism. In yeast, at least four pathways are involved in regulating nitrogen metabolism, with the Nitrogen Catabolite Repression (NCR) pathway (6) being the main orchestrator of the response to nitrogen starvation. Importantly, extensive natural variation in nitrogen consumption (7, 8) and starvation tolerance under wine fermentative conditions (3) have been found among yeast populations, demonstrating significant divergence in the regulatory mechanisms involved in the NCR pathway (9).

Yeast strains have different nitrogen consumption profiles, amino acid preferences, and tolerance to nitrogen scarcity during wine fermentation (7, 10). In this way, winemaking strains exhibit physiological adjustments to poor nitrogen environments while displaying good fermentation performance, a feature which is a hallmark of domestication (10-13). Using Quantitative Trait Loci (QTLs) approaches, several genes involved in differences in nitrogen consumption have been mapped (3, 7, 9, 14), including important nodes of the NCR pathway such as *GTR1* (15) and *RIM15* (16). Nevertheless, the regulation of gene expression and the modifying role of polymorphic transcription factors (TFs) in response to nitrogen scarcity in wine strains remains to be elucidated.

Transcriptional divergence originates from genetic variants, which can be identified through mapping of expression QTLs (1, 17). eQTLs might regulate the adjacent allele (*cis* eQTL), or affect one or multiple distant genes (*trans* eQTL). First-generation (F1) hybrids constructed from individuals of divergent lineages offer a refined approach to map *cis* factors responsible for expression divergence (18-21). In this F1 hybrid setup, the *trans* component is neglected as *trans* eQTLs affect both parental alleles in the same way,

therefore cancelling potentially different contributions. On the other hand, *cis* effects will remain allele-specific (22). Moreover, differences in the expression of each allele (Allele Specific Expression, ASE) are explained by the allele's local variants, which might control the physical accessibility of its promoter or regulatory region (23). This could be achieved by modulating the affinity of TF binding sites or affecting the regulation of the encoded RNA at a post-transcriptional level (24-27). Numerous studies have extensively quantified ASE in different model organisms (28-32). However, most of these studies have not incorporated a genome-wide experimental approach that assesses the *cis*-regulatory mechanisms underlying allelic expression variation. Coupling massive mRNA sequencing with assays that cut DNA *in vivo* at physically accessible chromatin regions, such as ATAC-seq (Assay for Transposase Accessible Chromatin), can portray a whole-genome profile of DNA accessibility to transcriptional regulators (33). In addition, ATAC-seq can also provide a genome-wide survey of transcription factor binding (TFB), allowing the *in silico* footprinting of TFB at open chromatin regions (34). ATAC-seq has been employed in yeast to investigate regulatory mechanisms driving aging (35), metabolism and cell division (36), pathogenesis (37), and cold-adaptation in an interspecies hybrid (20). Recent studies in mouse crosses have incorporated assays that profile allelic differences at the transcriptional regulatory level (38-40), demonstrating the suitability of these techniques to measure allelic imbalance in the regulation of gene expression.

By coupling RNA-seq and ATAC-seq , herein we measure allelic imbalance occurring at the level of gene expression, chromatin accessibility, and TF binding in an F1 hybrid between two divergent *S. cerevisiae* strains. We evaluated whether differences in nitrogen consumption between a winemaking strain (DBVPG6765), and an un-domesticated strain (YPS128) isolated from an oak tree were due to *cis*-regulatory variants modulated through environments differing in nitrogen availability. We report numerous events of allelic differences in chromatin accessibility between these two strains, remarkably few of which directly correlate with ASE. Furthermore, we show that one third of the allelic differences in gene expression and accessibility only occur under low nitrogen. By performing allele-specific TFB footprinting, we reveal TFs that potentially drive allelic expression differences, some of which have not been previously related to the regulation of nitrogen metabolism.

Results

A differential response to nitrogen scarcity is observed in two divergent yeast strains

To identify *cis*-regulatory variants driving gene expression divergence in the adaptation of yeast to nitrogen scarcity, we performed Allele-Specific Expression and Accessibility (ASE and ASA, respectively) employing RNA-seq and ATAC-seq (**Fig. 1**). For this, we used the WE x NA F1 hybrid grown under low or excess nitrogen concentrations (SM60 and SM300, respectively; see methods).

First, we assessed nitrogen consumption kinetics in the WE x NA hybrid and the parental strains in both fermentation conditions. We sampled fermentations at an early time point, i.e. 14 hours after synthetic grape must inoculation. In SM300, the parental NA strain exhibited lower consumption levels of total YAN than the hybrid and WE strains (p-value < 1e-04, ANOVA) **(Fig. 2A)**. Although, we did not find higher total YAN consumption in the hybrid, after examining each nitrogen source we found that the hybrid strain consumed higher levels of serine and alanine in SM300 compared to both parental strains (p-value < 0.05, ANOVA) **(Fig. 2B)** . We formally estimated heterosis and found that the consumption of six amino-acids under SM300 had a heterosis coefficient higher than 1 **(Fig. 2C)**. Among those aminoacids that are rapidly consumed in the hybrid, we found all amino-acids transported by Agp1p/Gnp1p (Serine, Threonine, Glutamine) (**Fig. 2C**). In contrast to SM300, under low nitrogen conditions YAN was almost depleted from the media after 14 hours of fermentation, particularly by the WE and WE x NA strains, with the NA strain exhibiting lower total YAN consumption (p-value < 0.05, ANOVA, **Fig. 2A**). In addition, we found differences in the consumption kinetics for 14 amino acids when comparing the hybrid and the two parental strains (p-value < 0.05, ANOVA), together with heterosis in valine consumption (**Fig. 2BC**). After 14 days of fermentation, the three genetic backgrounds showed no differences in total CO_2 loss under SM300 (**Fig. 2D**), though under low nitrogen conditions the NA strain had the lowest total CO_2 loss (p-value WE - NA = 0.01; p-value WE x NA – NA = 0.09, ANOVA), which indicates a stronger negative effect of nitrogen scarcity on the fermentation performance of the NA strain and nitrogen

152 starvation stress (10, 41). Furthermore, low nitrogen affected all strains' fermentation
153 kinetics compared to SM300, diminishing the maximum fermentation speed by 63%, 58%,
154 and 60% in NA, WE, and WE x NA strains, respectively (**Fig. 2E**). These results
155 demonstrate that the genotype affects fermentation performance during nitrogen scarcity,
156 and suggest a dominant inheritance of efficient nitrogen consumption in yeast.

157

158 **Low correlation between gene expression and chromatin accessibility in response**
159 **to fermentation under low nitrogen**

160 We collected mRNA from the WE x NA hybrid after 14 hours of fermentation under
161 SM300 and SM60. At this time point we observed primarily consumption of preferred
162 nitrogen sources under SM300 (NCR suppressed state), and complete consumption of
163 YAN under SM60 triggering nitrogen starvation stress (NCR active state) (**Figure 2B**) (10,
164 41). We found a total of 3,719 DEGs (Differentially Expressed Genes) between conditions,
165 of which 1,842 and 1,877 were upregulated or downregulated respectively in response to
166 low nitrogen (FDR < 0.05, **Table S1a**). Among the upregulated genes, we found 69 DEGs
167 previously classified as NCR sensitive (42). Interestingly 21 of these NCR sensitive genes
168 were determined among the top 30 genes that were most-induced by low nitrogen (**Table**
169 **S1b**). Enriched biological processes among low nitrogen induced genes were related to
170 transport, energy generation, detoxification, and oxido-reduction (**Table S2a**), while genes
171 associated with ribosomal biogenesis were significantly enriched among downregulated
172 genes (**Table S2b**).

173 To profile the chromatin accessibility landscape in response to low nitrogen, we performed
174 ATAC-seq in hybrid replicates collected in the aforementioned culture conditions (See **Fig.**
175 **S1** for ATAC-seq QC plots). First, we assessed the correlation between gene expression
176 and chromatin accessibility. We found a significant, though moderate correlation
177 (Spearman R=0.28 and R=0.30 for SM300 and SM60 respectively, p-value < 1 e-99),
178 between the amount of gene expression (RPKM) and the corresponding ATAC-seq signal
179 up to 1,000 bp upstream of the transcript start site (TSS). We further examined this
180 correlation by performing a window analysis between ATAC-seq coverage and total gene
181 expression (RPKM) on the upstream and downstream TSS region of every *S. cerevisiae*
182 gene (**Fig. 3A**). These results showed that the highest and most significant correlation
183 between gene expression of a given gene and its nearest ATAC-seq signal differed

between nitrogen levels. In particular, for SM300 we found a narrower signal region upstream of the TSS (-350 to -50 bp, Spearman R = 0.32, p-value < 1 e-135), while we estimated a wider region for SM60 (-400 to +50 bp, Spearman R = 0.33, p-value < 1 e-140) (white squares in **Fig. 3A**). These results show that different segments within the gene's regulatory regions respond to changing nitrogen concentrations. For the analyses shown hereafter, to evaluate the corresponding ATAC-seq signal we considered a regulatory region of 400 bp upstream of the TSS of each gene.

By analyzing 5,625 regulatory regions, we found an increase in chromatin accessibility in 376 differentially accessible regions (DARs) and a decrease in 875 DARs (FDR < 0.05) under SM60, representing 22.2% of the analyzed regions (**Fig. 3B**, **Table S1**). Generally, chromatin accessibility fold changes between conditions were lower in absolute magnitude for DARs (mean | Log2FC | = 0.38) than those expression fold changes observed for DEGs (mean | Log2FC | = 0.95) (p-value = 0.001, Mann-Whitney-Wilcoxon test, **Fig. 3C**). We found that regions regulating genes involved in glucose transport and urea metabolism were less accessible under low nitrogen conditions (**Table S2c**), while regions regulating genes involved in the metabolism of non-preferred nitrogen sources were more accessible when subjected to nitrogen scarcity (**Table S2d**).

We inspected whether differential gene expression between conditions correlated with chromatin accessibility. Despite the positive correlation between total gene expression and upstream ATAC-seq signals (**Fig. 3A**), we found a small overlap between DEGs and differential open chromatin (**Fig. 3B),** which was in agreement with the absence of correlation when comparing log2 fold-changes of chromatin accessibility and gene expression in response to low nitrogen (Pearson R = 0.015). Still, 444 DEGs and DARs showed downregulation and lower accessibility, and conversely, an induction in transcript levels with higher chromatin accesibility. Downregulated genes within this positively correlated set were enriched in processes related with cytoplasmic translation and vitamin biosynthesis, while upregulated genes were enriched for catabolism of allantoin and glyoxylate (**Fig. 3D**). In contrast, a large number of DEGs related with ribosome biogenesis showed no differences in accessibility at their regulatory regions across nitrogen conditions, despite being concertedly less expressed in response to low nitrogen (**Fig. S2A**). In addition, among non-DEGs, we found differential accessibility in genes related with sugar transport (**Fig. S2B**). These results indicate that approximately 12% of gene expression differences in the yeast genome positively correlated with chromatin

accessibility differences between conditions. We would like to highlight that while RNA-seq and ATAC-seq experiments were performed independently, although carefully sampling after the same period of time. Still the absence of correlation between gene expression fold changes and chromatin accessibility has been previously observed in yeast (20, 43). In summary, chromatin accessibility results complement those obtained by RNA-seq by providing novel evidence of regulatory rewiring under contrasting nitrogen conditions.

Allele-specific expression and chromatin accessibility under excess and low nitrogen conditions

We compared the transcriptional response to low nitrogen of each parental genome within a shared (WE x NA) *trans* environment by splitting RNA-seq alignments with SNP-informative reads (see Materials and Methods). We found that parental genomes display a highly similar transcriptional response to low nitrogen (Pearson R = 0.911, p-value < 2.2e-16), although differences in ASE could still be found. Therefore, we investigated ASE by performing a binomial test of read counts at 21,647 SNPs in 3,923 genes (1732 genes did not have SNPs or only had SNPs with fewer than 10 counts). Differential ASE was determined in 543 genes, of which 147 and 245 were solely found in SM300 and SM60, respectively, while 153 were significant under both conditions (**Fig. 4AB, Table S3**), the majority of which (152) maintained the ASE direction, i.e. bias in expression favored the same genotype in both conditions. Allelic ratios in both conditions were correlated (Pearson R = 0.69, P value < 1e-15), although they were lower than those described in other ASE assays in yeast (21, 44). Accordingly, a significant portion of ASE differences in yeast under contrasting nitrogen conditions are dependent on the *cis*-genotype x environment interaction, rather than primordially on the *cis*-genotype configuration.

Further dissection of the ASE of strongly upregulated genes under low nitrogen conditions (FC > 3, 251 DEGs) allowed us to identify key variants involved in NCR. For example, we found high WE allelic expression compared to the NA variant for the amino-acid permeases *GAP1* and *PUT4*, and for the ammonium permease *MEP2* (**Fig. 4B**). In contrast, the NA allele encoding for the NCR-sensitive allantoate transporter (*DAL5*) showed a strong bias in expression when compared to the WE allele. These results indicate that genetic variants in ammonium and specific amino-acid permeases are

overexpressed in wine strains when facing nitrogen scarcity, putatively due to *cis*-regulatory variants.

To assess these differences, allele-specific accessibility (ASA) was determined in each parental genome in response to low nitrogen within the shared *trans* environment (**Fig. 1C**). Interestingly we found a moderate correlation between the chromatin accessibility response of both parental genomes (Pearson R= 0.529, p-value < 1e-15), which was lower than that found for the gene expression response, suggesting a greater impact of environmentally-dependent differences in *cis* regulation acting on chromatin accessibility than on gene expression. We tested allelic imbalance in chromatin accessibility (binomial test), and were able to quantify ASA in 15,333 SNPs belonging to 4,822 regulatory regions. A total of 252 regions exhibited differential ASA, and from this, 113 and 69 regulatory regions showed ASA under either SM60 or SM300, respectively (**Table S3**). ASA was moderately correlated between nitrogen conditions (Pearson R = 0.61, P value < 1e-15), similar to what was found for ASE. Importantly, all regions that showed ASA in both conditions (70 in total) maintained the imbalance direction, i.e. bias in accessibility favored the same genotype in both conditions (**Fig. 4C**). Among enriched biological processes in regions that showed ASA in SM60, we found more accessible regulatory regions associated to the expression of aldehyde metabolic process genes and the response to oxidative stress in the WE and NA genomes respectively (**Table S4cd**).

Taking the differential ASE and ASA datasets together, we assessed the intersection between them in each condition. Under nitrogen excess, 13 genes displaying ASE also showed ASA in the same direction, while 5 genes showed ASA in the opposite direction **(Fig. 4D)**. Moreover, under low nitrogen, 28 genes showed ASE and ASA in the same direction, while 10 displayed ASA and ASE in opposite directions **(Fig. 4E)**. Overall, these results indicate a convergence between ASA and ASE in 41 genes, where open chromatin alleles had greater expression levels. However, only 5% and 9% of genes in ASE coincided with allelic imbalance in accessibility occurring in the same direction in SM300 and SM60, respectively, while the large majority of genes exhibiting ASE did not show ASA levels and viceversa. These results suggest that additional regulatory mechanisms, such as differences in transcription factor binding (TFB), could influence differences in allelic expression.

Allelic imbalance in transcription factor binding reveal environment-dependent regulatory mechanisms driving gene expression divergence

To explore the link between TFB and allele specific expression differences occurring in the WE x NA hybrid, we used ATAC-seq (Tn5) cut sites to infer TFB and generate genome-wide binding scores for 141 TFs. First, we calculated TFB scores inferred from cut sites from all WE x NA ATAC-seq alignments, finding 28,842 motifs likely bound by a TF in SM300 and/or SM60. To assess the correlation between TFB, gene expression and chromatin accessibility, we inspected fold changes of TFB scores in the hybrid across conditions, discovering a strong correlation with ATAC-seq fold-changes (Pearson R = 0.71, p-value < 1e-130) (**Fig. 5A**). However, no overall correlation between fold changes of TFB scores and gene expression differences between conditions was found (Pearson R=0.072, p-value 1.58e-31) (**Fig. 5A**), suggesting a complex interplay between global TFB and gene expression, while also evidencing a direct effect of chromatin accessibility on TFB scores. We re-examined the correlation between TFB scores and gene expression, but this time individually for each TF. In this way we found a significant correlation for 11 out of 141 TFs (Pearson R > 0.2, FDR < 0.1) **(Table S5a)**. The Cst6p TF showed the highest correlation between TFB and gene expression (Pearson R = 0.36, FDR = 0.03) **(Fig. 5B**). Cst6p encodes for a basic leucine zipper TF involved in the stress response (45), and our results suggest a role for Cst6p in response to low nitrogen conditions. Moreover, for all TFs tested, the binding scores were highly correlated with their ATAC-seq fold changes, ranging from Dal80p (Pearson R = 0.882, FDR = 3.08 e-25) to Fkh2p (Pearson R = 0.50, FDR= 3.14 e-6) (**Table S5a**) .

By utilizing ATAC-seq data from a hybrid strain, we were able to obtain allele specific TFB scores by splitting SNP-informative reads in the hybrid alignments (**Fig. 1E**). This allowed us to infer allele-specific binding (ASB) by estimating the TFB scores for each parental strain. We calculated ASB at 27,370 predicted binding motifs, which were on average at a distance of 23.2 bp to the nearest SNP, with 4,119 motifs (16%) overlapping with a SNP. We found a significant correlation between ASB and ASA (for SM60 Pearson R = 0.53, p-value < 1e-130). However, no genome-wide correlation was found between ASB and ASE in any condition (for SM60 Pearson R = 0.09, p-value = 4.36e-47). As previously done with the hybrid TFB scores, we also decided to inspect the correlation between ASB and ASE for each TF individually **(Table S5b)**. Under SM60, we observed that the allele-specific binding for six TFs correlated with ASE of their target genes (Pearson R > 0.2, FDR <

313 0.05). Among them, we found three out of the four GATA-type zinc finger TFs that
314 participate in NCR regulation; i.e. Gat1p, Dal80p, and Gzf3p (**Fig. 5C, Table S5b**).
315 Together with Gln3p, these GATA TFs share very similar binding motifs (shown in **Fig.**
316 **5C**), which suggest that these correlations might not be specific for any of these TFs in
317 particular. In the case of Gln3p, a shorter and less informative GATA motif was used,
318 which substantially increased the number of predicted binding sites compared to those of
319 the other GATA factors, suggesting a stronger influence of false positive binding sites on
320 the lack of correlation between Gln3p ASB and the ASE of its target genes under SM60
321 (Pearson R = 0.09). We crossreferenced our Dal80p *in silico* binding data under SM60
322 with a Dal80p Chip-seq dataset performed under a similar stress condition (42), and found
323 agreement for Dal80p binding at 36 (53%) of 67 predicted bound promoters.

324 In addition to GATA factors, we found that Skn7p, Swi4p, Tos8p, Yap5p, Tod6p, and
325 Yox1p TFs had a significant correlation between ASB and ASE under SM60 **(Table S5b)**.
326 Under the SM300 culture condition, we found a significant correlation between ASE and
327 ASB solely for Yap5p and Swi4p (Pearson R > 0.2, FDR < 0.1), and specifically for this
328 condition for Yap7p **(Table S5b)**. These results suggest a role for TFs that have not
329 previously been associated with nitrogen metabolism in the adaptation of winemaking
330 strains to nitrogen scarcity in grape must. Nevertheless, we expect that the significant
331 correlations found between ASB and ASE were confounded by the allele-specific ATAC-
332 seq signal, which prompted us to evaluate our ASB data in more detail.

333 We tested for significant differences in ASB (see Materials and Methods) and found 2,254
334 and 1,161 binding sites having differential ASB (out of 27,370 predicted binding motifs) in
335 SM60 and SM300 respectively (FDR < 0.1) **(Table S5cd)**. Only 22 % (623) of these
336 binding sites were differentially bound in both conditions **(Fig. 5D)**. Furthermore, 513
337 (22%) and 293 (25%) of the binding sites displaying ASB overlapped with a SNP in SM60
338 and SM300, respectively. This represents a higher proportion than that observed for all
339 tested binding sites (16%), suggesting that binding sites co-occurring with SNPs were
340 more likely to be differentially bound (Chi-square test, p-value = 7.3e-18).

341 Next, we wanted to identify TFs that are differentially bound to the WE or NA allele that
342 might drive ASE in absence of chromatin accessibility differences. Hence, we focused our
343 ASB analysis on those regulatory regions in which we did not find ASA but their genes
344 showed ASE. In SM60, of the 311 genes that showed ASE (but not ASA), we found 55
345 regulatory regions displaying ASB, representing 17% of the ASE differences (**Fig. 5E**, FDR

346 < 0.1). Also, 247 genes showed ASE but not ASA in SM300; of those, 30 showed ASB at
347 their regulatory regions (**Fig. 5E**, FDR < 0.1). Moreover, many of the binding sites showing
348 ASB did not alter allelic expression **(Fig. 5E)**. These results demonstrate the additional
349 contribution of ASB in absence of ASA towards differences in allelic expresion under low
350 nitrogen conditions.

351

352

353 **Identification of *cis*-acting variants driving allelic expression differences in**
354 **response to nitrogen scarcity**

355 We inspected in detail those DEGs whose expression was upregulated or
356 downregulated by low nitrogen in the hybrid context, which also exhibited ASB and ASE
357 but not ASA. We observed ASB in 23 and 13 regulatory regions, respectively, in DEGs
358 induced or repressed under low nitrogen (**Fig. 6A induced in SM60, Fig. S3 repressed in**
359 **SM60**). As an example of a binding site co-occurring with a SNP and displaying ASB, we
360 show the TFB site for the transcriptional repressor Mot2p localized at the regulatory region
361 of the *HXK1* gene, which encodes a hexokinase that phosphorylates hexoses for
362 subsequent glycolysis **(Fig. 6B)**. We found a 1.2 fold higher binding for the WE allele than
363 for the NA allele (FDR 0.08), consistent with higher expression of the *HXK1*-NA allelic
364 variant under low nitrogen. The causal variant could be a SNP (A > G) affecting the
365 binding motif of Mot2p in the NA genome. Another example is shown for a differentially
366 bound region in the *STF1* promoter, at which the binding sites of Gln3p and Pho2p co-
367 localize with a SNP that might drive allelic differences in expression found for the *STF1*
368 gene (**Fig. 6C**). Among downregulated DEGs, we observed differential binding at the
369 motifs for the TFs Msn2p, Msn4p, and Rgm1p present in the *THI4* promoter (**Fig. S3B**).

370 As an example of a binding site that is relatively far from a putative causal SNP, but that
371 exhibits large allelic binding differences, we found in the *MEP2* regulatory region (which
372 encodes for a ammonium permease) two sequentially occurring GATA-like sites for Gln3p,
373 both of which showed higher binding scores for the WE allele than for the NA variant (1.93
374 fold higher in WE under SM60, FDR < 0.003, **Fig. 6D**), coincident with the higher
375 expression of the *MEP2*-WE allele under SM60. Importantly, *MEP2* was highly induced
376 under low nitrogen, and the allelic differences in binding and expression were only
377 significant under SM60.

In summary, our results demonstrate that the specific identification of allele specific TFB, together with differences in chromatin accessibility, can shed light onto novel molecular targets and mechanisms driving phenotypic differences between yeast strains.

Discussion

In this work, we describe genome-wide allelic imbalance events at three levels - gene expression, chromatin accessibility, and transcription factor binding - which allowed us to expose *cis* mechanisms driving the adaptation to low nitrogen fermentation in winemaking yeast. The transcriptional response of the WE x NA hybrid to low nitrogen resembled that of other studies of yeast under nitrogen stress (10, 42, 46, 47). However, the numerous differences found in allele specific expression (ASE) and accessibility (ASA) suggest widespread variation in *cis* mechanisms regulating nitrogen metabolism among natural yeast populations. We found large differences in the profiles depicted by ATAC-seq and RNA-seq in response to nitrogen scarcity, even though both strategies provided similarly-enriched functional annotations among the concertedly regulated genes. We hypothesize that differences in accessibility translate into differences in gene expression at distinct times depending on the pathway involved. For example, we found increased accessibility and expression for genes associated with metabolism of non-preferred nitrogen sources (e.g. allantoin and carnitine). In contrast, we found less accessibility in six genes encoding hexose transporters without the involvment of differential gene expression, which could still have occurred at an earlier or later time-point during fermentation. Several hexose transporters are downregulated under nitrogen starvation during wine fermentation (48), with the exception of *HXT5* (49), for which we also report increased accessibility and expression under low nitrogen. In addition, we found that several ribosome biogenesis genes were less expressed under low nitrogen, but this response was absent when we inspected accessibility at their promoters, suggesting that this group of genes is regulated by mechanisms other than chromatin organization. In fact, the gene expression output that we measured with RNA-seq (mainly processed mRNA) can be influenced by mechanisms other than chromatin accessibility, including

transcription rate (50), mRNA turnover (51), and by regulatory elements such as TFs and non-coding RNAs (32).

Nitrogen availability is essential for complete wine fermentation, and nitrogen scarcity affects yeast biomass (52), fermentation performance and time to complete fermentation (3, 53). The domestication process selected wine yeasts to withstand nitrogen scarcity stress, while still maintaining good fermentation performance (12). For instance, ammonium is an excellent nitrogen source for yeast growth, and is rapidly consumed at early stages, but only if the concentration of other preferred nitrogen sources such as glutamine is low (41). Here, we found that the ammonium permease *MEP2* was highly induced under low nitrogen while also showing significant allelic bias favoring the WE allele. Our findings indicate that differential binding of Gln3p (or other GATA-like TFs) at the *MEP2* promoter might explain differences in allelic expression in the absence of differential chromatin accessibility, in agreement with Gln3p being a crucial regulator of *MEP2* expression (54). A similar case was found for the non-preferred nitrogen source proline, in which the transporter encoded by the *PUT4* gene was found in allelic bias favoring the WE allele, again in the absence of differences in chromatin accessibility. This finding is interesting since proline cannot be assimilated under oxygen-deprived conditions (55). Indeed, oxygenation has a significant effect upon wine fermentation, accelerating the fermentation rate and impacting the production of volatile compounds (56). Additionally, we found that the allantoate permease *DAL5* has a strong allelic bias favoring the NA strain. Allantoate is absent in wine must (57), and our results point to *cis* regulation orchestrating low priority uptake of allantoate in the WE background. Summarizing, the expression of genes involved in nitrogen transport was frequently found in allelic bias, in particular when cells suffered nitrogen scarcity. The expression of these transporters is mainly controlled by two pathways, the NCR and the SPS (Ssy1-Ptr3-Ssy5) sensor system, which are differentially activated depending on nitrogen availability. In wine fermentation, the SPS pathway maximizes the uptake of preferred nitrogen sources by inducing the expression of their specific permeases (41). On the other hand, the NCR pathway represses generic permeases and those involved in internalizing poor nitrogen sources, but under nitrogen insufficiency such repression is released by NCR deactivation (41). To highlight the involvement of these two pathways in the allelic differences in expression and regulation found among genes encoding for nitrogen-compound permeases, we show a summary of our findings in (**Fig. 7**).

Our findings indicate an important participation of GATA TFs in *cis*-regulatory divergence driving physiological differences under low nitrogen fermentation. Regions containing GATA motifs were more likely to have higher allele accessibility and/or allelic binding, an observation that correlates into higher allelic expression. Importantly, this link between imbalance in allelic expression and regulation was only significant under nitrogen stress, highlighting the role of environmental fluctuations on *cis*-acting causal variants driving eQTLs. In addition, we identified several binding motifs affected by variants that might constitute causal polymorphisms driving differential allelic expression. Certainly, allele-specific binding data obtained from digital TF footprinting serves to identify chromatin regions likely containing causal variants driving differences in allelic expression. These candidate regions that contain *cis*-regulatory variants could be subjected to experimental validation, such as precise genome editing or allele swap, to validate their role in determining phenotypic differences.

In conclusion, we demonstrate that joint allele-specific profiling of chromatin accessibility and gene expression of a divergent yeast cross unveil regulatory dynamics driving a considerable portion of transcriptome divergence. Our findings determine the contribution of chromatin organization towards allelic differences in expression. Importantly we detected that allele-specific TF binding adds a layer of regulation in the absence of differences in promoter accessibility. Our results improve our understanding of *cis* regulatory elements' role on nitrogen regulation and starvation adaptation in winemaking yeast.

Materials and methods

Yeast strains and culture conditions

We used *S. cerevisiae* strains DBVPG6765 (hereinafter referred to as Wine European 'WE') and YPS128 (hereinafter referred to as North American 'NA'), as previously described in (58, 59). An F1 hybrid (WE x NA) was constructed by mating haploid strains of opposite mating types of WE (MATα, *hoΔ::NatMX, ura3Δ::KanMX*) and NA (*MATa, hoΔ::HphMX, ura3Δ::KanMX*), which had been generated previously (59).

Fermentations were performed in synthetic wine must (SM) following the recipe of (10) with modifications (**Table S6**). Two SMs containing different Yeast Assimilable Nitrogen

(YAN) concentrations were used for fermentations, i.e. SM300 (300 mg/mL YAN, excess nitrogen) and SM60 (60 mg/mL YAN, low nitrogen).

Fermentation assays

Pre-cultures were started from single colonies collected from YPD plates. These were grown for 48 hours in 50 mL falcon tubes containing 10 mL of SM60 or SM300 at 28 °C under constant agitation (250 rpm). From these pre-cultures, $1x10^6$ cells/mL were inoculated into 50 mL of SM60 or SM300 for fermentation. We conducted fermentations in at least three biological replicates in 250 mL bottles sealed with a drilled rubber stopper coupled to an airlock filled with 80% glycerol to allow for CO_2 release. A 100 mm cannula was inserted into the rubber stopper to perform periodical sampling. All components were autoclaved and fermentors were assembled under sterility inside a flow cabin. Fermentations were performed under constant agitation using magnetic stirring at 25 °C. Must samples were collected by obtaining 0.5 to 1 mL of fermented SM. Fermentation kinetics were monitored by manually weighing the fermentors to determine CO_2 release every two days.

High performance liquid chromatography (HPLC)

Samples obtained from fermentations were processed for quantification of amino acids using HPLC (Shimadzu, USA) with a Bio-Rad HPX –87H column (60). Briefly, 100 µL of filtered must were incubated with 3 µL of DEEM (Sigma-Aldrich 87-13-8) for 30 min in a sonication bath at room temperature in a solution containing 580 µL of borate buffer (pH 9) and 250 µL of methanol. After sonication, samples were incubated for 2 hours at 70 °C. After the derivatization reaction, 20 µL of the processed samples were injected into a Shimadzu Prominence HPLC (Shimadzu, USA). The concentration of each amino acid was calculated using a calibration curve obtained from sequential dilutions of non-fermented SM.

RNA extraction and sequencing

Yeast cells of the WE x NA hybrid were collected for RNA sequencing after 14 hours of fermentation in SM300 or SM60 in triplicates. Yeast cells were washed three times with PBS buffer and total RNA was extracted following a hot-formamide protocol (61). RNA in formamide was treated with DNase I (Promega, USA) to remove genomic DNA traces and purified using the GeneJET RNA Cleanup and Concentration Micro Kit (Thermo Fisher Scientific K0841). RNA integrity was confirmed using a Fragment Analyzer

(Agilent, USA). The RNA-seq libraries were constructed using the TruSeq RNA Sample Prep Kit v2 (Illumina, USA). The sequencing was conducted using paired-end 100-bp reads on an Illumina HiSeq in a single lane for all samples.

ATAC-seq assay and sequencing

For ATAC-seq, fermentations of the WE and NA parental strains, together with the WE x NA hybrid, were performed in duplicates, and cells in SM60 and SM300 were sampled after 14 hours. Cells were quantified using a Neubauer counting chamber. Five million haploid WE and NA cells, and 2.5 million diploid WE x NA cells were spun down (1.8 g for 4 min at room temperature) and washed twice using SB buffer (1 M Sorbitol, 10 mM $MgCl_2$, 40 mM HEPES pH 7.5). Cells were treated with 50 mg/mL of zymolyase 20T (EUROMEDEX UZ1000-A) in 200 µL of SB for 30 min at 30 °C, after which cells were washed two times with SB buffer. Immediately after, cells were incubated for 30 min at 37 °C in 50 µL of transposition mix, containing 25 µL Nextera Tagment DNA Buffer (Illumina # 15027866), 22.5 µL H_2O, and 2.5 µL Nextera Tagment DNA Enzyme I (Illumina # 15027865). Subsequently, DNA was purified using the DNA Clean and Concentrator-5 kit (Zymo Research D4003) following the supplier's instructions.

Tagmented DNA was amplified and barcoded using Nextera Index i5 and i7 series PCR primers. The PCR reaction consisted of 25 µL NEBNext Hi-Fidelity 2x PCR Master Mix (New England Biolabs NEB.M0541S), 7.5 µL H_2O, 6.25 µL i5 primer (10 mM), 6.25 µL i7 primer (10 mM) and 5 µL tagmented DNA. PCR cycles were set as: 1 cycle: 72 °C for 5 min.; 1 cycle: 98 °C for 30 s.; 8 cycles: 98 °C for 10 s., 63 °C for 30 s., 72 °C for 1 min. Subsequently, the amplified ATAC-seq library was subjected to double-sided size-selection using magnetic beads (AMPure XP, Beckman Coulter BC-A63880). Firstly, 50 µL of the library were incubated with 20 µL of beads (0.4X), after which the supernatant was collected. Subsequently, a left-side selection was performed by incubating the library with 1.1X of beads, after which the supernatant was discarded. DNA bound to the beads was washed twice with freshly made 80% ethanol, and then eluted in 20 µL of H_2O. Library quality was assessed using a Fragment Analyzer (Agilent, USA) and quantified in Qubit (Thermofisher, USA). Sequencing was conducted using paired-end 75-bp reads on an Illumina NextSeq 500.

Allele-specific read mapping

To estimate allele-specific counts derived from RNA-seq and ATAC-seq reads and to account for the mapping of SNP-informative reads, we modified the *S. cerevisiae* reference genome (R64-1-1) using SNPsplit (62). Genome-wide SNPs data from the WE and NA strains were obtained from the *Saccharomyces* Genome Resequencing Project (63). This data was used to replace the reference genome nucleotide sequence at 17,425 sites, in which the same genotype occurs for these two strains, but differed against the S288c reference strain. Next, the modified reference sequence was masked at 81,169 polymorphic sites between the NA and WE strains. A genome index was built using Bowtie2 (64), and then used to map ATAC-seq and RNA-seq reads in the WE x NA hybrid (using Bowtie2 option 'very-sensitive'). Before allele-specific mapping, sequencing reads were processed using fastp (65) to trim low quality 3' ends (Q < 20) and to exclude reads shorter than 36 bp (-A : disable automatic adaptor trimming; -3 : trimming by quality at 3' end; -l 36 : min length of 36 bp). Alignments containing all hybrid mapped reads were used for differential analysis of gene expression and chromatin accessibility of the WE x NA hybrid response to low nitrogen. Furthermore, SNPsplit was used to divide the hybrid alignments in two bam files, each containing only SNP-informative reads corresponding to either parental background (approximately 16-22% of the total mapped reads for each parent, **Table S7**). These allele-specific alignment files were used to evaluate the response (gene expression, chromatin accessibility, and TF binding) of each parental genome to low nitrogen.

Hybrid RNA-seq analysis

RNA-seq read counts per gene in the WE x NA hybrid were obtained from bam alignments using featureCounts (66) and the modified R64-1-1 genome annotation (ENSEMBL, release date 2018-10). Genes with at least 30 read-counts across the three replicates in at least one condition were selected for further statistical analysis (5,625 out of 6,534 genes). The differential transcriptome response of the WE x NA hybrid to low nitrogen (SM60) was estimated using DESeq2 (67) (design= ~ condition).

ATAC-seq data analysis

ATAC-seq read alignments were processed as follows: i) PCR duplicates were identified and removed using Genrich (github.com/jsh58/Genrich), then ii) Reads mapped to blacklisted regions were removed (i.e. mitochondrial genome, ribosomal genes in chromosome 12, and subtelomeres); and iii) Only properly paired mapped reads (mates

mapped to the same chromosome, pairs mapped in convergent direction) were kept using sambamba (68). ATAC-seq coverages around the Transcription Start Site (TSS) of those genes that passed RNA-seq count filters were obtained using deepTools computeMatrix (10 kb bins) and plotHeatmap (69), which were then further processed for plotting with the ComplexHeatmap R package (70). TSSs were obtained from (71). Correlations between ATAC-seq coverage (RPKM) and gene expression (RPKM) were calculated using the cor R package (method = "spearman"). To match genes with their nearby ATAC-seq signal, we selected a regulatory region 400 bp upstream of the TSS for each gene. For 224 genes that lacked an annotated TSS we used the transcript start as TSS. The ATAC-seq signals of 5,625 regulatory regions were quantified by counting mapped reads using featureCounts. Differential responses in ATAC-seq in the WE x NA hybrid to low nitrogen (SM60) were estimated using DESeq2 (design= ~ condition).

Allelic imbalance analyses of RNA-seq and ATAC-seq data

To test for differential allelic imbalance in chromatin accessibility or gene expression, we used the R package MBASED on allele counts (72). Counts at each SNP were obtained using ASEReadCounter employing the pipeline described in {Mendelevich, 2020 #87). We excluded those SNPs from one parent that overlapped with an indel of the other parental strain, and those that fell within low mappability regions of either parent (determined after mapping parental DNA-seq reads to the reference genome). Besides, only SNPs with at least five counts in both parents in SM300 or SM60 were retained for further analysis. Allelic counts were obtained by summing up SNP-read counts at regulatory regions (ATAC-seq) or genes (RNA-seq). A binomial test implemented in MBASED was used to statistically evaluate allelic imbalance. Genes or regulatory regions with an adjusted *P*-value (BH correction) lower than 0.05 in at least two replicates were considered to display significant allelic imbalance.

Allele-specific transcription factor binding analyses from ATAC-seq footprints

Analysis of allele-specific transcription factor binding (TFB) from ATAC-seq footprints was performed using TOBIAS {Bentsen, 2020 #88}. Briefly, ATAC-seq alignments of the WE x NA hybrid were used to obtain Tn5 cut-sites which were corrected for cutting bias by TOBIAS ATACorrect. The binding signal per site was calculated using

TOBIAS ScoreBigWig (--fp-min 5 --fp-max 30) for 5,401 regulatory regions. To calculate TFB, we obtained binding motifs for 141 yeast TFs from the JASPAR database (73). Based on this, TFB was quantified at motifs occurring in the regulatory regions by TOBIAS BinDetect. Binding scores were further processed in R. To evaluate differences in binding scores, we used a linear model implemented by the limma R package (74). We excluded from this analysis those TFs that showed low expression in SM300 or SM60 (RPKM < 5), and motif sites that were not considered by TOBIAS as "bound" in any condition. We considered binding differences with a FDR < 0.1 as statistically significant. For the analysis of allele specific TFB, allele-specific mapped reads were identified with SNPsplit (see above). Allele-specific alignments were further processed for TOBIAS analyses, as previously indicated for the hybrid. In allele-specific analysis, we excluded motifs located at regions having low ATAC-seq coverage after SNP-splitting. Total and allele-specific alignment statistics are provided in **Table S7**.

Data availability

RNA-seq and ATAC-seq raw reads are available in SRA under the project PRJNA705961.

ACKNOWLEDGMENTS

The authors are grateful to Valentina Abarca and Antonio Molina for their assistance in maintaining yeast collections. We thank J.M Alvarez for his assistance with chromatin accessibility analyses and C. Meneses (UNAB) for his support in ATAC library construction and sequencing. We thank Michael Handford (Universidad de Chile) for language support.

FAC was supported by Comisión Nacional de Investigación Científica y Tecnológica CONICYT FONDECYT [1180161] and ANID - Programa Iniciativa Científica Milenio - ICN17_022. CAV was supported by CONICYT FONDECYT [3170404]. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

References

1. Cubillos FA. 2016. Exploiting budding yeast natural variation for industrial processes. Curr Genet 62:745-751.

2. Zhang W, Du G, Zhou J, Chen J. 2018. Regulation of Sensing, Transportation, and Catabolism of Nitrogen Sources in Saccharomyces cerevisiae. Microbiology and Molecular Biology Reviews 82:e00040-17.
3. Brice C, Sanchez I, Tesnière C, Blondin B. 2014. Assessing the mechanisms responsible for differences between nitrogen requirements of saccharomyces cerevisiae wine yeasts in alcoholic fermentation. Applied and environmental microbiology 80:1330-1339.
4. Natarajan K, Meyer MR, Jackson BM, Slade D, Roberts C, Hinnebusch AG, Marton MJ. 2001. Transcriptional profiling shows that Gcn4p is a master regulator of gene expression during amino acid starvation in yeast. Mol Cell Biol 21:4347-68.
5. Gasch AP, Spellman PT, Kao CM, Carmel-Harel O, Eisen MB, Storz G, Botstein D, Brown PO. 2000. Genomic expression programs in the response of yeast cells to environmental changes. Mol Biol Cell 11:4241-57.
6. Hofman-Bang J. 1999. Nitrogen catabolite repression in Saccharomyces cerevisiae. Molecular Biotechnology 12:35-71.
7. Jara M, Cubillos FA, Garcia V, Salinas F, Aguilera O, Liti G, Martinez C. 2014. Mapping genetic variants underlying differences in the central nitrogen metabolism in fermenter yeasts. PLoS One 9:e86533.
8. Ibstedt S, Stenberg S, Bagés S, Gjuvsland AB, Salinas F, Kourtchenko O, Samy JKA, Blomberg A, Omholt SW, Liti G, Beltran G, Warringer J. 2014. Concerted Evolution of Life Stage Performances Signals Recent Selection on Yeast Nitrogen Use. Molecular Biology and Evolution 32:153-161.
9. Kessi-Pérez EI, Molinet J, Martínez C. 2020. Disentangling the genetic bases of Saccharomyces cerevisiae nitrogen consumption and adaptation to low nitrogen environments in wine fermentation. Biological Research 53:2.
10. Brice C, Cubillos FA, Dequin S, Camarasa C, Martinez C. 2018. Adaptability of the Saccharomyces cerevisiae yeasts to wine fermentation conditions relies on their strong ability to consume nitrogen. PLoS One 13:e0192383.
11. Legras J-L, Galeote V, Bigey F, Camarasa C, Marsit S, Nidelet T, Sanchez I, Couloux A, Guy J, Franco-Duarte R, Marcet-Houben M, Gabaldon T, Schuller D, Sampaio JP, Dequin S. 2018. Adaptation of S. cerevisiae to Fermented Food Environments Reveals Remarkable Genome Plasticity and the Footprints of Domestication. Molecular Biology and Evolution 35:1712-1727.
12. Marsit S, Dequin S. 2015. Diversity and adaptive evolution of Saccharomyces wine yeast: a review. FEMS Yeast Research 15.
13. Su Y, Macías LG, Heras JM, Querol A, Guillamón JM. 2021. Phenotypic and genomic differences among S. cerevisiae strains in nitrogen requirements during wine fermentations. Food Microbiology 96:103685.
14. Cubillos FA, Brice C, Molinet J, Tisne S, Abarca V, Tapia SM, Oporto C, Garcia V, Liti G, Martinez C. 2017. Identification of Nitrogen Consumption Genetic Variants in Yeast Through QTL Mapping and Bulk Segregant RNA-Seq Analyses. G3 (Bethesda) 7:1693-1705.
15. Molinet J, Cubillos FA, Salinas F, Liti G, Martinez C. 2019. Genetic variants of TORC1 signaling pathway affect nitrogen consumption in Saccharomyces cerevisiae during alcoholic fermentation. PLoS One 14:e0220515.
16. Kessi-Perez EI, Araos S, Garcia V, Salinas F, Abarca V, Larrondo LF, Martinez C, Cubillos FA. 2016. RIM15 antagonistic pleiotropy is responsible for differences in fermentation and stress response kinetics in budding yeast. FEMS Yeast Res 16.

677 17. Albert FW, Kruglyak L. 2015. The role of regulatory variation in complex traits and disease.
678 Nat Rev Genet 16:197-212.
679 18. Tirosh I, Reikhav S, Levy AA, Barkai N. 2009. A yeast hybrid provides insight into the
680 evolution of gene expression regulation. Science 324:659-62.
681 19. Salinas F, de Boer CG, Abarca V, Garcia V, Cuevas M, Araos S, Larrondo LF, Martinez C,
682 Cubillos FA. 2016. Natural variation in non-coding regions underlying phenotypic diversity
683 in budding yeast. Sci Rep 6:21849.
684 20. Hovhannisyan H, Saus E, Ksiezopolska E, Hinks Roberts AJ, Louis EJ, Gabaldón T. 2020.
685 Integrative Omics Analysis Reveals a Limited Transcriptional Shock After Yeast Interspecies
686 Hybridization. Frontiers in Genetics 11.
687 21. Li XC, Fay JC. 2017. Cis-Regulatory Divergence in Gene Expression between Two Thermally
688 Divergent Yeast Species. Genome Biology and Evolution 9:1120-1129.
689 22. Cubillos FA, Stegle O, Grondin C, Canut M, Tisne S, Gy I, Loudet O. 2014. Extensive cis-
690 regulatory variation robust to environmental perturbation in Arabidopsis. Plant Cell
691 26:4298-310.
692 23. Kita R, Venkataram S, Zhou Y, Fraser HB. 2017. High-resolution mapping of cis-regulatory
693 variation in budding yeast. Proc Natl Acad Sci U S A 114:E10736-E10744.
694 24. Wittkopp PJ, Kalay G. 2011. Cis-regulatory elements: molecular mechanisms and
695 evolutionary processes underlying divergence. Nat Rev Genet 13:59-69.
696 25. Wittkopp PJ, Haerum BK, Clark AG. 2004. Evolutionary changes in cis and trans gene
697 regulation. Nature 430:85-8.
698 26. Chang J, Zhou Y, Hu X, Lam L, Henry C, Green EM, Kita R, Kobor MS, Fraser HB. 2013. The
699 molecular mechanism of a cis-regulatory adaptation in yeast. PLoS Genet 9:e1003813.
700 27. Sharon E, Chen SA, Khosla NM, Smith JD, Pritchard JK, Fraser HB. 2018. Functional Genetic
701 Variants Revealed by Massively Parallel Precise Genome Editing. Cell 175:544-557 e16.
702 28. Brem RB, Yvert G, Clinton R, Kruglyak L. 2002. Genetic dissection of transcriptional
703 regulation in budding yeast. Science 296:752-5.
704 29. Cubillos FA, Yansouni J, Khalili H, Balzergue S, Elftieh S, Martin-Magniette ML, Serrand Y,
705 Lepiniec L, Baud S, Dubreucq B, Renou JP, Camilleri C, Loudet O. 2012. Expression variation
706 in connected recombinant populations of Arabidopsis thaliana highlights distinct
707 transcriptome architectures. BMC Genomics 13:117.
708 30. Hasin-Brumshtein Y, Hormozdiari F, Martin L, van Nas A, Eskin E, Lusis AJ, Drake TA. 2014.
709 Allele-specific expression and eQTL analysis in mouse adipose tissue. BMC Genomics
710 15:471.
711 31. Li R, Jeong K, Davis JT, Kim S, Lee S, Michelmore RW, Kim S, Maloof JN. 2018. Integrated
712 QTL and eQTL Mapping Provides Insights and Candidate Genes for Fatty Acid Composition,
713 Flowering Time, and Growth Traits in a F(2) Population of a Novel Synthetic Allopolyploid
714 Brassica napus. Front Plant Sci 9:1632.
715 32. Hill MS, Vande Zande P, Wittkopp PJ. 2020. Molecular and evolutionary processes
716 generating variation in gene expression. Nature Reviews Genetics doi:10.1038/s41576-
717 020-00304-w.
718 33. Klein DC, Hainer SJ. 2020. Genomic methods in profiling DNA accessibility and factor
719 localization. Chromosome Research 28:69-85.
720 34. Tripodi IJ, Allen MA, Dowell RD. 2018. Detecting Differential Transcription Factor Activity
721 from ATAC-Seq Data. Molecules (Basel, Switzerland) 23:1136.
722 35. Hendrickson DG, Soifer I, Wranik BJ, Kim G, Robles M, Gibney PA, McIsaac RS. 2018. A new
723 experimental platform facilitates assessment of the transcriptional and chromatin
724 landscapes of aging yeast. eLife 7:e39911.

36. Gowans GJ, Schep AN, Wong KM, King DA, Greenleaf WJ, Morrison AJ. 2018. INO80 Chromatin Remodeling Coordinates Metabolic Homeostasis with Cell Division. Cell Reports 22:611-623.
37. Jenull S, Tscherner M, Mair T, Kuchler K. 2020. ATAC-Seq Identifies Chromatin Landscapes Linked to the Regulation of Oxidative Stress in the Human Fungal Pathogen Candida albicans. Journal of Fungi 6:182.
38. Rivera-Mulia JC, Dimond A, Vera D, Trevilla-Garcia C, Sasaki T, Zimmerman J, Dupont C, Gribnau J, Fraser P, Gilbert DM. 2018. Allele-specific control of replication timing and genome organization during development. Genome research 28:800-811.
39. Zhang S, Zhang H, Zhou Y, Qiao M, Zhao S, Kozlova A, Shi J, Sanders AR, Wang G, Luo K, Sengupta S, West S, Qian S, Streit M, Avramopoulos D, Cowan CA, Chen M, Pang ZP, Gejman PV, He X, Duan J. 2020. Allele-specific open chromatin in human iPSC neurons elucidates functional disease variants. Science 369:561-565.
40. Xu J, Carter AC, Gendrel A-V, Attia M, Loftus J, Greenleaf WJ, Tibshirani R, Heard E, Chang HY. 2017. Landscape of monoallelic DNA accessibility in mouse embryonic stem cells and neural progenitor cells. Nature genetics 49:377-386.
41. Crépin L, Nidelet T, Sanchez I, Dequin S, Camarasa C. 2012. Sequential Use of Nitrogen Compounds by Saccharomyces cerevisiae during Wine Fermentation: a Model Based on Kinetic and Regulation Characteristics of Nitrogen Permeases. Applied and Environmental Microbiology 78:8102-8111.
42. Ronsmans A, Wery M, Szachnowski U, Gautier C, Descrimes M, Dubois E, Morillon A, Georis I. 2019. Transcription-dependent spreading of the Dal80 yeast GATA factor across the body of highly expressed genes. PLOS Genetics 15:e1007999.
43. Tirosh I, Sigal N, Barkai N. 2010. Divergence of nucleosome positioning between two closely related yeast species: genetic basis and functional consequences. Molecular Systems Biology 6:365.
44. Naranjo S, Smith JD, Artieri CG, Zhang M, Zhou Y, Palmer ME, Fraser HB. 2016. Dissecting the Genetic Basis of a Complex cis-Regulatory Adaptation. PLOS Genetics 11:e1005751.
45. Liu G, Bergenholm D, Nielsen J. 2016. Genome-Wide Mapping of Binding Sites Reveals Multiple Biological Functions of the Transcription Factor Cst6p in Saccharomyces cerevisiae. mBio 7.
46. Wu J, Zhang N, Hayes A, Panoutsopoulou K, Oliver SG. 2004. Global analysis of nutrient control of gene expression in Saccharomyces cerevisiae during growth and starvation. Proceedings of the National Academy of Sciences of the United States of America 101:3148-3153.
47. Boer VM, de Winde JH, Pronk JT, Piper MDW. 2003. The Genome-wide Transcriptional Responses of Saccharomyces cerevisiae Grown on Glucose in Aerobic Chemostat Cultures Limited for Carbon, Nitrogen, Phosphorus, or Sulfur *. Journal of Biological Chemistry 278:3265-3274.
48. Palma M, Madeira SC, Mendes-Ferreira A, Sá-Correia I. 2012. Impact of assimilable nitrogen availability in glucose uptake kinetics in Saccharomyces cerevisiae during alcoholic fermentation. Microbial Cell Factories 11:99.
49. Buziol S, Becker J, Baumeister A, Jung S, Mauch K, Reuss M, Boles E. 2002. Determination of in vivo kinetics of the starvation-induced Hxt5 glucose transporter of Saccharomyces cerevisiae. FEMS Yeast Res 2:283-91.

50. McManus CJ, May GE, Spealman P, Shteyman A. 2014. Ribosome profiling reveals post-transcriptional buffering of divergent gene expression in yeast. Genome Research 24:422-430.
51. Pai AA, Cain CE, Mizrahi-Man O, De Leon S, Lewellen N, Veyrieras J-B, Degner JF, Gaffney DJ, Pickrell JK, Stephens M, Pritchard JK, Gilad Y. 2012. The Contribution of RNA Decay Quantitative Trait Loci to Inter-Individual Variation in Steady-State Gene Expression Levels. PLOS Genetics 8:e1003000.
52. Crépin L, Sanchez I, Nidelet T, Dequin S, Camarasa C. 2014. Efficient ammonium uptake and mobilization of vacuolar arginine by Saccharomyces cerevisiae wine strains during wine fermentation. Microb Cell Fact 13:109.
53. BELL S-J, HENSCHKE PA. 2005. Implications of nitrogen nutrition for grapes, fermentation and wine. Australian Journal of Grape and Wine Research 11:242-295.
54. Lorenz MC, Heitman J. 1998. The MEP2 ammonium permease regulates pseudohyphal differentiation in Saccharomyces cerevisiae. The EMBO journal 17:1236-1247.
55. Brandriss MC, Magasanik B. 1979. Genetics and physiology of proline utilization in Saccharomyces cerevisiae: enzyme induction by proline. J Bacteriol 140:498-503.
56. Peltier E, Bernard M, Trujillo M, Prodhomme D, Barbe J-C, Gibon Y, Marullo P. 2018. Wine yeast phenomics: A standardized fermentation method for assessing quantitative traits of Saccharomyces cerevisiae strains in enological conditions. PLOS ONE 13:e0190094.
57. Gutiérrez A, Beltran G, Warringer J, Guillamón JM. 2013. Genetic basis of variations in nitrogen source utilization in four wine commercial yeast strains. PloS one 8:e67166-e67166.
58. Liti G, Haricharan S, Cubillos FA, Tierney AL, Sharp S, Bertuch AA, Parts L, Bailes E, Louis EJ. 2009. Segregating YKU80 and TLC1 alleles underlying natural variation in telomere properties in wild yeast. PLoS Genet 5:e1000659.
59. Cubillos FA, Louis EJ, Liti G. 2009. Generation of a large set of genetically tractable haploid and diploid Saccharomyces strains. FEMS Yeast Res 9:1217-25.
60. Gómez-Alonso S, Hermosín-Gutiérrez I, García-Romero E. 2007. Simultaneous HPLC analysis of biogenic amines, amino acids, and ammonium ion as aminoenone derivatives in wine and beer samples. J Agric Food Chem 55:608-13.
61. Shedlovskiy D, Shcherbik N, Pestov DG. 2017. One-step hot formamide extraction of RNA from Saccharomyces cerevisiae. RNA Biol 14:1722-1726.
62. Krueger F, Andrews S. 2016. SNPsplit: Allele-specific splitting of alignments between genomes with known SNP genotypes [version 2; peer review: 3 approved]. F1000Research 5.
63. Bergström A, Simpson JT, Salinas F, Barré B, Parts L, Zia A, Nguyen Ba AN, Moses AM, Louis EJ, Mustonen V, Warringer J, Durbin R, Liti G. 2014. A high-definition view of functional genetic variation from natural yeast genomes. Mol Biol Evol 31:872-88.
64. Langmead B, Salzberg SL. 2012. Fast gapped-read alignment with Bowtie 2. Nature Methods 9:357-359.
65. Chen S, Zhou Y, Chen Y, Gu J. 2018. fastp: an ultra-fast all-in-one FASTQ preprocessor. Bioinformatics 34:i884-i890.
66. Liao Y, Smyth GK, Shi W. 2013. featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. Bioinformatics 30:923-930.
67. Love MI, Huber W, Anders S. 2014. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biology 15:550.
68. Tarasov A, Vilella AJ, Cuppen E, Nijman IJ, Prins P. 2015. Sambamba: fast processing of NGS alignment formats. Bioinformatics (Oxford, England) 31:2032-2034.

69. Ramírez F, Ryan DP, Grüning B, Bhardwaj V, Kilpert F, Richter AS, Heyne S, Dündar F, Manke T. 2016. deepTools2: a next generation web server for deep-sequencing data analysis. Nucleic Acids Research 44:W160-W165.
70. Gu Z, Eils R, Schlesner M. 2016. Complex heatmaps reveal patterns and correlations in multidimensional genomic data. Bioinformatics 32:2847-2849.
71. Chereji RV, Bryson TD, Henikoff S. 2019. Quantitative MNase-seq accurately maps nucleosome occupancy levels. Genome Biology 20:198.
72. Mayba O, Gilbert HN, Liu J, Haverty PM, Jhunjhunwala S, Jiang Z, Watanabe C, Zhang Z. 2014. MBASED: allele-specific expression detection in cancer tissues and cell lines. Genome Biology 15:405.
73. Fornes O, Castro-Mondragon JA, Khan A, van der Lee R, Zhang X, Richmond PA, Modi BP, Correard S, Gheorghe M, Baranašić D, Santana-Garcia W, Tan G, Chèneby J, Ballester B, Parcy F, Sandelin A, Lenhard B, Wasserman WW, Mathelier A. 2019. JASPAR 2020: update of the open-access database of transcription factor binding profiles. Nucleic Acids Research 48:D87-D92.
74. Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, Smyth GK. 2015. limma powers differential expression analyses for RNA-sequencing and microarray studies. Nucleic Acids Research 43:e47-e47.

Figures

FIG. 1. Allele-specific expression and chromatin accessibility profiling to reveal molecular mechanisms orchestrating gene expression divergence in response to nitrogen scarcity. A) Representative haploid strains of the WE (Wine-European) and NA (North-American) lineages of S. cerevisiae were selected to construct a WE x NA hybrid which was used to perform fermentations in synthetic wine media under low and excess nitrogen conditions (SM60 and SM300, respectively). Large boxes denote a cis regulatory region (R.R.) and an ORF (arrow on top), whilst smaller boxes depict polymorphisms occurring in each genetic backgrounds. B) RNA-seq and ATAC-seq were employed to profile the hybrid's transcriptome and open chromatin landscape at regulatory regions between nitrogen conditions. Grey bars denote the common pool of hybrid reads. C) RNA-seq and ATACseq reads that mapped to parent-specific single-nucleotide polymorphisms (SNPs) were used to determine Allele-Specific Expression (ASE) and Allele-Specific Accessibility (ASE and ASA, respectively). Reads highlighted in colours indicate different pools according to their parental origin. D) ATAC-seq cut sites were used to infer in silico footprinting of transcription factor binding (TFB) in the hybrid. E) Allele-specific ATAC-seq alignments were separated by parental origin to obtain allele-specific ATAC-seq cut sites which were used to infer allele-specific binding (ASB) from TFB in silico footprinting.

FIG. 2. Fermentation kinetics and nitrogen consumption in WE and NA parental strains and the WE x NA hybrid. A) The boxplots show the total Yeast Assimilable

Nitrogen (YAN) consumed by the NA, WE, and WE x NA strains after 14 hours of fermentation in synthetic wine must containing excess or low nitrogen (SM300 and SM60, respectively). Consumption of each aminoacid is shown as dots. Asterisks denote significance difference (Tukey post-hoc), * < 0.05, ** < 0.001, *** < 1e-5. B) The bars indicate consumption fraction of each aminoacid in SM300 and SM60 after 14 hours of fermentation by the WE x NA (yellow), WE (green), and NA (purple) strains. Letters indicate significant differences between groups (ANOVA, Tukey post hoc test). Error bars show the standard deviation (SD). C) The bars show the heterosis coefficient for the consumption of each aminoacid in SM300 and SM60. Values > 1 indicate best parent heterosis, values < -1 indicate worst-parent heterosis. D) Total production of CO2 at the end of fermentation process of each indicated strain and culture condition. Letters indicate significant differences among groups (ANOVA, Tukey post hoc test) and error bars denote SD. E) Fermentation rates of the WE, NA, and WE x NA strains in SM300 and SM60 culture conditions.

FIG. 3. Hybrid transcriptome and regulatory region accessibility in response to fermentation in low nitrogen wine must. A) The heatmaps show the Pearson correlation between ATAC-seq coverage and total gene expression of the WE x NA hybrid after 14 hours of fermentation in SM300 and SM60. For each gene, different regulatory region segments were evaluated using a sliding window approach starting from 1,000 bp upstream of the transcription start site (TSS) to 200 bp upstream of the TSS, and ending from 100 bp upstream of the TSS to 500 bp downstream of the TSS. The white boxes indicate the regions with the highest correlation in SM300 and SM60. B) Upset plot showing the number of differentially expressed genes (DEGs) and differentially accessible regions (DARs), the set sizes, and the intersect among combinations. C) Average absolute Log2 fold changes of the ATAC-seq and RNA-seq datasets. D) Number of genes occurring in GO terms (Biological Processes) enriched among common DEGs and DARs.

FIG. 4. Allelic imbalance in accessibility and gene expression in the WE x NA hybrid. A) Upset plot showing the number of genes in differential allelic imbalance under SM300 and SM60 in the WE x NA hybrid, and the intersection between conditions. Blue bars indicate that allelic ratios were maintained between conditions, while red bars indicate that allelic ratios were inverse between conditions. B) The bar plots show the expression ratio of genes with significant ASE and induced by SM60 (FDR < 0.05, Log2 fold-change > 3). Genes with allelic ratios higher than 0.5 had the higher expression of the WE allele, while

genes with allelic ratios lower than 0.5 had higher expression of the NA allele. Blue bars indicate genes that also showed differential ASA in the same direction of ASE, i.e. bias favoring the same genotype in accessibility and expression. C) Upset plot showing the numbers of regulatory regions in differential ASA in SM300 and SM60 and the intersection across conditions. D and E) The upset plot shows the number of differential ASE and ASA in SM300 (D) and SM60 (E), and the intersection of these within the same condition.

FIG. 5. Allele-specific differences in transcription factor binding. A) Correlation between ATAC-seq Log2 fold changes (Log2FC) and Transcription Factor Binding (TFB) Log2FC (blue dots), and RNA-seq Log2FC and TFB Log2FC (red dots) in response to low nitrogen in the WE x NA hybrid. B) Correlation between Cst6p binding scores and expression in response to low nitrogen of its target genes. C) Correlation between the allelic expression ratio of the target genes of Gat1p, Dal80p, and Gzf3p and the allelic binding ratio of these transcription factors (TFs) at their target genes regulatory regions. Allelic ratios higher than 0.5 indicate higher expression or binding of the WE allele. Sequence logos indicate the motif used for each TF to scan binding sites throught the genome. D) The Venn diagram displays the number of binding sites showing significant ASB in SM300 or SM60, and in both conditions. E) The Venn diagrams indicate the number of regulatory regions having at least one binding site showing ASB, and the number which were associated with regions showing ASE or ASA, or both, in SM60 and SM300.

FIG. 6. ASB in genes induced under low nitrogen and showing ASE. A) Heatmap depicting the allele-specific binding scores of 61 TFs bound to 23 regulatory regions of genes induced by low nitrogen. These genes showed allelic imbalance in gene expression (dot plot) but no allelic differences in accessibility. TFB scores are shown as z-score. B-C-D) Three examples of allelic imbalance in TFB explaining ASE. The bar plot shows the bias-corrected Tn5 cut site signal obtained as reported by TOBIAS. Negative values (red bars) surrounded by positive values (blue bars) suggest the presence of a transcription factor binding footprint. The range of values for each bar plot is indicated. ATAC-seq coverage demonstrates similar accessibility in both parental backgrounds. Allele specific data (cut sites and coverage) were obtained after splitting SNP-informative reads from the hybrid ATAC-seq alignments. Green bars show the location of predicted motifs found at differentially bound regions. Logos show the consensus motifs.

FIG. 7. Allelic imbalance profile of genes encoding permeases involved in nitrogen uptake. Summary of the allelic bias in expression and accessibility found for genes encoding transporters involved in nitrogen uptake. Genes are colored according to the pathway by which they are mainly regulated (NCR or SPS). The preferred substrates indicated for each transporter are based on those described previously (41). Data obtained from ATAC-seq or RNA-seq are shown with blue or red blackgrounds, respectively. The bars show the average fold-change induction (purple) or repression (green) in response to low nitrogen. The line plots show the allelic bias favoring either the WE (> 0.5) or the NA (< 0.5) allele.

Supplemental data

Figure S1. ATAC-seq Quality Control plots. A) Fragment-size distribution of each ATAC-seq library mapped to the yeast genome. Peaks corresponding to nucleosome-free regions, mono-nucleosome and di-nucleosome are observed, indicating a successful ATAC-seq experiment. B) Aggregated ATAC-seq coverage upstream and downstream of the TSS of every *S. cerevisiae* gene.

Figure S2. GO enrichment analyses of ATAC-seq and RNA-seq differentially regulated genes. Number of DEGs or DARs occurring in enriched GO terms of different set of interest

Figure S3. ASB in downregulated nitrogen-responsive genes showing ASE but not ASA. A) Heatmap depicting the allele-specific binding scores of 51 TFs at 13 regulatory regions of genes repressed by low nitrogen. These genes showed allelic imbalance in gene expression (dot plot) but no allelic differences in accessibility. TFB scores are shown as z-score. B-C) Two examples of allelic imbalance in TFB explaining ASE. The bar plot shows the bias-corrected Tn5 cut site signal obtained as reported by TOBIAS. Negative values (red bars) surrounded by positive values (blue bars) suggest the presence of a transcription factor binding footprint. The range of values for each bar plot is indicated. ATAC-seq coverage demonstrates similar accessibility in both parental backgrounds. Allele specific data (cut sites and coverage) were obtained after splitting SNP-informative reads from the hybrid ATAC-seq alignments. Green bars show the location of predicted motifs found at differentially bound regions. Logos show the consensus motifs.

953 **Table S1**. A) Set of genes with differential expression and/or differential accessibility at
954 their regulatory region in response to low nitrogen. B) List of the 30 genes that are most
955 upregulated by low nitrogen

956 **Table S2**. Top 20 enriched GO terms of the DEGs repressed (A) or induced (B) in SM60.
957 Top 20 enriched GO terms of the DARs repressed (C) or induced (D) in SM60.

958 **Table S3**. ASE and ASA ratio counts and statistical test

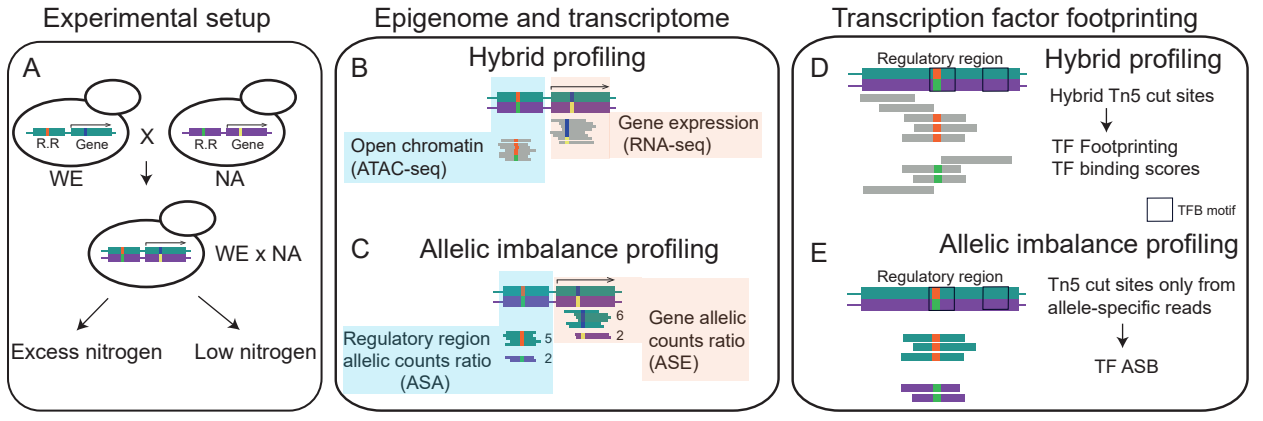
959 **Table S4**. GO enrichment of genes in ASE with higher WE (A) or NA (B) allelic expression.
960 GO enrichment of genes in ASA with higher WE (A) or NA (B) allelic accessibility.

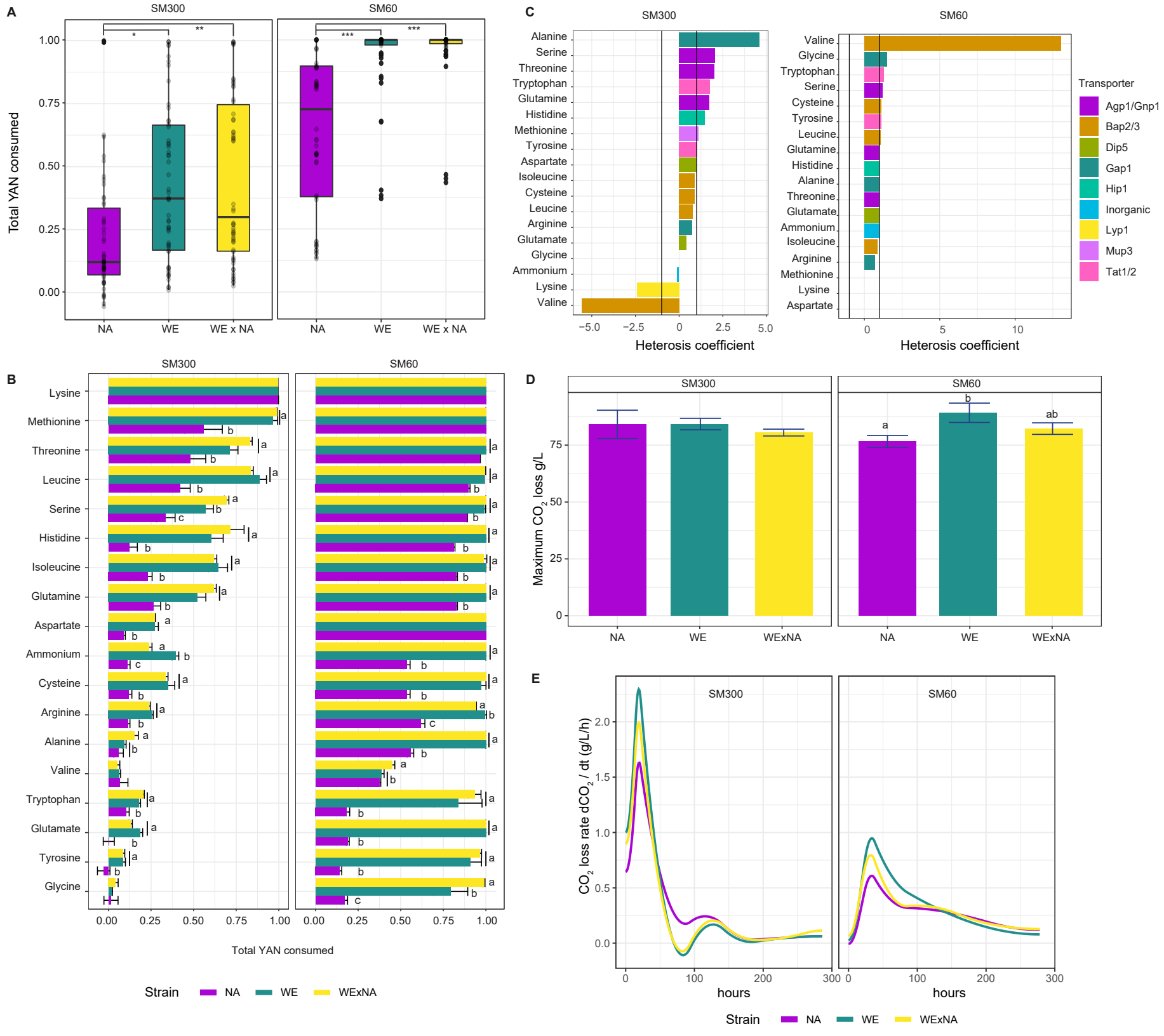
961 **Table S5**. A) Correlation between TFB, RNA-seq, and ATAC-seq for fold changes
962 between conditions. B) Correlation between ASB, ASE, and ASA using counts or score
963 ratios. Motifs with significant ASB in SM60 (C) and SM300 (D)

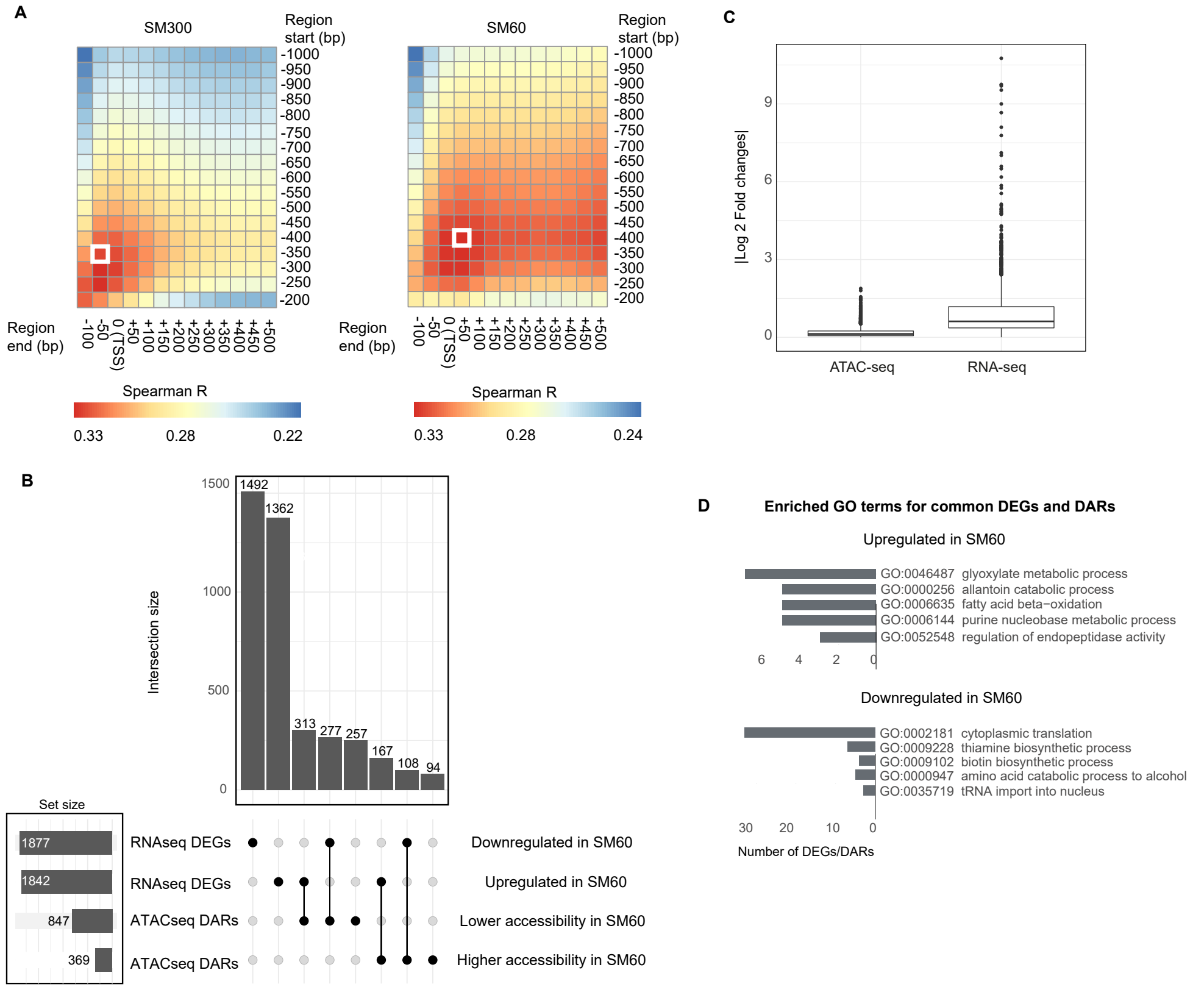
964 **Table S6**. Composition of SM300 and SM60

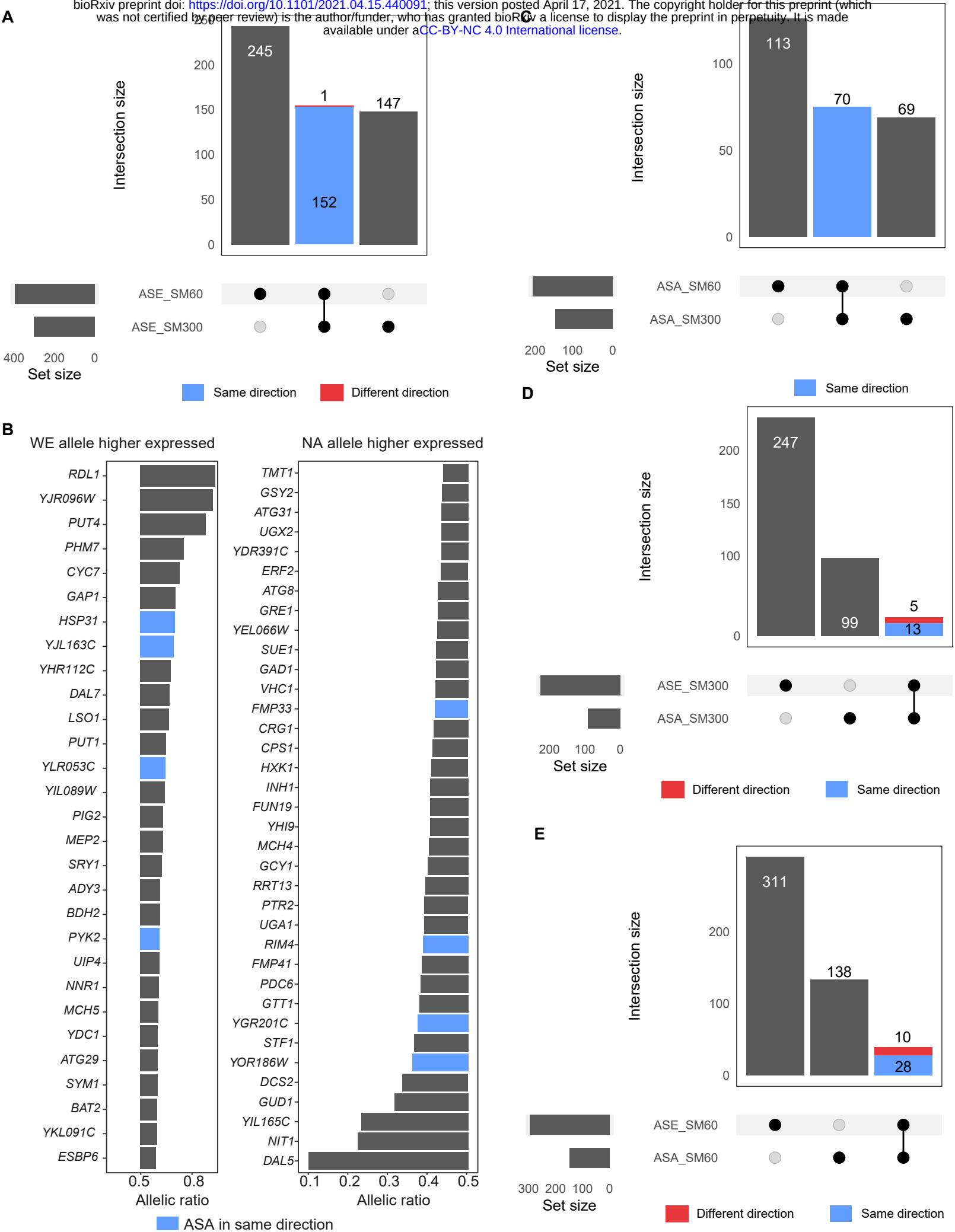
965 **Table S7**. Alignment statistics of a) RNA-seq and b) ATAC-seq

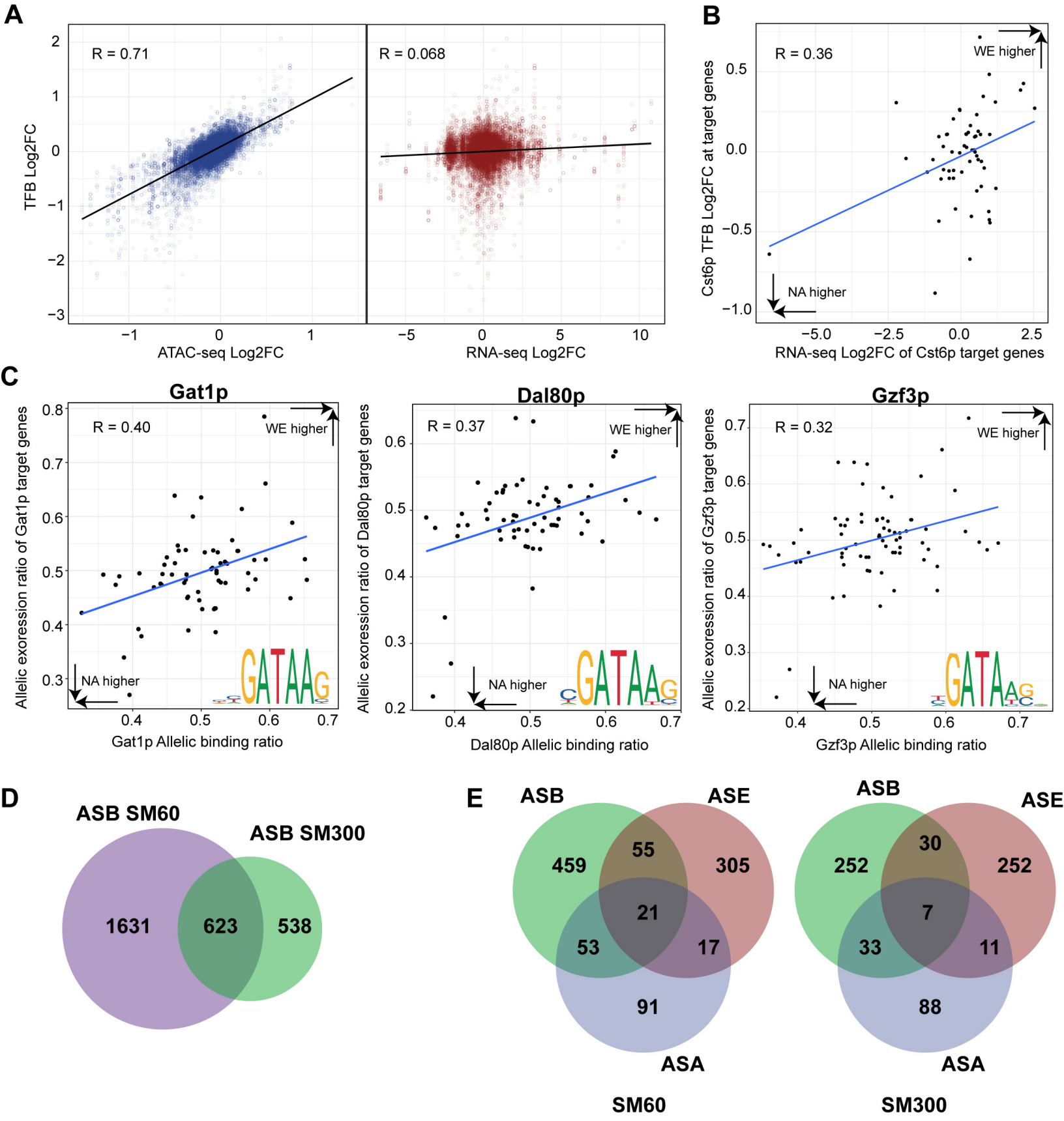
966

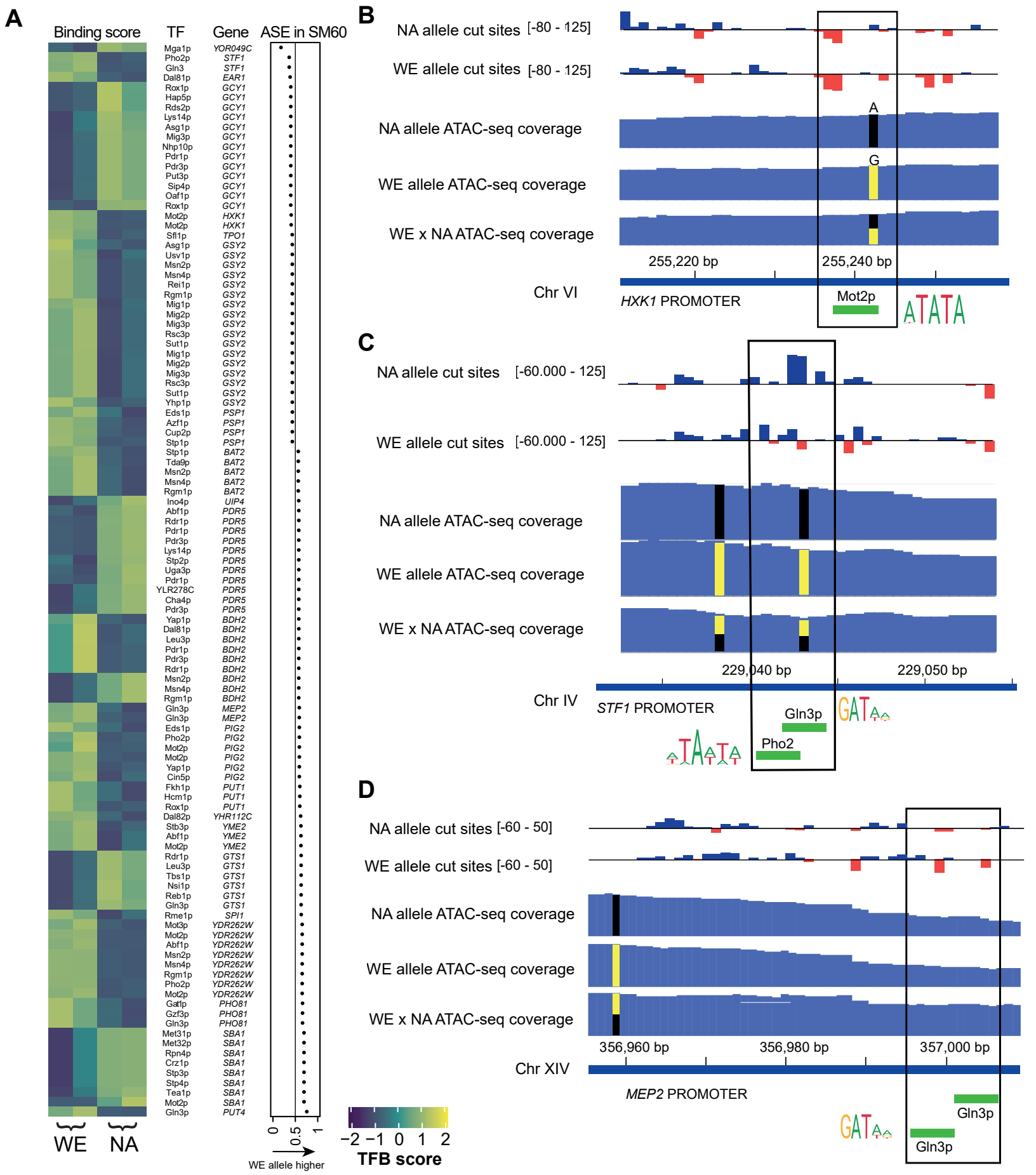

Experimental setup
A
R.R
Gene
X
WE
R.R
Gene
NA
WE x NA
Excess nitrogen
Low nitrogen
Epigenome and transcriptome
Hybrid profiling
B
Open chromatin
(ATAC-seq)
Gene expression
(RNA-seq)
C
Allelic imbalance profiling
Regulatory region
allelic counts ratio
(ASA)
5
2
6
2
Gene allelic
counts ratio
(ASE)
Transcription factor footprinting
D
Regulatory region
Hybrid profiling
Hybrid Tn5 cut sites
TF Footprinting
TF binding scores
TFB motif
E
Allelic imbalance profiling
Regulatory region
Tn5 cut sites only from
allele-specific reads
TF ASB

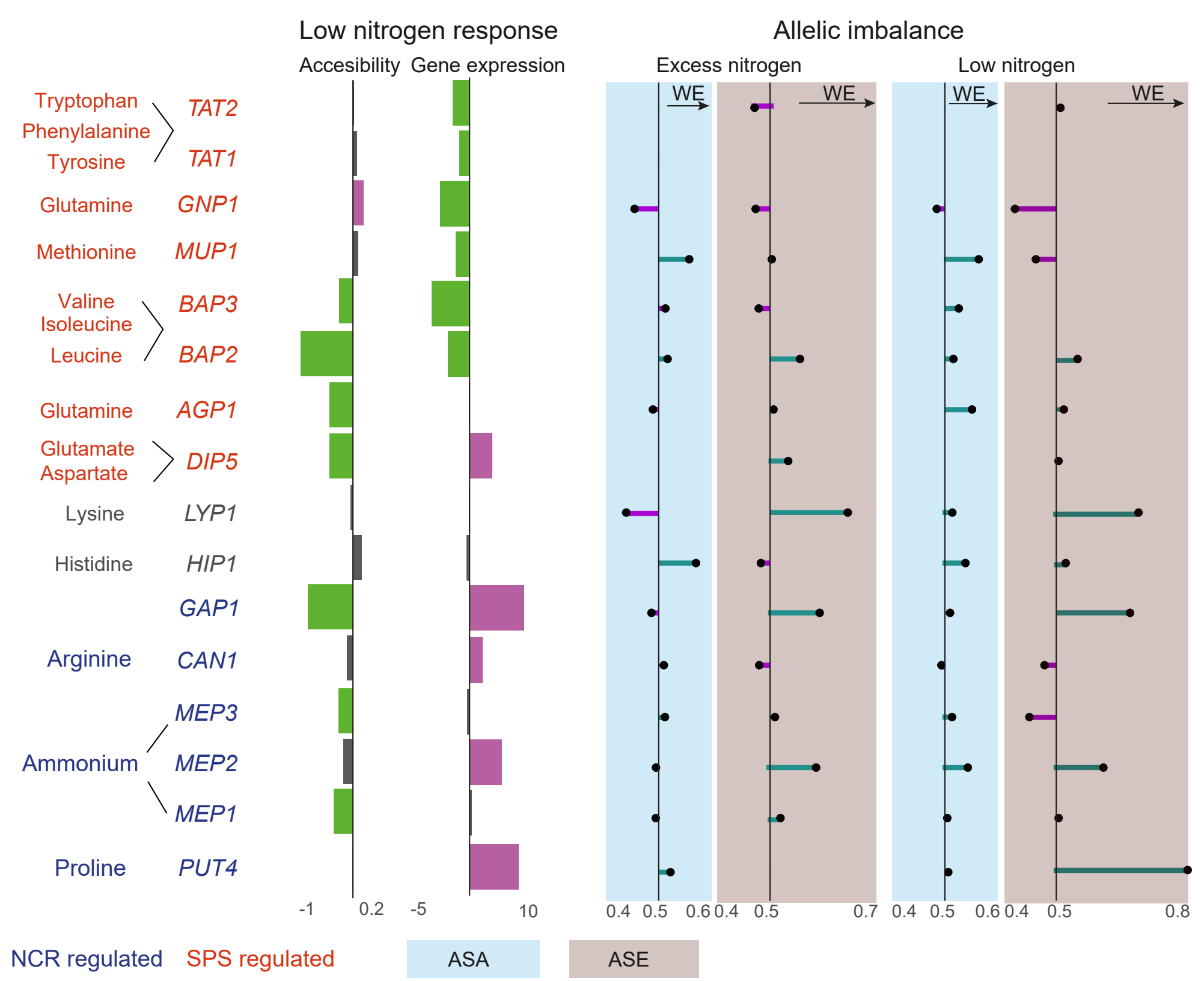

A
SM300
SM60
Total YAN consumed
1.00
0.75
0.50
0.25
0.00
*
**
***
***
NA
WE
WE x NA
NA
WE
WE x NA
C
SM300
SM60
Alanine
Serine
Threonine
Tryptophan
Glutamine
Histidine
Methionine
Tyrosine
Aspartate
Isoleucine
Cysteine
Leucine
Arginine
Glutamate
Glycine
Ammonium
Lysine
Valine
−5.0
−2.5
0
2.5
5.0
Heterosis coefficient
Valine
Glycine
Tryptophan
Serine
Cysteine
Tyrosine
Leucine
Glutamine
Histidine
Alanine
Threonine
Glutamate
Ammonium
Isoleucine
Arginine
Methionine
Lysine
Aspartate
0
5
10
Heterosis coefficient
Transporter
Agp1/Gnp1
Bap2/3
Dip5
Gap1
Hip1
Inorganic
Lyp1
Mup3
Tat1/2
B
SM300
SM60
Lysine
Methionine
Threonine
Leucine
Serine
Histidine
Isoleucine
Glutamine
Aspartate
Ammonium
Cysteine
Arginine
Alanine
Valine
Tryptophan
Glutamate
Tyrosine
Glycine
0.00
0.25
0.50
0.75
1.00
Total YAN consumed
Strain
NA
WE
WExNA
D
SM300
SM60
Maximum CO_2 loss g/L
0
25
50
75
a
b
ab
NA
WE
WExNA
E
SM300
SM60
CO_2 loss rate dCO_2 / dt (g/L/h)
0.0
0.5
1.0
1.5
2.0
0
100
200
300
hours
Strain
NA
WE
WExNA


A
SM300
Region start (bp)
-1000
-950
-900
-850
-800
-750
-700
-650
-600
-550
-500
-450
-400
-350
-300
-250
-200
Region end (bp)
-100
-50
0 (TSS)
+50
+100
+150
+200
+250
+300
+350
+400
+450
+500
Spearman R
0.33
0.28
0.22
SM60
Region start (bp)
Region end (bp)
Spearman R
0.33
0.28
0.24
C
|Log 2 Fold changes|
9
6
3
0
ATAC-seq
RNA-seq
B
Intersection size
1500
1000
500
0
1492
1362
313
277
257
167
108
94
Set size
1877
1842
847
369
RNAseq DEGs
RNAseq DEGs
ATACseq DARs
ATACseq DARs
Downregulated in SM60
Upregulated in SM60
Lower accessibility in SM60
Higher accessibility in SM60
D
Enriched GO terms for common DEGs and DARs
Upregulated in SM60
GO:0046487 glyoxylate metabolic process
GO:0000256 allantoin catabolic process
GO:0006635 fatty acid beta−oxidation
GO:0006144 purine nucleobase metabolic process
GO:0052548 regulation of endopeptidase activity
6
4
2
0
Downregulated in SM60
GO:0002181 cytoplasmic translation
GO:0009228 thiamine biosynthetic process
GO:0009102 biotin biosynthetic process
GO:0000947 amino acid catabolic process to alcohol
GO:0035719 tRNA import into nucleus
30
20
10
0
Number of DEGs/DARs






A
Binding score
TF
Gene
ASE in SM60
Mga1p YOR049C
Pho2p STF1
Gln3 STF1
Dal81p EAR1
Rox1p GCY1
Hap5p GCY1
Rds2p GCY1
Lys14p GCY1
Asg1p GCY1
Mig3p GCY1
Nhp10p GCY1
Pdr1p GCY1
Pdr3p GCY1
Put3p GCY1
Sip4p GCY1
Oaf1p GCY1
Rox1p GCY1
Mot2p HXK1
Mot2p HXK1
Sfl1p TPO1
Asg1p GSY2
Usv1p GSY2
Msn2p GSY2
Msn4p GSY2
Rei1p GSY2
Rgm1p GSY2
Mig1p GSY2
Mig2p GSY2
Mig3p GSY2
Rsc3p GSY2
Sut1p GSY2
Mig1p GSY2
Mig2p GSY2
Mig3p GSY2
Rsc3p GSY2
Sut1p GSY2
Yhp1p GSY2
Eds1p PSP1
Azf1p PSP1
Cup2p PSP1
Stp1p PSP1
Stp1p BAT2
Tda9p BAT2
Msn2p BAT2
Msn4p BAT2
Rgm1p BAT2
Ino4p UIP4
Abf1p PDR5
Rdr1p PDR5
Pdr1p PDR5
Pdr3p PDR5
Lys14p PDR5
Stp2p PDR5
Uga3p PDR5
Pdr1p PDR5
YLR278C PDR5
Cha4p PDR5
Pdr3p PDR5
Yap1p BDH2
Dal81p BDH2
Leu3p BDH2
Pdr1p BDH2
Pdr3p BDH2
Rdr1p BDH2
Msn2p BDH2
Msn4p BDH2
Rgm1p BDH2
Gln3p MEP2
Gln3p MEP2
Eds1p PIG2
Pho2p PIG2
Mot2p PIG2
Mot2p PIG2
Yap1p PIG2
Cin5p PIG2
Fkh1p PUT1
Hcm1p PUT1
Rox1p PUT1
Dal82p YHR112C
Stb3p YME2
Abf1p YME2
Mot2p YME2
Rdr1p GTS1
Leu3p GTS1
Tbs1p GTS1
Nsi1p GTS1
Reb1p GTS1
Gln3p GTS1
Rme1p SPI1
Mot3p YDR262W
Mot2p YDR262W
Abf1p YDR262W
Msn2p YDR262W
Msn4p YDR262W
Rgm1p YDR262W
Pho2p YDR262W
Mot2p YDR262W
Gat1p PHO81
Gzf3p PHO81
Gln3p PHO81
Met31p SBA1
Met32p SBA1
Rpn4p SBA1
Crz1p SBA1
Stp3p SBA1
Stp4p SBA1
Tea1p SBA1
Mot2p SBA1
Gln3p PUT4
WE
NA
0
0.5
1
WE allele higher
−2 −1 0 1 2
TFB score
B
NA allele cut sites [-80 - 125]
WE allele cut sites [-80 - 125]
NA allele ATAC-seq coverage
WE allele ATAC-seq coverage
WE x NA ATAC-seq coverage
A
G
255,220 bp
255,240 bp
Chr VI
HXK1 PROMOTER
Mot2p
C
NA allele cut sites [-60.000 - 125]
WE allele cut sites [-60.000 - 125]
NA allele ATAC-seq coverage
WE allele ATAC-seq coverage
WE x NA ATAC-seq coverage
229,040 bp
229,050 bp
Chr IV
STF1 PROMOTER
Gln3p
Pho2
D
NA allele cut sites [-60 - 50]
WE allele cut sites [-60 - 50]
NA allele ATAC-seq coverage
WE allele ATAC-seq coverage
WE x NA ATAC-seq coverage
356,960 bp
356,980 bp
357,000 bp
Chr XIV
MEP2 PROMOTER
Gln3p
Gln3p


Low nitrogen response
Accesibility
Gene expression
Allelic imbalance
Excess nitrogen
Low nitrogen
WE
Tryptophan
Phenylalanine
Tyrosine
TAT2
TAT1
Glutamine
GNP1
Methionine
MUP1
Valine
Isoleucine
Leucine
BAP3
BAP2
Glutamine
AGP1
Glutamate
Aspartate
DIP5
Lysine
LYP1
Histidine
HIP1
GAP1
Arginine
CAN1
MEP3
Ammonium
MEP2
MEP1
Proline
PUT4
-1
0.2
-5
10
0.4
0.5
0.6
0.7
0.8
NCR regulated
SPS regulated
ASA
ASE