Complex genetic and epigenetic regulation deviates gene expression from a global transcriptional program

Mónica Chagoyen[1] and Juan F Poyatos[2]

[1] Computational Systems Biology Group (CNB-CSIC), Darwin 3, 28049, Madrid, Spain
[2] Logic of Genomic Systems Laboratory (CNB-CSIC), Darwin 3, 28049, Madrid, Spain

Email: *monica.chagoyen@cnb.csic.es* (MC), *jpoyatos@cnb.csic.es* (JFP)

Abstract

Environmental or genetic perturbations lead to gene expression changes. While most analyses of these changes emphasize the presence of qualitative differences on just a few genes, we now know that changes are widespread. This large-scale variation has been linked to the exclusive influence of a global transcriptional program determined by the new physiological state of the cell. However, given the sophistication of eukaryotic regulation, we expect to have a complex structure of deviations from the global program. Here, we examine the regulatory landscape that contributes to these deviations. Using data of *Saccharomyces cerevisiae* expression in different nutrient conditions, we first propose a five-component genome partition as a framework to understand expression variation. In this framework, we recognize invariant genes, whose regulation is dominated by the global program, specific genes, which substantially depart from it, and two additional classes that respond to intermediate regulatory schemes. Whereas the invariant class shows a considerable absence of specific regulation, the rest is enriched by regulation at the level of transcription factors (TFs) and epigenetic modulators. We nevertheless find markedly different strategies in how these classes deviate. On the one hand, there are TFs that act in an exclusive way between partition constituents, and on the other, the action of chromatin modifiers is significantly diverse. The balance between regulatory strategies ultimately modulates the action of the general transcription machinery, and therefore limits the possibility of establishing a single program of expression change at a genomic scale.

Introduction

A global program of regulation impacts the expression of most genes (Liang et al., 1999, Dennis et al., 2004, Zaslaver et al., 2009, Scott et al., 2010). This program depends on the availability of individual components of the cellular expression machinery, e.g., free RNA polymerases, co-factors, ribosomes, etc., which affects the rates of transcription and/or translation and thus determines the overall state of the cell.

The study of this program goes back to the research of the early school of bacterial physiologists who introduced the notion of steady state of cellular growth, in which growth rate acted as a valid proxy for cell physiology and consequently for the global program (Schaechter et al., 1958, Neidhardt and Magasanik, 1960, Maaløe, 1979). Indeed, these early studies documented that the macromolecular composition of cells is a function of the growth rate. With such experimental approach, the general question of how gene expression is influenced by physiology is addressed by studying how expression depended on growth rate.

The research has been taken up in more recent work (Zaslaver et al., 2009, Scott et al., 2010, Klumpp and Hwa, 2014, Bosdriesz et al., 2015), emphasizing a framework of distribution of limited resources associated with gene expression, i.e., several cellular parameters manifest as resource trade-offs. Changes in expression as a function of growth rate can be understood by means of coarse genomic partitions that gain differential access to these resources according to separate functional categories; the minimal partition being that between ribosomal and metabolic genes (Zaslaver et al., 2009, Scott et al., 2010, Hui et al., 2015). Likewise, a recent model proposed a broader picture of resource allocation in which only a small group of genes, specific genes, deviate from the global expression program that is constraint by the remaining resources not engaged in the activation of the specific response (Keren et al., 2014).

Understanding what precise control mechanisms decouple gene expression from such global transcriptional program is therefore of great interest. For instance, results in bacteria have demonstrated the prevalence of the global expression program, while they have lowered the importance of transcription factors (TFs) controlling the assumed deviations from it. TFs seem only to complement the action of the global regulation (Berthoumieux et al., 2014, Gerosa et al., 2014), in combination with a few metabolites (Kochanowski et al., 2017). Results in eukaryotes are however lacking (Keren et al., 2014, Metzl-Raz et al., 2017), partly because one has to study a more complex basal regulatory machinery.

In this work, we wanted to characterize which molecular elements of control influence the deviation from the global transcriptional program in eukaryotes. For this, we have outlined in detail the groups of genes that can *a priori* be more sensitive to the global program. We do this by introducing a partition of the yeast genome into five components, which incorporates previous models (Zaslaver et al., 2009, Scott et al., 2010, Keren et al., 2014). In this framework, we are able to recognize invariant genes, whose regulation seems to be dominated by the global program, specific genes, which are those that seemingly deviate the most from it, and two classes of intermediate genes that would integrate both global and specific programs. We then study regulatory factors that act on the components of each element of the partition. We focus both on the influence of TFs and chromatin modifiers. Our results allow us to better appreciate the integration between the specific and global mechanisms controlling genome-wide expression in eukaryotes, as well as the genetic and epigenetic factors that contribute to this integration. More broadly, our study emphasizes the limitations of implementing a unique genome-wide program of expression control.

Results

A five-component partition captures large-scale changes of gene expression. We first examined the changes with growth rate in the promoter activity (PA) of a subset of ~900 *Saccharomyces cerevisiae* genes in 10 growing conditions (Keren et al., 2014). Studying absolute PA values Keren et al. presented a binary partition that helped explain these changes, recognizing a global proportional response in most genes and a specific one in a much smaller set. To focus on resource reallocation, we considered here instead fractional activities, i.e., the fraction of PA of each gene in a given nutrient condition [out of the summed activity of all genes in the dataset (Maaløe, 1979, Zaslaver et al., 2009)], and quantified their change for each pair of conditions (from low to high growth rate).

We interpreted these changes by delimiting a five-component partition of all genes (Methods). Figure 1A shows a descriptive case (glycerol to glucose growing conditions). Genes whose fractional PA remains approximately invariant (along the diagonal in Fig. 1A) constitute the first partition element. One can broadly introduce other four components: 1/positive genes (whose fractional PA moderately increases between conditions), 2/negative genes (fractional PA decreases to only a limited extent), 3/specifically activated genes (fractional PA becomes much larger), and 4/ specifically repressed genes (fractional PA becomes much smaller). Note that by reducing the partition to two components, specific and global, the stronger allocation of "expression resources" to specific genes in glycerol as compared to glucose becomes manifest (Fig. 1.B, white/grey pie charts). This leaves less resources to biosynthesis (reduction of the global component) affecting growth rate (Keren et al., 2014).

Our analysis identifies a fine-grained structure within global genes (invariant, positive and negative genes), each class following a precise proportional response between

conditions (Fig.S1). Within this structure, we also observed how the resources taken by positive genes in a condition of high growth rate limit in turn those available for the expression of negative genes (Fig. 1B, brown/red/blue pie charts). The resource reallocation is also revealed by the (absolute) PA response to growth that exhibit this type of genes, PA that cannot be simply explained in terms of the associated changes in growth rate (Maaløe, 1979) (Keren et al., 2014) (Fig. 1C).

With respect to the type of genes that comprise each component, the invariant class is enriched by transcription regulation and ribosomal proteins; the latter being more extensively observed in the positive class. Indeed, positive genes are enriched by ribosomal genes (~65% of genes code for small or large subunits of the ribosome), while negative genes are enriched in ATP metabolic processes, e.g., oxidative phosphorylation or the TCA cycle. Lastly, activated and repressed genes indicated specific expression programs related to the particular carbon source, and the pathways of the central metabolism that facilitate the transition between fermentation or respiration routes (Table S1, Methods).

The five-component partition can be delineated on a genomic scale. To generalize the preceding analysis to a genome-wide scale we analysed a DNA microarray dataset of yeast cells exhibiting a range of growth rates for several limiting nutrients (Brauer et al., 2008). We first applied singular value decomposition (SVD) to the fractional gene expression quantified on each nutrient separately (Methods). Notably, the first and second SVD components (Fig. 2A) explain >90% of the variance in each condition. As a result, fractional expression of each gene can be approximated by the linear combination of these two components (Fig. 2B), which exhibited an analogous trend in all nutrients (Fig. S2).

Moreover, while the first component ($\vec{v}_1$) does not change with growth rate, we observed that the second component ($\vec{v}_2$) exhibits a monotonic behaviour (Fig, 2A for glucose, and Fig. S2 for other nutrients). We can thus interpret the first element of the linear combination as the baseline fractional expression of the gene, and the second element as its monotonic behaviour with growth rate (Fig. 2B). This interpretation enables us to generalize the partition framework previously introduced with PA data.

Therefore, a change in the loading of $\vec{v}_1$ (a_i) in two different nutrients implies that the corresponding gene is specific, as it changes between conditions, and global otherwise (Fig. 2C) (note that $\vec{v}_1$ is quantitatively similar in all nutrients, Fig. S2). This supports the framework in (Keren et al., 2014) (Fig. 1B). Comparison of these gene loadings in the six nutrients revealed that they are fairly similar (minimal correlations found of ~0.96), i.e., much of the gene response is global.

In contrast, the second component ($\vec{v}_2$) provides a quantitative score (the second loading, b_i) to classify genes as invariant, positive or negative, as before (Fig. 2C, Methods). Some genes have the same classification in two different nutrients, but this does not have to be necessarily the case. With the use of the second component we can also evaluate how the response to growth rate depends on the exact nutrient (Brauer et al., 2008). We found that nutrient condition particularly matters in the range of slow growth, and that some nutrients trigger a more similar response with growth than others (Gutteridge et al., 2010). Finally, the functional analysis of genes within each class agrees with PA data and previous reports (Table S2).

The partition composition presents different transcriptional regulation. TFs are the most direct elements that can deviate the expression of genes from the global transcription program. Before examining this effect, we asked how TFs themselves are framed in the previous partition. After assembling a transcriptional regulatory network with existing data (Methods), we observed that most constituent TFs (122 of a total of 133 composing the network) exhibit similar basal fractional expression (a_i loadings) across all pairwise condition changes, i.e., they are global genes. Within this set, 31% presents a dominant invariant response (b_i~0 in >3 nutrients, of a total of 6), with five genes acting as invariant in all six conditions (*rsc1*, *mbp1*, *pho2*, *rgr1*, and *swi6*). Two of these (*mbp1*, *swi6*) are at the top of the network hierarchy (being involved in the mitotic cell cycle), and two are elements of relevant complexes that interact with RNA polymerase II (*rsc1* of the RSC chromatin complex, and *rgr1*/*med14* of the mediator complex); they can be considered as elements of a general transcriptional machinery, for which maintaining its concentration invariant across conditions could be essential. Moreover, 32% of global TFs are dominantly negative, and only 4% positive. Of note, some of the TFs whose expression decreases with growth (b_i<0) are positive regulators of transcription in response to stress (e.g., *bur6*, *gcn4*, *rpn4*) what justifies their overexpression at low growth rates.

To what extent is the regulation of target genes dependent on which component of the partition they belong to? We labelled target genes as global if they showed a global response (similar a_i loadings) in >8 pairwise change of conditions (total of 15). Genes are considered specific otherwise. Global genes are less regulated on average than specific ones [by 3.09 TFs *vs.* 5.06 TFs, $p = 1.20\ 10^{-4}$, two-sample Kolmogorov-Smirnov (KS) test]. Within global genes, we described as invariant –following again the second loading score, b_i – those which exhibit this pattern in >3 conditions. Global and invariant genes are less regulated on average than global and *not* invariant genes (by 2.56 TFs *vs.* 3.3 TFs, $p = 8.16\ 10^{-13}$, two-sample KS test). Finally, global and positive

genes are slightly more regulated than global and negative genes (by 3.33 TFs *vs.* 3. 27 TFs, p = 0.0018, two-sample KS test). Overall, specific genes are subjected to more regulation (larger number of TFs) (Fig. 3A), while global and invariant genes are the ones that show the least.

Specific TFs regulate genes belonging to each partition sector. Although Fig. 3A shows how the structure of transcriptional interactions is reflected differentially in the components of the partition it does not assure us when these interactions are active, e.g., (Berthoumieux et al., 2014). For this we examined several features.

We first inspected if target genes presenting a particular growth response are enriched by TFs showing the very same response, as similarity of the responses could imply that part of the regulatory structure is active. We thus computed –for each target gene– the fraction of its regulators that behave as negative, invariant, or positive (TF_{neg},TF_{inv},TF_{pos}, respectively) with growth rate in a given condition. Figure 3B shows the mean of the fractions for target genes whose response is negative, invariant or positive. Negative TFs are more likely to be found acting on target genes that are also negative (higher mean TF_{neg} on negative genes), while invariant (TF_{inv}) and positive (TF_{pos}) TFs regulate more often invariant and positive target genes, respectively (the latter signal is weaker and depends on the particular condition, Fig. S3). Thus, TFs that exhibit the same behaviour as their cognate target gene tend to appear, on average, dominant on its regulation; part of the regulatory structure seems then functional.

To further test the active effect of TFs, we measured the correlation of the response to growth rate between any particular gene and all its cognate TFs ("regulatory coherence", Methods). Specific genes showed stronger regulatory coherence than global ones (Fig. S4A), and remain coherent in more nutrient conditions (Fig. S4B), both results implying an active contribution of TFs to deviate gene expression from the

global program. Moreover, Fig. 3C shows those TFs whose action is particularly coherent *per* partition component (Methods). One can identify here two broad groups, which control global or specific genes [this is supported by earlier reports (Fazio et al., 2008)]. In this way, within the regulatory network we find TFs that act more significantly on different types of genes. Notably, those that work on global genes are higher up in the network hierarchy (Methods). We also noted that some these (significantly coherent) TFs are involved in chromatin remodelling (Cyc8, Ume6, Spt6, Msn4, Abf1, Msn2, Nhp6A, acting on global ones), or chromatin organization (Spt3, Spt2, Pho4, FKh2, Sin3, Spt20, Wtm2, Wtm1, Hif1, acting on specific genes). We examine epigenetic aspects next.

The partition composition also reveals distinctive epigenetic regulation. To inspect the function of epigenetic control mechanisms, we first quantified the proportion of general transcription factors (GTFs) found within the set of TFs acting on a given gene (Fourel et al., 2002). GTFs (Rap1, Abf1, Reb1, Cbf1, and Mcm1) usually have little intrinsic regulatory activity and together with the presence of chromatin remodelers (in particular, RSC –Remodels Structure of Chromatin) control an alleged general machinery of expression. We observed that GTFs constitute a larger and significant fraction in the regulation of positive genes, while the opposite is observed for negative ones (Fig. 4A).

GTFs are also connected to particularly fragile nucleosome promoter architectures (Xi et al., 2011), a connection recently examined (Kubik et al., 2015). Using this data, we computed the nucleosome landscape for the different gene classes (Methods). Promoters of positive genes are indeed enriched in fragile nucleosomes (Fig 4.B) while both negative and invariant genes typically lack these structures. This suggests that positive genes are more sensitive to deviate from the global program of expression (implemented by a general transcriptional machinery) by means of chromatin

modulation. Enrichment of other promoter features contribute to this model (Fig 3.C), like the absence of TATA boxes, the action of TFIID over SAGA [but this precise grouping has been recently revaluated (Baptista et al., 2017)], the presence of nucleosomal free regions closer to the transcriptional starting site (shNFRs) (partially associated to the previous score of fragile nucleosomes), and the dominant effect of *trans* variability (Fig. S5) (Choi and Kim, 2008).

We examined additional chromatin modifiers using a previously assembled compendium (Steinfeld et al., 2007). Figure 5 shows the effects of mutating different types of *trans*-acting chromatin regulators on the genes constituting the partitions. Note here that growth rate reduction can be connected to many of these deletions, so we controlled for the possible contribution of cell cycle population shifts as described (Methods). This enables us to better identify expression changes due to regulation (O'Duibhir et al., 2014). With the exception of histone acetyltransferases (HATs) and TAF-related factors (TAFs), the effect of most chromatin modifiers is dominant in specific genes (Fig. 5A, Methods). Within global genes we found three main configurations (Fig. 5B): 1) Epigenetic regulators acting as part of a general machinery (HATs –including SAGA–, TAFs and methyltransferases) whose mutation causes a general decrease in expression, very particularly in invariant and positive classes. Indeed, work by (Baptista et al., 2017) and (Warfield et al., 2017) demonstrated that SAGA and TFIID are recruited to pol II promoters genome-wide and that each complex is generally required for pol II transcription, i.e., its mutation would lead to genome-wide decrease of gene expression. 2) Regulators (histones, and chromatin remodelers) acting in a dual manner: increasing the expression of negative genes (remodeler as repressor) or reducing their expression in positives (remodeler as activator). This underlines the enrichment of negative and positive classes by stress and ribosomal genes, respectively, which are largely regulated in opposite manner (Bajić and Poyatos, 2012); a dual role of remodelers as activators and repressors have been

previously reported (Sudarsanam et al., 2000, Holstege et al., 1998). And 3) regulators as broad repressors that represent regulation by gene silencing.

Discussion

Could one interpret genome-wide expression changes as deviations from a global program of transcriptional control? In this work, we aim to answer this question by capturing these deviations in a five-component partition of the yeast genome. The analysis of relative expression values is necessary here as it helps us to appreciate expression reallocation among partition constituents. Therefore, this integrates an earlier model (Keren et al., 2014) that reduced variation between conditions to a proportional response shared by most genes and induced by a global transcriptional program (with only a limited number of specifically expressed genes), while also discriminating three subclasses within such response. Invariant genes, that best follow the global program, and positive and negative genes, which were broadly defined in other studies as growth-related genes (Regenberg et al., 2006, Castrillo et al., 2007, Brauer et al., 2008). The biological significance of our framework is reinforced by its differential promoter and regulatory architecture ranging from a model of almost passive control to one exhibiting complex combinatorial regulation.

More explicitly, invariant genes are those subjected to less regulation by TFs, regulation that increases among the rest of global classes, and between these and the specific ones. Specific genes also show a stronger regulatory coherence than global genes (similarity of expression response to that of the TFs acting on them). In addition, among TFs whose action is particularly coherent we identify two groups that almost exclusively regulate global or specific genes: the action of the TF network is somehow segregated. Beyond TF regulation, we can discriminate two broad promoter

architectures. Those that are TATA enriched and shNFR/TFIID depleted (global negative and specific genes), and those that are TATA depleted and shNFR/TFIID enriched (global invariant and positive genes). Notably, these features are similarly observed in metazoans promoters (Lenhard et al., 2012) (Type I promoters, genes expressed in a tissue-specific manner, and Type II promoters, ubiquitously expressed genes, respectively).

That (global) positive genes are moderately controlled by TFs (like negative) but depleted in TATA box (unlike negative) could suggest certain expression features (e.g., high level of transcription, Fig. S6) and alternative modes of regulation. Indeed, positive genes are enriched in fragile nucleosomes, which highlights the regulatory role of nucleosomal stability. This is supported by the particular action of GTFs on these genes [as GTFs fine-tune nucleosomal stability (Xi et al., 2011, Kubik et al., 2015)]. In addition, we find the expression of global genes being adjusted in distinctive manner by epigenetic modifiers, with three main configurations: 1/ HATs, TAFs and methyltransferases working as general activators of invariant and positive genes, 2/ Histone and chromatin remodelers working in a dual manner; repressors of negative genes and activators of positive ones, and 3/ Gene silencing elements acting as general repressors of negative genes.

The five-component partition in a broader context.

Some of the previous features discussed in our framework of the five-component partition match those observations related to environmental stress response genes (ESR) (Gasch et al., 2000), so it is interesting to examine how this set fits into our partition. ESR genes included two complementary subsets, which are enriched in our global negative and repressed genes (induced ESR genes), or global positive genes (repressed ESR genes, Methods). This confirms the suggestion of previous studies that stress response genes were not responding directly to stress but rather to the

associated decrease in growth rate. More generally, two models to coordinate gene expression to the available nutrients can be imagined: a feed-forward regulation by signalling pathways that predict growth rate in a certain environmental condition, or a feedback mechanism, which senses growth rate, or other related internal cell variable, and then modifies expression (Levy and Barkai, 2009). In this context, a passive resource allocation model could explain that the global program is responding always to the environment, although indirectly (as it can only use those resources that were not consumed in the mounting of the specific response). This validates, for instance, that ribosomal genes follow the feed-forward model (Levy et al., 2007). The fine-grained structure of the global class (invariant/positive/negative) could nevertheless monitor growth rate, at least partially, with the feedback being mediated by epigenetic mechanisms (see below).

If, as suggested by (Hansen and O'Shea, 2015), TFs can mostly transmit qualitative (presence/absence of a particular nutrient) rather than quantitative (amount of nutrient) information, how can we then explain the monotonic variation of fractional expression with nutrient dilution in the chemostat of the genes in the negative and positive partition components? One way is that metabolism, which is highly sensitive to the limiting nutrient (Boer et al., 2010) acts as regulator of the epigenetic factors discussed above. Indeed, several metabolites (e.g., GlcNAc, NAD+, acetyl-CoA, alpha KG, ATP) are known to regulate transcription through interactions with enzymes involved in epigenetic modifications (Lu and Thompson, 2012). For example, acetyl-CoA induces cell growth and proliferation by promoting the acetylation of histones at growth genes (Cai et al., 2011) (histone acetylation affects rather similarly specific and global genes, Fig. 5A, what supports its potential role as a widespread mechanism). Another explanation is that the monotonic variation observed is the result of cell population shifts with growth rate, instead of changes in single-cell resource allocations. Note that these shifts cannot be attributed to the fact that slow growing cells enlarged their G1

cell cycle phase as neither (Brauer et al. 2008) nor us observed a bias in positive/negative genes with any particular phase of cell cycled genes.

In this work we have studied changes in fractional expression but not in mRNA abundances. It is known that the global program dictates that the faster a population of cells growths, the higher the promoter activity (rate of RNA synthesis) (Keren et al., 2014) or total mRNA abundance (rate of RNA synthesis and degradation) (Athanasiadou et al., 2016). We expect most (if not all) gene products to follow this (absolute) global program, with potential additional layers of regulation (which are nutrient and gene dependent) that increment or decrement mRNA levels. The invariant group best describes the absolute global program, while positive genes are slightly above and negative genes slightly bellow this program (but all of them incrementing mRNA levels or promoter activities) (e.g., Fig. 1C). On the other hand, it would be interesting to quantify the degree to which single cells can present a distribution of resources that is separated from the model here discussed (Gasch et al., 2017), as well as to understand the mechanisms that lead to such divergence.

In summary, although one could argue that cellular physiology can indeed determine a global transcriptional program of gene expression control, our work highlights that this program is mediated by the integration of genetic and epigenetic modes of regulation, what limits the prospect of “simplifying” our understanding of genome-wide expression change.

Materials and Methods

Promoter activity (PA) data. (Keren et al., 2014) measured the activities of ~900 *S. cerevisiae* promoters in 10 different growing conditions using a library of fluorescent reporters. For each strain in every growth condition, promoter activity was obtained as the YFP production rate per OD per second in the window of maximal growth.

Genome partition based on PA data. Fractional promoter activity (f_{PA}) for each growth condition and ratios of f_{PA}s for each pair of conditions (with increasing growth rate) were calculated. We then computed the absolute distance of these ratios to ratio 1 (i.e., same f_{PA} in both conditions), and defined as invariant genes the top 350 genes (distance closest to 0) and as activated (repressed) the bottom 50 with ratio >0 (< 0). The rest of genes with ratio >0 (<0), and both f_{PA}s > 10^{-4}, were designated as global positive (negative). We used the "typical" class of a gene (the most frequently occurring category that a gene presents in all pairwise growth rate changes) to select the examples of Fig. 1C, and to characterize the partition in functional terms (Table S1).

Microarray data. (Brauer et al., 2008) grew yeast cultures in chemostats under different continuous culture conditions (six different limiting nutrients each at six dilution rates) and measured mRNA abundance with two-colour microarrays. Since the original reference channel for all samples corresponded to a particular glucose condition, which mixes the response of different nutrients, we reanalysed the data without this reference by considering the red processed signal as independent channel ('t Hoen, 2004), and normalizing by the corresponding sum for each case (to obtain a fractional score). SVDs were computed on this processed data.

Genome partition based on microarray data. Global genes are those whose difference on the loadings of the 1st component (a_i's) between two conditions is less, or equal, than three standard deviations of all gene differences (in absolute values). Genes are otherwise considered specific (activated or repressed if the difference of a_i's is positive or negative, respectively). Moreover, absolute values of the loadings of the

2nd component (b_i's) were sorted to define those with smallest values (top 2500) as invariant genes, with the rest being positive or negative (determined by the sign of b_i). To define the partition, we classify as global those genes that act as global in >8 pairwise conditions (out of 15). Global genes acting as invariant in > 3 conditions (recall that the total number is 6) are labelled as invariant. Global and not invariant genes appearing more times as positive than as negative (in all 6 conditions) are categorized as positive, and likewise for negative. Specific genes which appear more times as positive than as negative (again in all 6 conditions) are categorized as activated, and analogously for repressed.

Regulatory network. We obtained regulatory data from http://yeastmine.yeastgenome.org. No microarray data is considered for the TF info; only data from different manuscripts using chromatin immunoprecipitation, chromatin immunoprecipitation-chip, chromatin immunoprecipitation-seq, combinatorial evidence, and computational combinatorial evidence for a total of 20,673 interactions with 133 TFs involved (Brauer et al., 2008). We also calculated the hierarchical organization of the network (Jothi et al., 2009). Bas1, Mbp1, Med6, Spt7 and Swi6 appear at the top of the hierarchy.

Fragile nucleosome data. Nucleosome occupancy and position have been measured by analysis of MNase-digested chromatin. Recent work noted that certain nucleosomes were extremely sensitive to this digestion, and thus obtained a quantitative score of nucleosome fragility that we used for our analysis, Table S6 in (Kubik et al., 2015).

Chromatin compendium. This set includes 170 gene expression profiles for chromatin-regulation related mutations (expressed in $\log_2$ ratios) taken from 26 different publications (Steinfeld et al., 2007). It covers more than 60 potential interacting chromatin modifiers such as histone acetyltransferases (HATs; the NuA4, HAT1 and SAGA complexes), histone deacetylases (HDACs; the RPD3, HDA1 and SET3 complexes), histone methyltransferases (the COMPASS complex), ATP-dependent chromatin remodelers (the SWI/SNF, SWR1, INO80, ISWI and RSC complexes), and

other chromatin-affecting genes and cofactors such as Spt10, Sir proteins and the TATA-binding protein (TBP). We normalized each dataset to unit variance (Choi and Kim, 2008). For Fig. 5A, we took absolute values to estimate the strength of the chromatin regulator effect.

Regulatory coherence. We identified the set of TFs regulating each gene, and quantified the Pearson's correlation coefficient between the expression vector (as function of growth rate) of each TF within the set and the target gene to then take the mean. This is the (mean) regulatory coherence in a given nutrient condition. Randomizing expression vectors for each gene, 1000 times, we obtained a score of significance for each gene's regulatory coherence. With this, we identified a list of genes displaying significant regulatory coherence. Identification of TFs acting more significantly on each partition component is computed by first measuring how often it acts on significantly coherent genes, within the five-component partition grouping, and then estimating a null value by randomization of the partition classes.

Removal of the slow growth signature. We took the full data in (Kemmeren et al., 2014) to obtain the slow growth profile and remove the slow growth signature in the epigenetic data following (O'Duibhir et al., 2014). In brief, the slow growth profile is obtained as the first-mode approximation of the data after SVD decomposition. To compare with the epigenetic compendium data, we chose the column of this approximation with the largest norm as the slow growth signature. The correlation with the slow growth signature is removed by transforming the epigenetic data in Gram-Schmidt fashion by subtracting from their projection onto the basis vector given by the normalized slow growth profile.

ESR genes. There are 281 stress-induced and 585 stress-repressed genes –as defined in (Gasch et al., 2000)– within the set of genes delineating the five-component partition. A subset of global negative genes and specific repressed genes corresponds to stress-induced (232 out of 2053, and 10 out of 70, respectively), while a subset of global positive genes corresponds to stress-repressed (485 out of 1914). Note that

most of the features discussed in the main text associated to the five-component partition remain when controlling for ESR genes.

Acknowledgments

This work was supported by grant FIS2016-78781-R from the Spanish Ministerio de Economía y Competitividad (J.F.P.). We thank John C. Matese for his help with the analysis of the microarray data.

References

Athanasiadou, N., Neymotin, B., Brandt, N., Miller, D., Tranchina, D., Gresham, D., 2016. Growth Rate-Dependent Global Amplification of Gene Expression. bioRxiv https://doi.org/10.1101/044735.

Bajić, D., Poyatos, J.F., 2012. Balancing noise and plasticity in eukaryotic gene expression. BMC Genomics 13, 343.

Baptista, T., Grünberg, S., Minoungou, N., Koster, M.J.E., Timmers, H.T.M., Hahn, S., Devys, D., Tora, L., 2017. SAGA Is a General Cofactor for RNA Polymerase II Transcription. Molecular Cell 68, 130-143.e5.

Berthoumieux, S., de Jong, H., Baptist, G., Pinel, C., Ranquet, C., Ropers, D., Geiselmann, J., 2014. Shared control of gene expression in bacteria by transcription factors and global physiology of the cell. Molecular Systems Biology 9, 634–634.

Boer, V.M., Crutchfield, C.A., Bradley, P.H., Botstein, D., Rabinowitz, J.D., 2010. Growth-limiting Intracellular Metabolites in Yeast Growing under Diverse Nutrient Limitations. Molecular Biology of the Cell 21, 198–211.

Bosdriesz, E., Molenaar, D., Teusink, B., Bruggeman, F.J., 2015. How fast-growing bacteria robustly tune their ribosome concentration to approximate growth-rate maximization. FEBS Journal 282, 2029–2044.

Brauer, M.J., Huttenhower, C., Airoldi, E.M., Rosenstein, R., Matese, J.C., Gresham, D., Boer, V.M., Troyanskaya, O.G., Botstein, D., 2008. Coordination of Growth Rate, Cell Cycle, Stress Response, and Metabolic Activity in Yeast. Molecular Biology of the Cell 19, 352–367.

Cai, L., Sutter, B.M., Li, B., Tu, B.P., 2011. Acetyl-CoA Induces Cell Growth and Proliferation by Promoting the Acetylation of Histones at Growth Genes. Molecular Cell 42, 426–437.

Castrillo, J.I., Zeef, L.A., Hoyle, D.C., Zhang, N., Hayes, A., Gardner, D.C., Cornell, M.J., Petty, J., Hakes, L., Wardleworth, L., Rash, B., Brown, M., Dunn, W.B., Broadhurst, D., O'Donoghue, K., Hester, S.S., Dunkley, T.P., Hart, S.R., Swainston, N., Li, P., Gaskell, S.J., Paton, N.W., Lilley, K.S., Kell, D.B., Oliver, S.G., 2007. Growth control of the eukaryote cell: a systems biology study in yeast. Journal of Biology 6, 4.

Choi, J.K., Kim, Y.-J., 2008. Epigenetic regulation and the variability of gene expression. Nature Genetics 40, 141–147.
Dennis, P.P., Ehrenberg, M., Bremer, H., 2004. Control of rRNA Synthesis in Escherichia coli: a Systems Biology Approach. Microbiology and Molecular Biology Reviews 68, 639–668.
Fazio, A., Jewett, M.C., Daran-Lapujade, P., Mustacchi, R., Usaite, R., Pronk, J.T., Workman, C.T., Nielsen, J., 2008. Transcription factor control of growth rate dependent genes in Saccharomyces cerevisiae: A three factor design. BMC Genomics 9, 341.
Fourel, G., Miyake, T., Defossez, P.-A., Li, R., Gilson, É., 2002. General Regulatory Factors (GRFs) as Genome Partitioners. Journal of Biological Chemistry 277, 41736–41743.
Gasch, A.P., Spellman, P.T., Kao, C.M., Carmel-Harel, O., Eisen, M.B., Storz, G., Botstein, D., Brown, P.O., 2000. Genomic Expression Programs in the Response of Yeast Cells to Environmental Changes□D. Molecular Biology of the Cell 11, 17.
Gasch, A.P., Yu, F.B., Hose, J., Escalante, L.E., Place, M., Bacher, R., Kanbar, J., Ciobanu, D., Sandor, L., Grigoriev, I.V., Kendziorski, C., Quake, S.R., McClean, M.N., 2017. Single-cell RNA sequencing reveals intrinsic and extrinsic regulatory heterogeneity in yeast responding to stress. PLOS Biology 15, e2004050.
Gerosa, L., Kochanowski, K., Heinemann, M., Sauer, U., 2014. Dissecting specific and global transcriptional regulation of bacterial gene expression. Molecular Systems Biology 9, 658–658.
Gutteridge, A., Pir, P., Castrillo, J.I., Charles, P.D., Lilley, K.S., Oliver, S.G., 2010. Nutrient control of eukaryote cell growth: a systems biology study in yeast. BMC Biology 8, 68.
Hansen, A.S., O'Shea, E.K., 2015. Limits on information transduction through amplitude and frequency regulation of transcription factor activity. eLife 4.
Holstege, F.C.., Jennings, E.G., Wyrick, J.J., Lee, T.I., Hengartner, C.J., Green, M.R., Golub, T.R., Lander, E.S., Young, R.A., 1998. Dissecting the Regulatory Circuitry of a Eukaryotic Genome. Cell 95, 717–728.
Hui, S., Silverman, J.M., Chen, S.S., Erickson, D.W., Basan, M., Wang, J., Hwa, T., Williamson, J.R., 2015. Quantitative proteomic analysis reveals a simple strategy of global resource allocation in bacteria. Molecular Systems Biology 11, e784–e784.
Jothi, R., Balaji, S., Wuster, A., Grochow, J.A., Gsponer, J., Przytycka, T.M., Aravind, L., Babu, M.M., 2009. Genomic analysis reveals a tight link between transcription factor dynamics and regulatory network architecture. Molecular Systems Biology 5.
Kemmeren, P., Sameith, K., van de Pasch, L.A.L., Benschop, J.J., Lenstra, T.L., Margaritis, T., O'Duibhir, E., Apweiler, E., van Wageningen, S., Ko, C.W., van Heesch, S., Kashani, M.M., Ampatziadis-Michailidis, G., Brok, M.O., Brabers, N.A.C.H., Miles, A.J., Bouwmeester, D., van Hooff, S.R., van Bakel, H., Sluiters, E., Bakker, L.V., Snel, B., Lijnzaad, P., van Leenen, D., Groot Koerkamp, M.J.A., Holstege, F.C.P., 2014. Large-Scale Genetic Perturbations Reveal Regulatory Networks and an Abundance of Gene-Specific Repressors. Cell 157, 740–752.
Keren, L., Zackay, O., Lotan-Pompan, M., Barenholz, U., Dekel, E., Sasson, V., Aidelberg, G., Bren, A., Zeevi, D., Weinberger, A., Alon, U., Milo, R., Segal, E., 2014. Promoters maintain their relative activity levels under different growth conditions. Molecular Systems Biology 9, 701–701.
Klumpp, S., Hwa, T., 2014. Bacterial growth: global effects on gene expression, growth feedback and proteome partition. Current Opinion in Biotechnology 28, 96–102.

Kochanowski, K., Gerosa, L., Brunner, S.F., Christodoulou, D., Nikolaev, Y.V., Sauer, U., 2017. Few regulatory metabolites coordinate expression of central metabolic genes in *Escherichia coli*. Molecular Systems Biology 13, 903.
Kubik, S., Bruzzone, M.J., Jacquet, P., Falcone, J.-L., Rougemont, J., Shore, D., 2015. Nucleosome Stability Distinguishes Two Different Promoter Types at All Protein-Coding Genes in Yeast. Molecular Cell 60, 422–434.
Lenhard, B., Sandelin, A., Carninci, P., 2012. Metazoan promoters: emerging characteristics and insights into transcriptional regulation. Nature Reviews Genetics 13, 233–245.
Levy, S., Barkai, N., 2009. Coordination of gene expression with growth rate: A feedback or a feed-forward strategy? FEBS Letters 583, 3974–3978.
Levy, S., Ihmels, J., Carmi, M., Weinberger, A., Friedlander, G., Barkai, N., 2007. Strategy of Transcription Regulation in the Budding Yeast. PLoS ONE 2, e250.
Liang, S.-T., Bipatnath, M., Xu, Y.-C., Chen, S.-L., Dennis, P., Ehrenberg, M., Bremer, H., 1999. Activities of constitutive promoters in Escherichia coli. Journal of Molecular Biology 292, 19–37.
Lu, C., Thompson, C.B., 2012. Metabolic Regulation of Epigenetics. Cell Metabolism 16, 9–17.
Maaløe, O., 1979. Regulation of the Protein-Synthesizing Machinery—Ribosomes, tRNA, Factors, and So On, in: Goldberger, R.F. (Ed.), Biological Regulation and Development. Springer US, Boston, MA, pp. 487–542.
Metzl-Raz, E., Kafri, M., Yaakov, G., Soifer, I., Gurvich, Y., Barkai, N., 2017. Principles of cellular resource allocation revealed by condition-dependent proteome profilingAbstract. eLife 6, e28034.
Neidhardt, F.C., Magasanik, B., 1960. Studies on the role of ribonucleic acid in the growth of bacteria. Biochimica et Biophysica Acta 42, 99–116.
O'Duibhir, E., Lijnzaad, P., Benschop, J.J., Lenstra, T.L., van Leenen, D., Groot Koerkamp, M.J., Margaritis, T., Brok, M.O., Kemmeren, P., Holstege, F.C., 2014. Cell cycle population effects in perturbation studies. Molecular Systems Biology 10, 732–732.
Regenberg, B., Grotkj, T., Åkesson, M., Bro, C., Hansen, L.K., Brunak, S., Nielsen, J., 2006. Growth-rate regulated genes have profound impact on interpretation of transcriptome profiling in Saccharomyces cerevisiae. Genome Biology 13.
Schaechter, M., MaalOe, O., Kjeldgaard, N.O., 1958. Dependency on Medium and Temperature of Cell Size and Chemical Composition during Balanced Growth of Salmonella typhimurium. Journal of General Microbiology 19, 592–606.
Scott, M., Gunderson, C.W., Mateescu, E.M., Zhang, Z., Hwa, T., 2010. Interdependence of Cell Growth and Gene Expression: Origins and Consequences. Science 330, 1099–1102.
Steinfeld, I., Shamir, R., Kupiec, M., 2007. A genome-wide analysis in Saccharomyces cerevisiae demonstrates the influence of chromatin modifiers on transcription. Nature Genetics 39, 303–309.
Sudarsanam, P., Iyer, V.R., Brown, P.O., Winston, F., 2000. Whole-genome expression analysis of snf/swi mutants of Saccharomyces cerevisiae. Proceedings of the National Academy of Sciences 97, 3364–3369.
't Hoen, P.A.C., 2004. Intensity-based analysis of two-colour microarrays enables efficient and flexible hybridization designs. Nucleic Acids Research 32, 41e – 41.
Warfield, L., Ramachandran, S., Baptista, T., Devys, D., Tora, L., Hahn, S., 2017. Transcription of Nearly All Yeast RNA Polymerase II-Transcribed Genes Is Dependent on Transcription Factor TFIID. Molecular Cell 68, 118-129.e5.

Xi, Y., Yao, J., Chen, R., Li, W., He, X., 2011. Nucleosome fragility reveals novel functional states of chromatin and poises genes for activation. Genome Research 21, 718–724.

Zaslaver, A., Kaplan, S., Bren, A., Jinich, A., Mayo, A., Dekel, E., Alon, U., Itzkovitz, S., 2009. Invariant Distribution of Promoter Activities in Escherichia coli. PLoS Computational Biology 5, e1000545.

Figures

Fig. 1

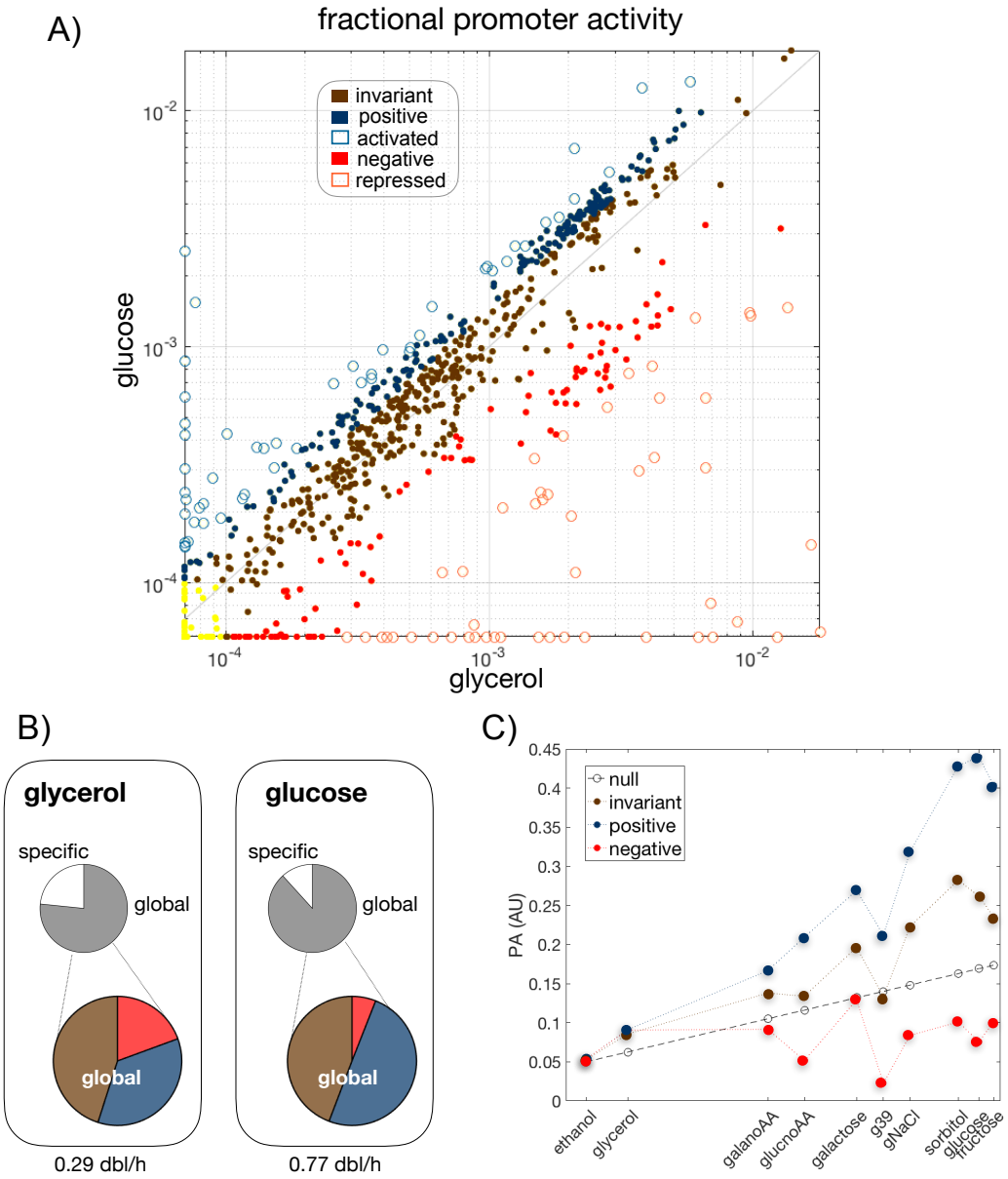

Figure 1. Partition of gene expression in five sectors. A) Fractional promoter activity (fPA) between two example conditions. Promoters can be classified into five categories (inset) depending on how their fPA changes (yellow dots indicate those with very low activity in both conditions). **B)** Repressed and activated promoters constitute the specific response, whose fPA is greater the lower the growth rate (at the cost of the promoters changing in a global manner). Global promoters are constituted by one invariant type and two other subclasses whose fPA depends on the growing condition. Note how the portion of expression, within global genes, of positive genes increases with growth rate, while it decreases for negative ones (colours as in A). **C)** Absolute promoter activity (PA) response of a typical invariant, positive and negative gene that corresponds to the *mrs11*, *rps6A* and *atp5*, respectively (conditions sorted by increasing growth rate). A null model of the dependence of PA with growth rate is given by the ratio of growth rates (empty circles). Gene categories within the global group clearly separate from the null (see main text).

Fig. 2

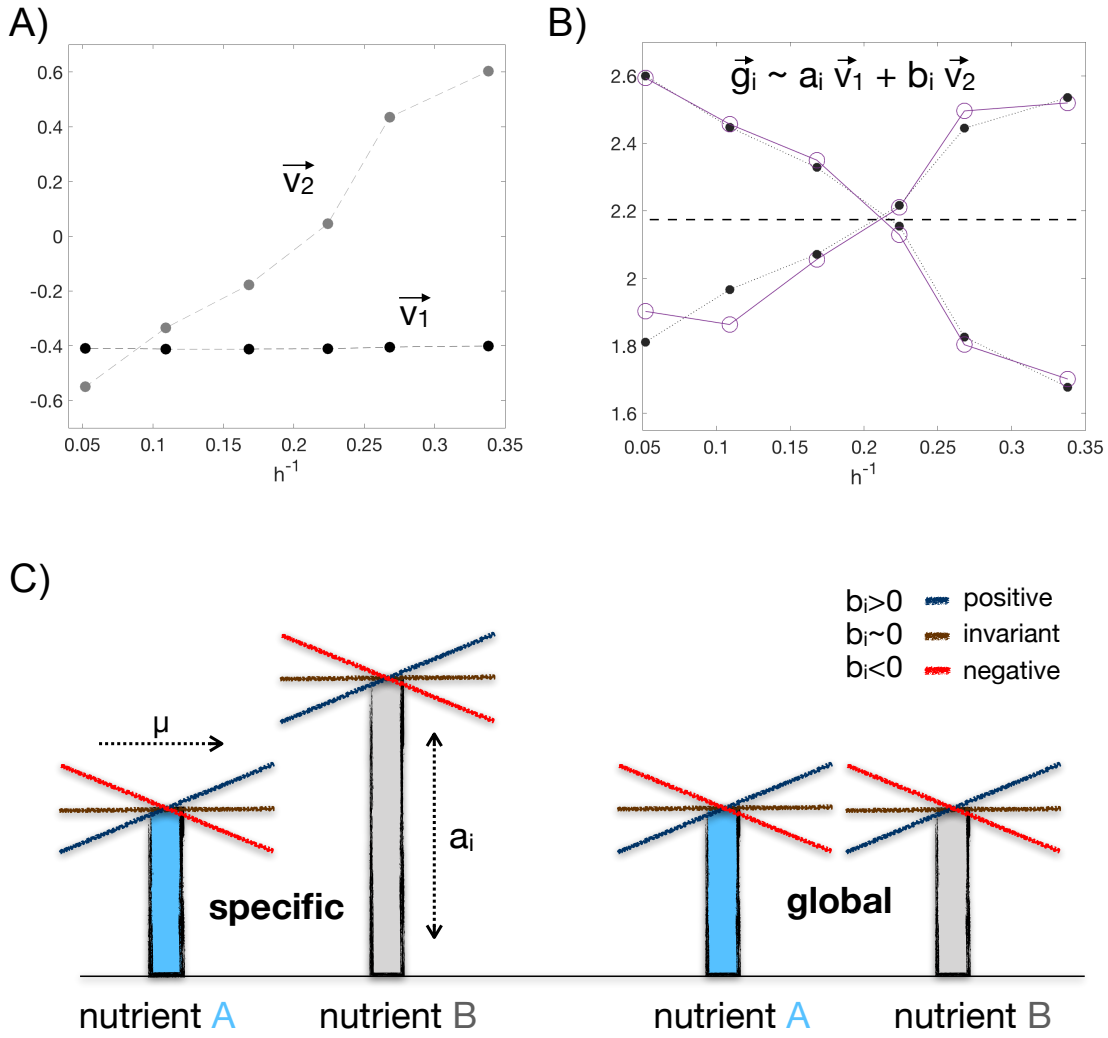

Figure 2. Genome-wide partition of gene expression in five sectors based on SVD. A) SVD components $\vec{v}_1$ and $\vec{v}_2$ describe baseline fractional expression and dependence with growth, respectively, and together explain most of the expression variance. **B)** The fractional expression of every gene $\vec{g}_i$ as a function of growth rate can be approximated by a linear combination of these two components, with loadings a_i and b_i. We show two examples (purple circles denote the expression vector, while the black dots correspond to the two-component approximation; lines added to help visualization) with the same baseline (dashed line; same a_i) but whose expression increases ($b_i>0$) or decreases ($b_i<0$) with growth. Data in A) and B) corresponds to growth in limiting glucose conditions. **C)** A given gene can be considered specific or global when its baseline fractional expression (bars) changes (left) or does not change (right) with different nutrients (different or comparable a_i, respectively). Beyond the baseline value on each condition, expression can increase, decrease or remain constant with growth-rate (μ) change. We thus compare a_i's between pairwise conditions (total of 15) to define a gene as specific or global, whereas the second loadings b_i's on each condition (total of 6) enable us to determine if the gene is positive, invariant or negative.

Fig. 3

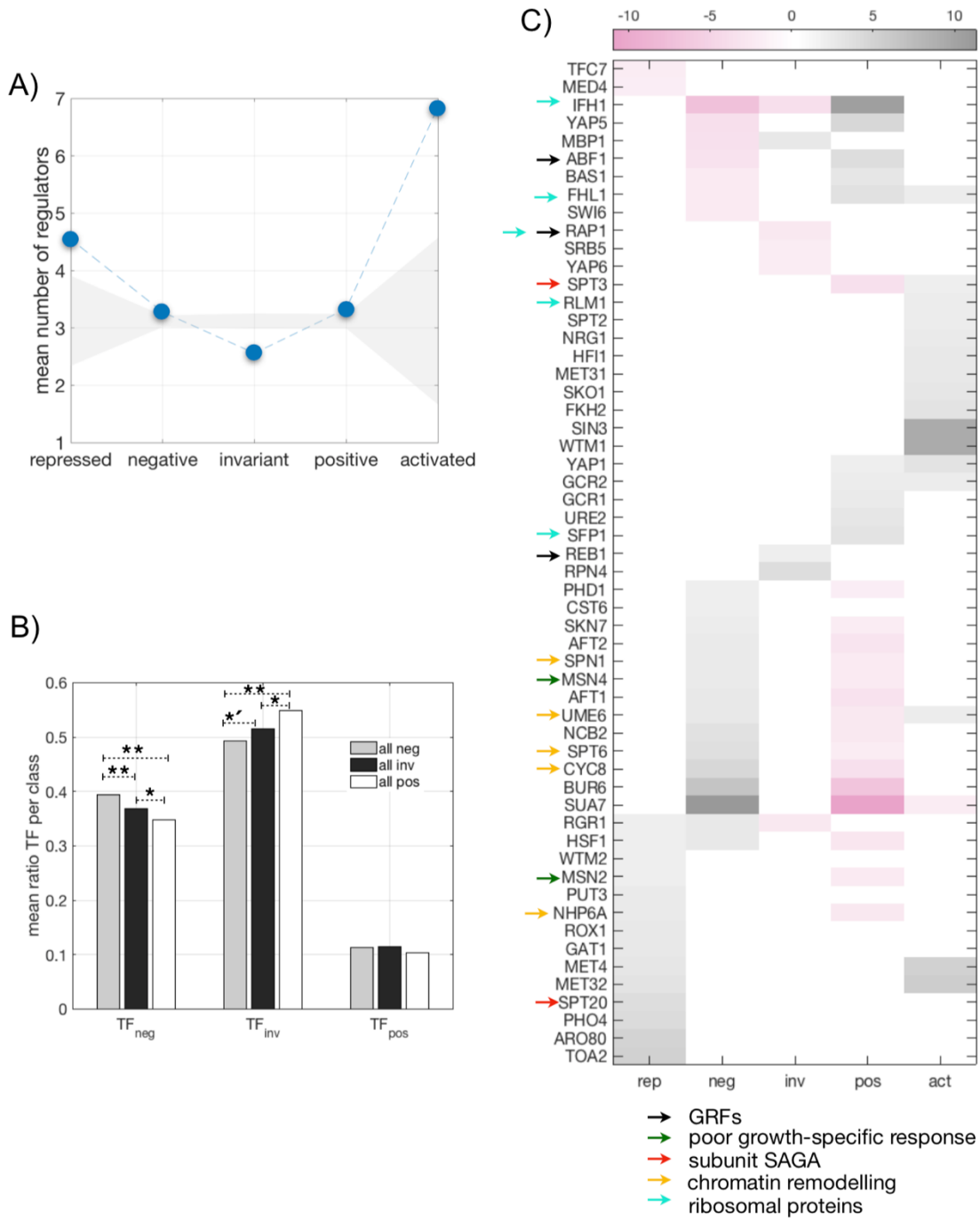

Figure 3. Genes within each partition sector show differential transcriptional regulation. **A)** Mean number of regulators acting on genes as a function of their response to growth (blue dots; dashed line to help visualization). Grey shading denotes the average null values +/- 2 standard deviations obtained by randomization. **B)** Mean ratio of fraction of TFs of a given class with respect to growth (e.g., TF_{neg} denotes TFs which are negative genes) for each group of target genes (also for a given class; here we do not distinguish between global and specific). Histogram obtained in glucose conditions, see also Fig. S5 for other nutrients (** $p < 0.001$, *´ $p < 0.01$, * $p < 0.05$, two-sided KS test). **C)** Regulators that act dominantly in genes showing a significant regulatory coherence. Colour denotes z-score with respect to a null obtained by randomization (positive values denoting enrichment). Properties of some regulators are also included (arrows). See main text for details.

Fig. 4

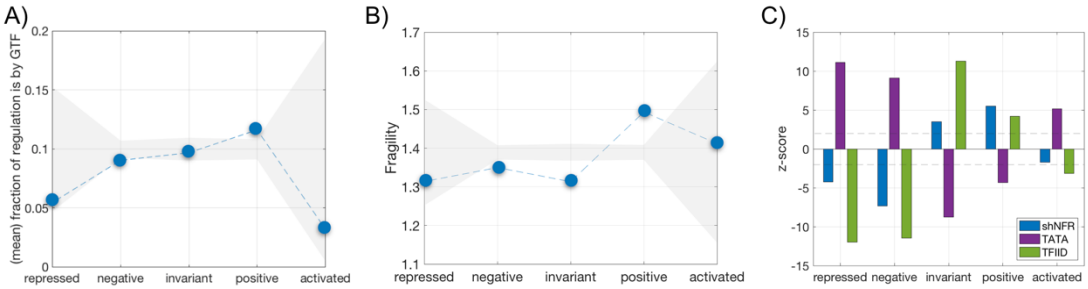

Figure 4. Genes within each partition sector show differential epigenetic regulation. A) Mean fraction of global transcriptional regulators (Rap1, Reb1, Cbf1 and Mcm1) within the full set of regulators acting on each gene. Grey shading denotes the average null values +/- 2 standard deviations obtained by randomization. Dashed lines to help visualization. **B)** Mean nucleosomal fragility. Shading/lines as before. **C)** Enrichment of nucleosomal free regions (shNFR, blue), presence of TATA boxes (pink), or action of TFIID global factor (green) as function of response class (measured as z-score with respect to a null by randomization; dashed line indicates z-score = +/- 2).

Fig. 5

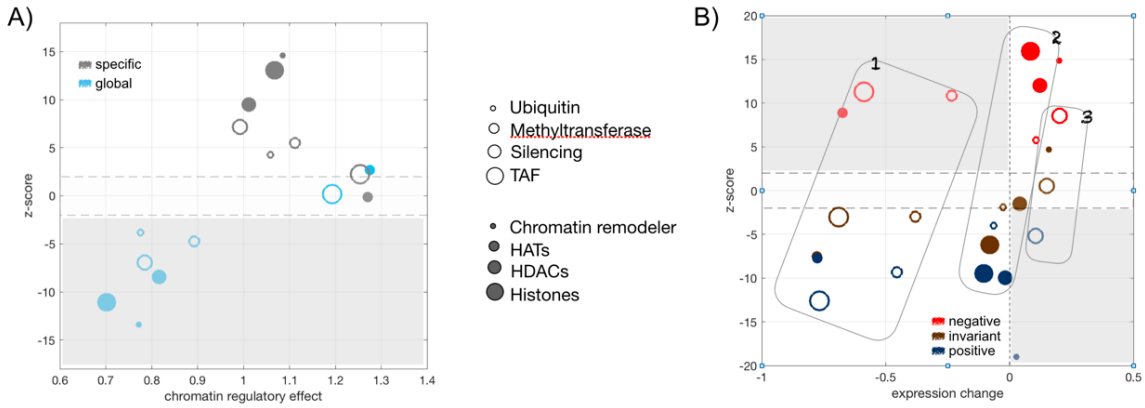

Figure 5. Chromatin modifiers act differentially on genes within each partition. A) The chromatin regulatory effect quantifies change in gene expression (absolute value) due to mutations in chromatin modifiers. Most modifiers show a significant effect on specific genes as compared to a null (obtained by randomization; grey shading indicates non-significant values, and dashed lines denote z-scores of +/-2). **B**) Mutations in chromatin modifiers reveal their diverse regulatory role when acting on negative, invariant, or positive genes. We observed three main categories: 1/ modifiers acting as activators (mutation decreases significantly the expression of invariant/positive genes), 2/ dual activator/repressor (mutation decreasing/increasing significantly expression of positive/invariant or negative genes, respectively), and 3/ repressors (mutation increasing significantly expression of negative genes). Significance and shading as before (note that a significant increase or decrease in expression correspond to positive or negative z-scores, respectively). Different type of modifiers corresponds to filled/empty circles of different size, whose colour denote the precise target gene class considered.

Supplementary Information

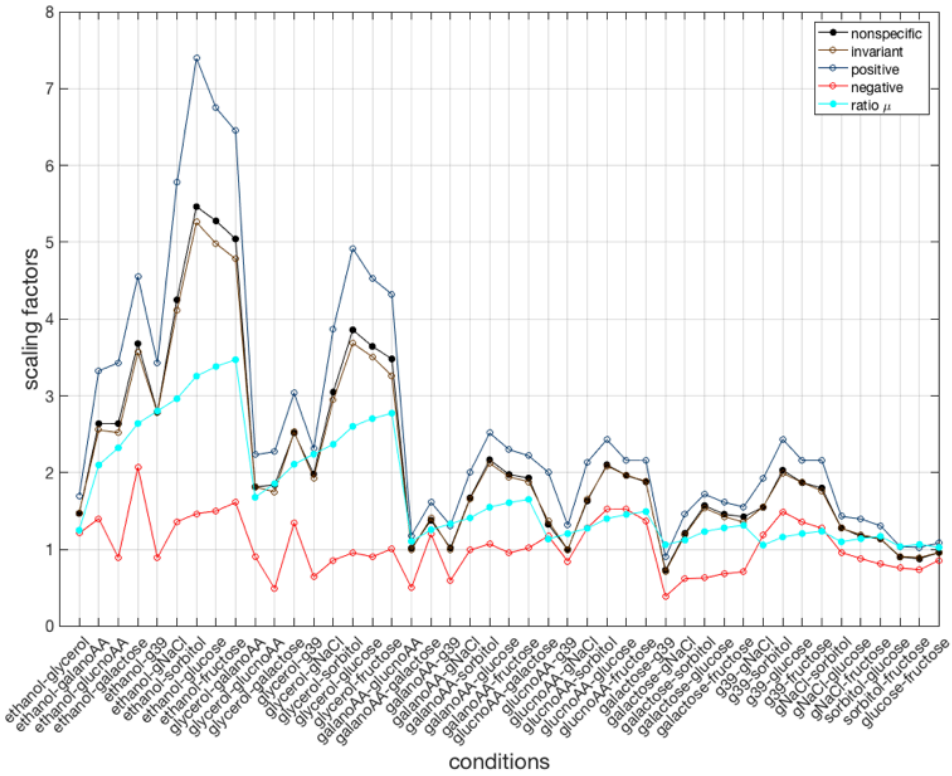

Figure S1. Scaling factors between pairs of conditions. A single proportionality (scaling) factor describes the change of promoter activity for different subsets of promoters according to the five-component partition. The three classes within the global promoters (invariant, negative, positive) clearly show different scaling. Shown also a null that correspond to the ratio of growth rates between conditions.

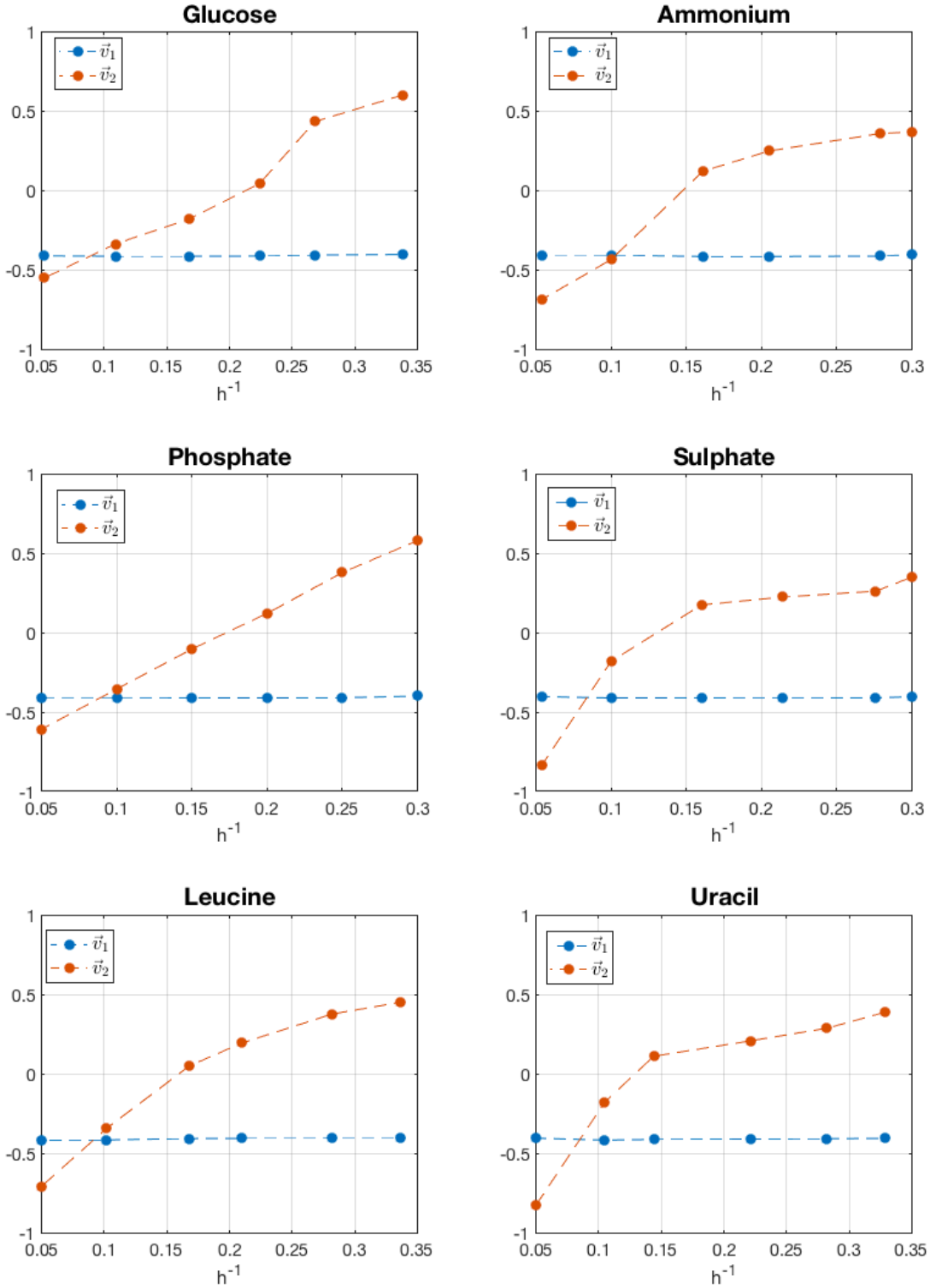

Figure S2. SVD components. First and second SVD components exhibited an analogous trend in all conditions what underlines a core response. As a result, expression of each gene can be approximated by the linear combination of these two components on each nutrient.

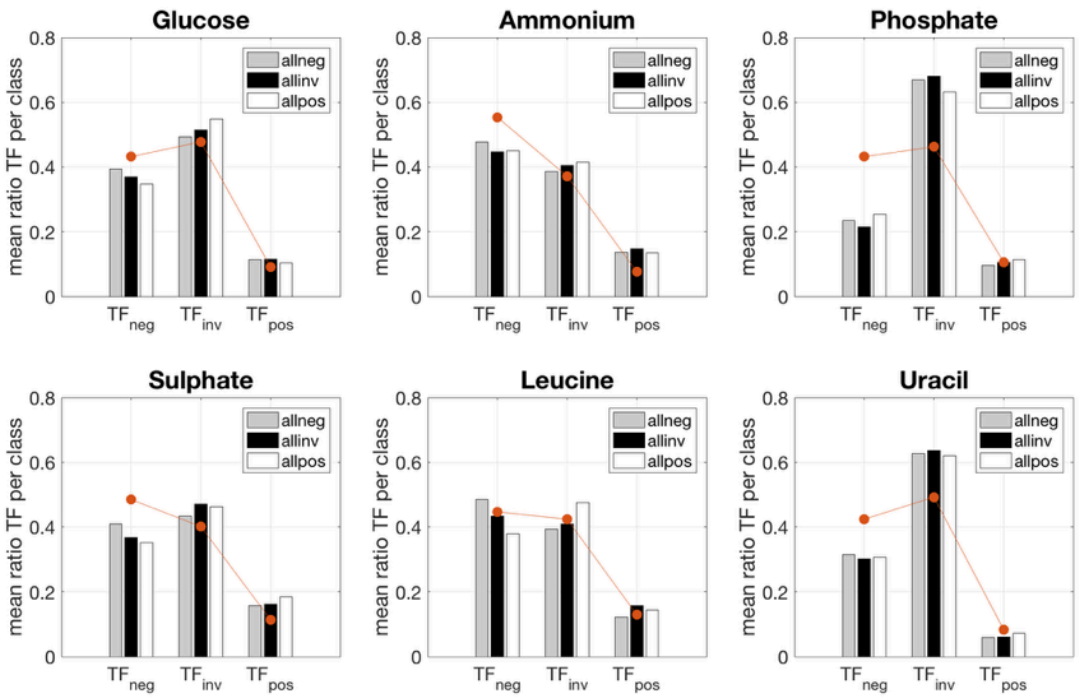

Figure S3. Association between class of TF and class of cognate target gene for all nutrients. Fraction of TF class (negative/invariant/positive) acting on target genes divided also with respect to growth response (negative/invariant/positive; global and specific genes were included that we denoted as allneg, etc.). Mean values of each grouping are shown in bars, while the orange curves show the distribution of each class of TF on each condition.

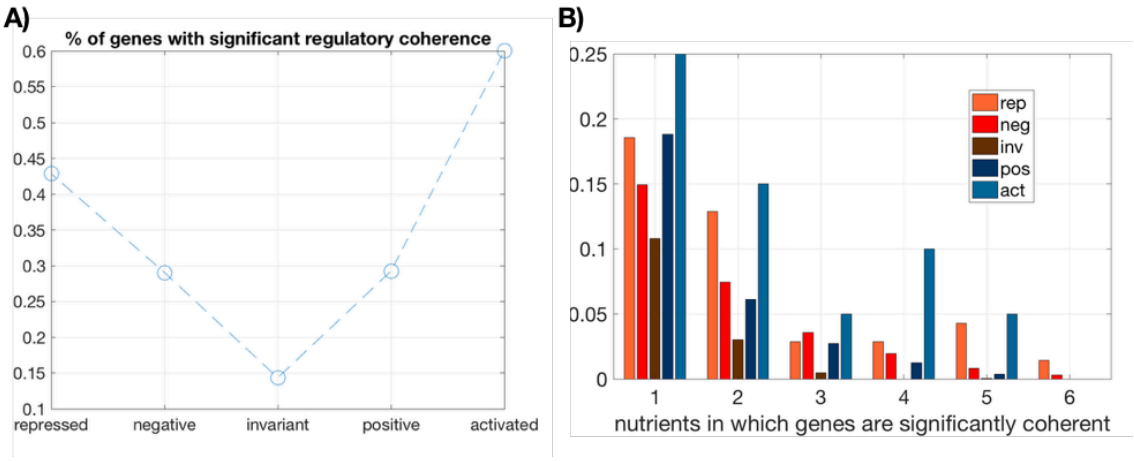

Figure S4. Regulatory coherence and five-component partition. A) Percentage of genes within each class that exhibit a significant regulatory coherence with respect to a null in which classes were assigned randomly (1000 randomizations; significance implies z-scores > 2). Note that invariant genes show minimal coherence. B) Distribution of genes exhibiting significant regulatory coherence in 1 to 6 different nutrient conditions. Specific genes show more cases of genes significantly coherent in more different conditions, while invariant genes showed the opposite.

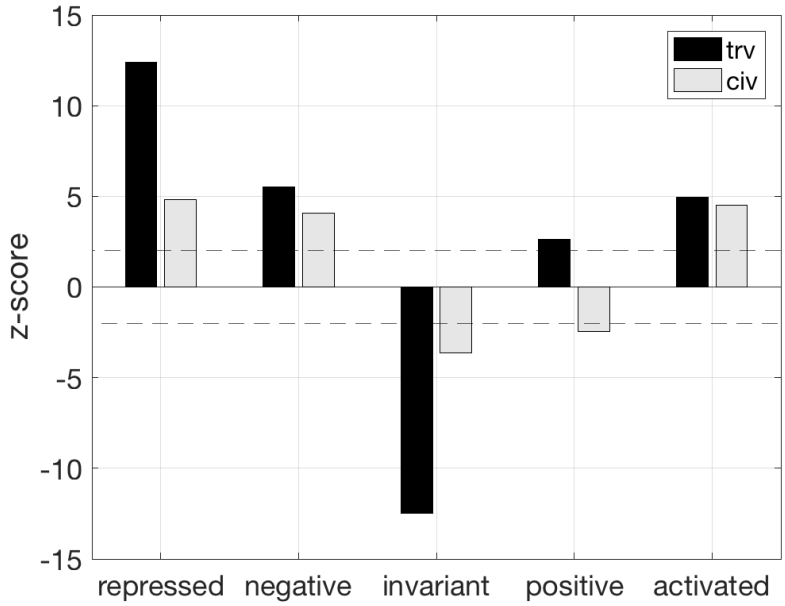

Figure S5. *Cis* and *trans* variability with respect to the five-component partition. A cross between a standard laboratory yeast strain and a wild isolate allowed the computation of *cis* and *trans* effects on transcriptional variance (Choi and Kim, 2008). For each partition, we quantified the mean of these measures and showed the associated z-score with respect to a null by randomization; dashed line indicates z-score = +/- 2. Positive genes show dominant effects associated to *trans* variability (trv and civ denote *trans* and *cis* variability, respectively).

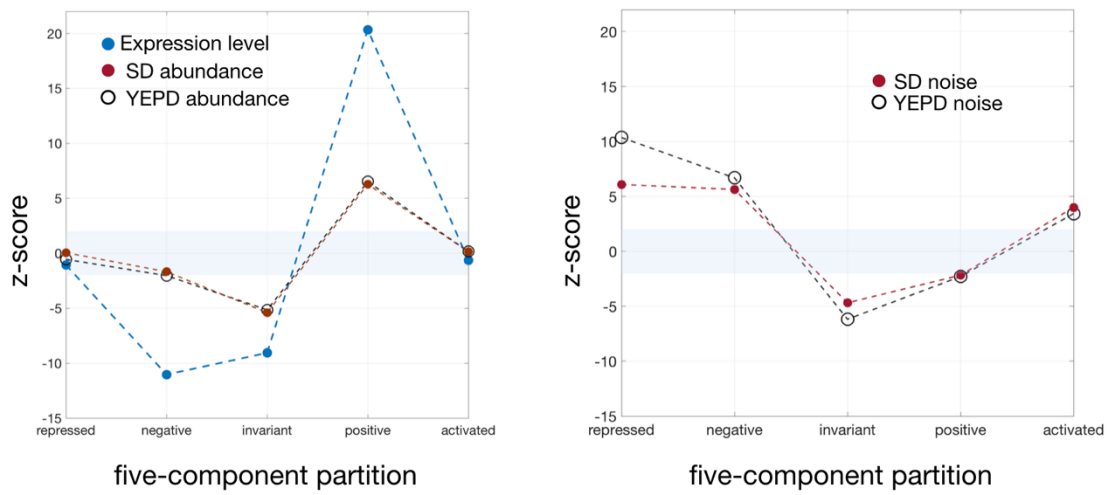

Figure S6. Expression abundance and noise with respect to the five-component partition. Mean expression and protein abundance (A) and protein noise (B) with respect to the five-component partition as compared to a null in which classes were assigned randomly (10000 randomizations; y-axis is plotting the associated z-score, shading correspond to z-score values within a range of -/+ 2; SD/YEPD denote poor/rich growing conditions). Global and positive genes showed high expression and low noise, a signal that was associated to the presence of fragile nucleosomes in the promoter and the action of general transcription factors [both enriched in global positive genes, see main text and (Bajić and Poyatos, 2012*)* for details on data].