Longitudinal network theory approaches identify crucial factors affecting sporulation efficiency in yeast

Camellia Sarkar[1,7,9], Saumya Gupta[2,7], Rahul Kumar Verma[1], Himanshu Sinha[3,4,5,8], Sarika Jalan[1,6,8]

[1]Centre for Biosciences and Biomedical Engineering, Indian Institute of Technology Indore, Khandwa Road, Simrol, Indore 453552, India

[2]Department of Cancer Biology, Dana Farber Cancer Institute, Department of Genetics, Harvard Medical School, Boston, MA 02215, USA

[3]Department of Biotechnology, Bhupat and Jyoti Mehta School of Biological Sciences, Indian Institute of Technology Madras, Chennai 600036, India

[4]Initiative for Biological Systems Engineering (IBSE), Indian Institute of Technology Madras, Chennai 600036, India

[5]Robert Bosch Centre for Data Sciences and Artificial Intelligence (RBCDSAI), Indian Institute of Technology Madras, Chennai 600036, India

[6]Complex Systems Lab, Discipline of Physics, Indian Institute of Technology Indore, Khandwa Road, Simrol, Indore 453552, India

[7]These authors contributed equally to this paper.

[8]Corresponding authors: Himanshu Sinha (sinha@iitm.ac.in) and Sarika Jalan (sarikajalan9@gmail.com)

[9]Current address: Institute of Biological Psychiatry, Mental Health Centre Sct. Hans, Roskilde, DK-4000, Denmark

ABSTRACT

Integrating network theory approaches over longitudinal genome-wide gene expression data is a robust approach to understand the molecular underpinnings of a dynamic biological process. Here, we performed a network-based investigation of longitudinal gene expression changes during sporulation of a yeast strain, SK1. Using global network attributes, viz. clustering coefficient, degree distribution of a node, degree-degree mixing of the connected nodes and disassortativity, we observed dynamic changes in these parameters indicating a highly connected network with inter-module crosstalk. Analysis of local attributes, such as clustering coefficient, hierarchy, betweenness centrality and Granovetter's weak ties showed that there was an inherent hierarchy under regulatory control that was determined by specific nodes. Biological annotation of these nodes indicated the role of specifically linked pairs of genes in meiosis. These genes act as crucial regulators of sporulation in the highly sporulating SK1 strain. An independent analysis of these network properties in a less efficient sporulating strain helped to understand the heterogeneity of network profiles. We show that comparison of network properties has the potential to identify candidate nodes contributing to the phenotypic diversity of developmental processes in natural populations. Therefore, studying these network parameters as described in this work for dynamic developmental processes, such as sporulation in yeast and eventually in disease progression in humans, can help in identifying candidate factors which are potential regulators of differences between normal and perturbed processes and can be causal targets for intervention.

INTRODUCTION

Network theory has been used to understand the cellular organization and reprogramming of biological processes (Barabási and Oltvai 2004; Buganim et al. 2012; Gitter et al. 2012). Significant changes in network architecture determined from genome-wide expression studies have identified a few crucial genes such as transcription factors as hubs (Luscombe et al. 2004; Buganim et al. 2012). Transcript abundance of nodes (genes) in a biological network is under genetic control (Brem et al. 2002), and this has led to the understanding of how genetic variation is mechanistically associated with gene expression changes that underlie physiological differences (Ehrenreich et al. 2010; Cubillos et al. 2011). These studies have resulted in the development of exciting approaches aimed at predicting phenotypic consequences of genetic perturbations (Shen et al. 2010). The methods used in these studies perform comparative gene expression analysis such as clustering methods (Eisen et al. 1998), bootstrapping clustering (Kerr and Churchill 2001), four-stage Bayesian model (Wakefield et al. 2003), Gaussian mixture models with a modified Cholesky-decomposed covariance structure (McNicholas and Murphy 2010), etc. However, these gene-centric methods tend to overlook local patterns where these genes are similar based on only a subset (subspace) of attributes, for example, expression values. This led to an implementation of pattern similarity based bi-clustering approaches to gene expression data that could find bi-clusters among co-regulated genes under the different subset of experimental conditions (Roy et al. 2013).

Sporulation in budding yeast, Saccharomyces cerevisiae, is a linear developmental process initiated under extreme nutrient starvation, involving meiotic cell divisions leading to spore formation (Neiman 2005; 2011). Several genetic, biochemical, and genome-wide transcriptome analyses have elucidated the cascade of transcriptional regulatory processes during sporulation (Fig. 1A). These analyses identified over 1,000 genes that change their expression during meiosis (Chu et al. 1998; Primig et al. 2000). These expression studies categorized the differentially expressed genes into clusters and elucidated their roles in meiosis and sporulation. These studies led to the identification of critical regulatory nodes that are responsible for cells transitioning between different developmental stages during sporulation, viz. *IME1*, initiator of meiosis and *NDT80*, a regulator of meiotic divisions (Chu et al. 1998; Primig et al. 2000). Several sporulation studies have attempted to determine its underlying network structure (Wang et al. 2005; Shen et al. 2010; Ding and Wang 2011). By

applying mechanistic models to transcriptional data, for a few crucial genes of sporulation network such as *NDT80*, regulatory mechanisms of this developmental program has been studied (Wang et al. 2005). Most these genome-wide expression studies are done in high sporulation efficiency SK1 strain (90% in 48h, Keeney 2009), while most molecular studies on sporulation are performed on the low sporulating S288c strain (5-10% in 48h; Fig. 1B, Deutschbauer and Davis 2005).

The next step in interpreting gene expression profiles is to go beyond the gene-centric techniques and employ more global approaches for a more comprehensive understanding of how gene expression profiles are specifically related to the regulatory circuitry of the genome (Huang 1999). Network theory provides an efficient framework for capturing structural properties and dynamical behaviour of a range of systems spanning from society (Jalan et al. 2014) to biology (Rai et al. 2014; Shinde et al. 2015). Here we used the network theory approach to investigate how genome-wide transcriptional regulatory networks vary across time and how the determination of various network parameters can help in identifying crucial hubs in this dynamic process. By integrating time-resolved transcriptomics data with the known physical gene interaction network of yeast, we created discrete longitudinal networks of yeast sporulation at multiple time points for the SK1 strain. Using global and local network parameters during sporulation in SK1 strain, we understood its longitudinal networks and identified the nodes that get highly perturbed as this strain enters into the meiotic pathway. An independent longitudinal analysis of these network parameters in the S288c strain that sporulates less efficiently than SK1 was done to understand the heterogeneity of network profiles between these two strains. Identification of crucial nodes and genes helped us to understand how variation in the genetic background can lead to a variable phenotype, sporulation efficiency in this case.

RESULTS AND DISCUSSION

Global properties of longitudinal transcriptional regulatory networks of SK1 in sporulation medium

We constructed discrete longitudinal transcriptional regulatory networks of SK1 strain in sporulation medium. This network is composed of 12 sub-networks for each time point during sporulation. Nodes in each sub-network consisted of genes that were differentially expressed with respect to the first-time point (i.e. t_0 = 0h) by at least 2-fold. We first asked how do the

different sub-networks of SK1 strain interact with one-another and form functional modules during the sporulation process. For this, we calculated the clustering coefficient for all sub-networks that measures the local cohesiveness between the nodes (Watts and Strogatz 1998). A high value of clustering coefficient of a node depicts high connectivity among its neighbours. We evaluated the average value of clustering coefficient, $\langle C \rangle$, for each time point. As expected for various biological networks (Albert and Barabási 2002), a high value of $\langle C \rangle$ was observed for the networks at all time points in SK1 as compared to the corresponding random networks (Fig. 2, Supplementary Table S1) as expected (Watts and Strogatz 1998). To follow these changes, we independently investigated the early (1-4h, as established for this strain previously (Chu et al. 1998; Primig et al. 2000)), middle (4-8h), and late phases of sporulation (later than 8h) in SK1 (Fig. 1B). By comparing $\langle C \rangle$ the time points, a sharp increase in its value was observed twice for SK1 (Fig. 2B). This increase was at 4h and 9h, signifying modularity in the SK1 network at these time points. Since these early (4h) and middle (9h) time-points are when major meiotic decisions such as those regarding entry into Meiosis I (MI) and Meiosis II (MII), respectively (Chu et al. 1998; Primig et al. 2000) take place, a significant clustering coefficient $\langle C \rangle$ may play a role in decision making. Furthermore, keeping in view the manner in which we constructed these sporulation networks, a high $\langle C \rangle$ meant that many of the neighbour target genes of a transcription factor also acted as transcription factors for the other neighbour target genes of that same transcription factor.

A closer look at the longitudinal network architecture of sporulation of SK1 revealed that it showed a heterogeneous degree distribution. Degree distribution $\langle k \rangle$ is a network parameter defining the number of edges between various nodes. In this biological network, a heterogeneous degree distribution could mean that a few transcription factors dominate the entire network, as is observed in most real-world networks (Albert and Barabási 2002). The sub-networks exhibited a wide range of network sizes across different sporulation time points (Table 1). The larger and denser sub-networks were indicative of extensive regulatory changes in the SK1 strain. We found that there was a drastic increase in the number of genes having significantly high or low expression values in the consecutive time points at the onset of sporulation in the early phase (Table 1), which could be due to cells transitioning from

mitotic growth to initiate meiosis. The value of average degree distribution $\langle k \rangle$ showed a sudden increase in first two points but remained constant across the remaining time points (Fig. 2A). This extensive reprogramming of gene expression early in sporulation as the cells prepare to enter meiotic cell division (Gupta et al. 2016) was revealed as an abrupt increase in the involvement of genes with sporulation progression in the early phase. However, as the sporulation progressed to later phases, the rate of change in network size reduced. Interestingly, despite changes in the early sporulation phase, the ratio of the number of differentially expressed transcription factors (N_{TF}) and target genes remained almost constant across all the time points (Table 1). The proportion of regulatory genes remaining constant throughout sporulation indicated that it might be an intrinsic property of the sporulation process (Supplementary Fig. S1).

A change in the number of connections modulates the intrinsic properties of a network (Albert and Barabási 2002). We investigated the impact of this change on SK1 during the three sporulation phases. Similar to the network size, the number of connections (N_c) increased drastically in the early phase. However, the rate of increase in the number of connections was high as compared to the rate of increase in their size, which could be attributed to the appearance of more number of high degree nodes in the second-time point (Table 1). The nodes having high degree refer to genes that regulate a large number of other genes. These highly interacting genes are known to be important in various cellular processes (Rai et al. 2014). In the middle and late phase of sporulation, when processes involved in meiotic divisions occur (Chu et al. 1998), we observed that the number of connections did not show a considerable change since more than 75% of the genes remained same across different time points in this phase (see Supplementary Information).

We next analysed how the interacting patterns impacted the overall structure of the underlying networks. We examined the degree-degree mixing of the connected nodes across the three phases of sporulation in SK1. Disassortativity is a parameter that measures the correlation in the degrees of the nodes in a network and provides an understanding of the dislikelihood in connectivity of the underlying systems (Newman 2002). In gene regulatory networks, highly connected nodes avoid linking directly to each other and instead connect to genes with only a few interactions, thus exhibiting disassortative topology (Yook et al. 2005). This behaviour of the nodes leads to a reduction in crosstalk between different functional modules and increase in the robustness of the networks by localising the effects of deleterious perturbations (Maslov

and Sneppen 2002). Pearson (degree-degree) correlation coefficient (r) was calculated for the sub-networks at all time points. As expected for gene regulatory networks, sporulation sub-networks in SK1 exhibited disassortativity at all time points (Fig. 2C). A high negative value correlation (r = -0.45, Supplementary Table S1) was observed during the early phase of sporulation in SK1 suggesting that the strain was highly resilient to perturbations during early transcriptional events of sporulation (Maslov and Sneppen 2002). After the early phase, in SK1, disassortativity values reach a steady state at middle sporulation phase, suggesting an increase in inter-module crosstalk.

Local properties of longitudinal regulatory networks of SK1 in sporulation medium

Having analysed the global properties of the sporulation networks, we next studied the impact of the local architecture of SK1 network on the phenotypic profile of the strain. We were interested in investigating how the number of neighbours of nodes denoted by node degree was associated with their neighbour connectivities (interactions between the neighbours of the node of interest) evaluated regarding clustering coefficient. All the networks in SK1 exhibited negative degree-clustering coefficient correlation (Supplementary Fig. S2) as observed in various other real-world networks, indicating the existence of hierarchy in these underlying networks (Barabási and Oltvai 2004). A hierarchical architecture implies that sparsely connected nodes are part of highly clustered areas, with communication between the different highly clustered neighbourhoods maintained by a few hubs. We quantified this hierarchy (h), also termed as global reaching centrality in networks (Mones et al. 2012) and found that the networks were more hierarchical at the beginning of sporulation process (Fig. 2D). A high value of hierarchy is associated with modularity in the network. For instance, in case of metabolic networks, hierarchical structure indicates that the sets of genes sharing common neighbour are likely to belong to the same functional class (Ravasz et al. 2002). A low value of h indicates more random interactions in the underlying networks. SK1 showed the highest value of hierarchy at time point 1 (h = 0.5, around 1h, Supplementary Table S1). A decrease in the hierarchy was observed until the middle phase of sporulation which further diminished in the late phase in SK1, yet again emphasising on the importance of transcriptional regulatory activities in the early phase, on the sporulation process. This inference is in line with the earlier experimental investigations carried out to understand sporulation (Gupta et al. 2015), thus demonstrating the success of this network-based approach in unravelling the underlying genetic basis.

For a network, betweenness centrality is a measure of network resilience (Newman 2001), and it estimates the number of shortest paths (the minimum number of edges traversed between each of the pairs of nodes) that increases if a node is removed from the network (Newman 2006). Usually, nodes with a high degree have high betweenness centrality and are known to bridge different communities in the network. However, in a network, there exist some nodes, which despite having a low degree have relatively high betweenness centrality (Jalan et al. 2014). In the case of gene regulatory networks, such nodes (genes or transcription factors) are involved in less number of regulatory interactions, but these interactions are with different signalling pathways. Thus, these nodes are expected to have particular significance in the underlying networks as their removal can result in a breakdown in the regulatory pathways. Furthermore, in very few cases, a target gene, known to have a low degree may also have relatively higher betweenness centrality than the other target genes if several transcription factors are simultaneously regulating it. For our longitudinal sporulation networks, we identified a few important sporulation-associated genes showing this property (Fig. 3A). These were known regulators of respiratory stress and starvation, namely *STP2* (Merz and Westermann 2009), *PMA1* (Ding et al. 2009) and *RPL2B* (Davey et al. 2012), processes involved in regulation of early phases of sporulation. These results show the importance of this network attribute in identifying nodes with regulatory roles in a dynamic biological process.

The above analyses helped us to identify influential genes underlying initial transition from mitotic to meiotic cycle which is a significant change in transcriptional profile. We next identified a few interactions that might be instrumental in regulating the sporulation process by considering a significant proposition from sociology, Granovetter's weak ties hypothesis (Granovetter 1973). This hypothesis states that the degree of overlap of two individuals' friendship networks varies directly with the strength of their tie to one another. In the networks, the ties having low overlap in their neighbourhoods (i.e. less number of common neighbours) are termed as the weak ties (Onnela et al. 2007). The weak ties that have high link betweenness centrality are the ones known to bridge different communities (Szell and Thurner 2010). Such weak ties revealed through our analysis of different transcriptional regulatory networks are listed in Table 3. Interestingly, we found a repetitive occurrence of the same weak ties in consecutive time points for SK1 indicating their phase-specific importance in yeast sporulation. For instance, *DAL81-ACE2* and *CDC14-ACE2* were repetitive weak ties with high link betweenness centrality in consecutive sub-networks of

SK1. To assess the functional importance of these weak ties, we investigated the characteristic properties of the end nodes of these weak ties. Unlike social networks where the end nodes of weak ties are low degree nodes (Sarkar et al. 2016), in the sporulation networks of SK1, the nodes forming weak ties were high degree nodes. An example of this was *BAS1*, a Myb-related transcription factor involved in amino acid metabolism and meiosis (Mieczkowski et al. 2006). In addition to *BAS1*, other significant sporulation regulatory genes identified in SK1 were *RIM101*, a pH-responsive regulator of an initiator of meiosis (Su and Mitchell 1993); *IME2*, a serine-threonine kinase activator of *NDT80* and meiosis (Honigberg and Purnapatre 2003); *CDC14*, a protein phosphatase required for meiotic progression (McDonald et al. 2005); *HCM1*, an activator of genes involved in respiration (Rodriguez-Colman et al. 2010) (see Table 3). These results showed that genes showing weak ties and high betweenness centrality are meiosis-associated genes that form essential bridges in SK1. Moreover, *DAL81*, a nitrogen starvation regulator (Marzluf 1997) and *ACE2*, a regulator of G_1/S transition in the mitotic cell cycle (Spellman et al. 1998) were identified as end nodes of repetitive weak ties in SK1, suggesting their probable regulatory role in the sporulation process that requires further investigation (Table 3).

Longitudinal transcriptional regulatory networks of S288c in sporulation medium

SK1 is a strain that sporulates efficiently with 90-95% of all its cells entering sporulation within 8-10h in sporulation medium. However, a genetically divergent standard laboratory strain S288c is not an ideal strain to study sporulation as it has low sporulation efficiency (5-10% in 48h; Deutschbauer and Davis 2005). Furthermore, several studies have provided evidence that S288c is not the ideal strain for studying the allelic spectrum in nature, constituting artificial combinations of alleles that have never together been exposed to natural selective pressure (Liti et al. 2009). Whole genome resequencing of 39 yeast isolates was performed to study the genetic variation in genome sequences of diverse strains of yeast, and amongst these strains, S288c diverged strongly by being a phenotypic extreme, across multiple environmental conditions including stresses (Warringer et al. 2011). We were interested in understanding how S288c responds to a stressful environment like sporulation induction, and if by studying global and local network properties, we could identify sporulation-specific differences between S288c and SK1. For this, similar to SK1, we constructed longitudinal transcriptional regulatory networks of S288c strain in sporulation medium that comprised of 8 sub-networks at each time point during sporulation (Gupta et al.

2015). To determine the interactions between these sub-networks and identify the time frame in which S288c forms functional modules, we calculated the average clustering coefficient, $\langle C \rangle$, for all sub-networks of S288c. While a sharp increase in $\langle C \rangle$ was observed two times for SK1 coinciding with the early and middle time points in sporulation medium, in S288c, this value was not found to show these characteristic increases (Fig. 4B). A low $\langle C \rangle$ during middle sporulation in S288c might suggest a lack of interaction between crucial transcription factors and their target genes, leading to low sporulation in this strain. This inability of S288c to form functional modules indicates an inherent property of the strain in responding to stress environments. This inability could be due to a delay in relaying information between functional modules in a cell possibly affecting S288c stress tolerance.

When we studied other global network properties of S288c strain, we observed that it exhibited a drastic increase in the number of significantly expressed genes in its early hours in sporulation medium (Table 3) that was similar to SK1. However, the rate of increase in the number of connections (N_c) was much higher in the case of SK1 as compared to S288c. For instance, in the early time points in sporulation medium, S288c had a two-fold increase in the number of connections, whereas, SK1 exhibited a four-fold increase (Tables 1 and 2, Fig. 4A). Moreover, the average degree $\langle k \rangle$ for SK1 was higher in the first few time points compared to S288c with first two time points being significant in SK1 ($\langle k \rangle$ = 2.9 and 5.3) and only first time point in S288c ($\langle k \rangle$ = 3.0, see Figs. 1 and 3, Supplementary Tables S1 and S2).

Since the same repository base network forms the basis for all the interactions for both the strains, a change in the number of connections is possible only if old nodes (genes) disappear or new nodes arise in the networks. Incidentally, as observed in SK1, S288c networks exhibited negative degree-clustering coefficient correlation (Supplementary Fig. S3), indicating the existence of hierarchy in these underlying networks (Barabási and Oltvai 2004). As time progresses, a fall in the number of connections was observed in S288c compared to SK1, with S288c consistently showing lower values for all time points (see Tables 1 and 2). Incidentally, this decrease in the number of connections can be attributed to the disappearance of the high degree node *BAS1*, a Myb-related transcription factor involved in amino acid metabolism and meiosis (Mieczkowski et al. 2006), discussed above as the high-degree end

node of weak ties in the SK1 network. Interestingly, in the early time points of S288c, *BAS1* contributed to approximately 50% of the connections (Table 2) even though it is not one of the known regulators of sporulation (Neiman 2011). Its disappearance as time progressed in sporulation medium in S288c was reflected in the number of connections. Furthermore, surprisingly, its involvement in early regulatory processes and disappearance as time progressed, was observed in both the strains. On the one hand, this indicated the specific significance of this gene intrinsic to the early phase of sporulation; on the other hand, it reflected the drastic changes in the regulatory activities as time progresses in sporulation medium. Furthermore, in the later time-points (approximately the time when all the SK1 cells have entered sporulation), the number of connections almost doubled for S288c, and two known stress-responsive regulators, namely *MSN4* and HSF1 (Görner et al. 1998) with a large number of edges appeared in this phase (Table 2). *MSN4*, a known stress response regulator is one of the prime initiators of sporulation by regulating *NDT80*. While it is difficult to speculate whether it is a cause or a consequence – the absence of *MSN4* in the early S288c sub-networks could be one of the reasons behind its poor sporulation ability. This aspect is further exemplified by disassortativity values, which in SK1 was significantly high (close to -0.45) in the early phase of sporulation remaining resilient initially and then a steady increase in module crosstalk. However, in S288c, while the network remained resilient, but module crosstalk fluctuated until very late (Fig. 4C, Supplementary Table S2). This fluctuation implied that the necessary crosstalk between functional modules occurred early in SK1 but was still going on or was random and unstable until later time points for the S288c strain. Furthermore, we observed a significant increase in the hierarchy at around 8h in sporulation medium for S288c, which in comparison was high in early sporulation for SK1 (Fig. 4D). This late increase in hierarchy again suggested that there was a delay in increased modularity in S288c, which could adversely affect its sporulation process. Crucial regulatory molecular decisions are needed to be taken by a cell in a finite time-window, especially for phenotypes related to developmental processes and stress-related phenotypes. Since metabolic decisions are taking place very rapidly in the early time points of sporulation, if this early signal is relayed later in the S288c, then the finite time-frame is lost. Hence, delayed appearance of crucial regulatory events can be construed as a very critical factor affecting the sporulation process in S288c, and it is possible that the strain remains primarily in the quiescent state resulting in poor sporulation efficiency.

Next, we were interested in identifying influential genes in the S288c longitudinal

transcriptional regulatory network in sporulation medium. Interestingly, genes showing the property of low degree and high betweenness centrality were sporulation-associated genes such as *IME1* (Neiman 2011) and *TOS4* (Horak et al. 2002) involved in the initiation of meiosis and DNA replication checkpoint response, respectively. These sporulation-associated genes showed this network property in the early phase in SK1, but in S288c this was only observed at later time points (Supplementary Tables S3, S4). These results again suggested that this late appearance of crucial early sporulation genes as bridges that could transfer information between regulatory modules during early sporulation might be the cause for sporulation not proceeding in S288c. We then examined if identifying instrumental node interactions occurring within 8h in sporulation medium could provide us with candidate molecular pathways that plausibly cause the delay in cross-talk between functional modules of the S288c strain. We observed the repeated occurrence of *BAS1-RTT107*, *BAS1-TYE7*, *YAP6-BAS1* and *ASK10-HMO1* in consecutive time-points of S288c sub-networks. We observed apart from *BAS1* (discussed above), genes associated with mitotic functions as end nodes of these weak ties, such as *RTT107* for DNA repair (Leung et al. 2011), *TYE7* for glycolytic gene expression (Sato et al. 2000), *YAP6* for carbohydrate metabolism (Hanlon et al. 2011), *ASK10* for glycerol transport (Beese et al. 2009) and *HMO1* for DNA structure modification (Murugesapillai et al. 2014) (Fig. 3B). While in SK1 meiosis-associated genes formed essential bridges (Fig. 3A), in S288c, genes involved in mitotic functions formed these bridges (Fig. 3B). It implies how differences in weak ties in regulatory networks can help us understand the dramatic differences observed in phenotypes between these strains. It is important to note that since the genome-wide expression data used in these analyses is a bulk expression of a large number of cells, a variation in the number of cells undergoing sporulation, which is the case between SK1 and S288c, can affect the genes identified in our network analyses. Further, it is possible that due to cascading consequences of early dysregulation observed in S288c, many of the following sporulation processes in S288c are dysregulated as well, resulting in changes in network properties compared to SK1. It is not possible to disentangle these two, but our results show that network property analyses were able to identify known differences between SK1 and S288c. Going ahead, it would be interesting to further investigate these genes and their influence on the robustness of the longitudinal network architectures in these strains.

CONCLUSION

This study presents a novel framework for assessing the molecular underpinnings underlying a phenotype using time-resolved gene expression profile. This framework helped reveal the characteristic signatures of a phenotype and identified candidate genes contributing to and affecting the phenotype. Inferences drawn based on the comparative investigations of structural attributes of the longitudinal sporulation networks of the two strains revealed that late appearance of early regulators and delayed crosstalk between functional modules might be critical for progression of sporulation process in SK1 and the plausible reasons behind the low sporulation efficiency of S288c. This speculation is especially interesting since most causative genetic variants known to contribute to sporulation efficiency variation have been observed in genes either showing early role in sporulation or affecting genes with an early regulatory role in sporulation (Deutschbauer and Davis 2005; Ben-Ari et al. 2006; Gupta et al. 2014; 2016).

Application of genome-wide strategies to elucidate the molecular networks in multiple genetic backgrounds provides us with the opportunity to understand the impact of natural variation. Studying these network properties for variation in causal genes would further help in understanding specific molecular effects in the different temporal phases of the phenotype. The strategies adopted in this work can be extended to assess the impact of molecular perturbations in the already known core interaction network of an organism (Carter et al. 2013; Gasch et al. 2016). Moreover, application of such network analyses on gene expression datasets for disease progression in complex diseases such as cancer and metabolic disorders can help identify specific nodes perturbing the underlying molecular pathways that can be the focus of personalised medicine and drug target discovery.

METHODS

Network construction

For constructing the transcriptional regulatory sporulation network, the known static regulatory interactions were overlaid on the time-resolved transcriptomics data of the two strains. The overlaying of the regulatory network and temporal transcriptome data created the longitudinally integrated sporulation networks. The static network known for yeast contains all the known regulatory interactions between all the yeast transcription factors (TF) and their target genes (TG). These interactions were obtained from YEASTRACT database (Teixeira et

al. 2013), a curated repository of regulatory associations in S. cerevisiae, based on more than 1,200 literature references.

Gene expression data were obtained from previously published studies – for SK1 (Lardenois et al. 2011), and S288c (Gupta et al. 2015). These datasets contained gene expression of 6,926 genes across 13 different time points in linear scale (0h to 12h with 1h intervals termed as T0 to T12, respectively) in SK1 and 9 different time points in logarithmic scale (0h, 30m, 45m, 1h10m, 1h40m, 2h30m, 3h50m, 5h40m, 8h30m termed as T0 to T8, respectively) in S288c. Gene expression analysis was performed as described previously (Gupta et al. 2015). In brief, all time points were normalized together using vsn (Huber et al. 2002) and the log2 transformed expression values obtained after normalisation were smoothed using locfit. Fold differences in expression values were calculated for all the time-points relative to t = 0h (t0), as follows:

$$Y'_{SK1(t_n)} = Y_{SK1(t_n)} - Y_{SK1(t_0)} \quad (1)$$

$$Y'_{S288c(t_n)} = Y_{S288c(t_n)} - Y_{S288c(t_0)} \quad (2)$$

such that Y is the expression value of a transcript for a strain (SK1 or S288c) at a specific time point n and Y' is the transformed expression value. One of the reasons behind considering fold-change values is to mask the effect of any possible transcriptional noise arising in the system.

At each time point, differentially expressed genes were identified by setting the threshold value on log_2 fold differences as 1.0. Hence, genes that were considered over-expressed or repressed showed at least a 2-fold difference with respect to the first time point t_0 (i.e. t = 0h) (Gupta et al. 2015). This step again reduces the possibility of picking up any gene that contributes to transcriptional noise.

The longitudinal sporulation network was constructed by overlaying the experimentally determined yeast sporulation-specific gene expression values on the yeast static network. For each time point of each strain, only those TF-TG pairs were considered that both showed either overexpression or repression. These pairs were included in the subnetwork for that specific time point and thus, subnetworks for each time point were constructed for each strain. For comparison of the gene names obtained from YEASTRACT and the sporulation gene

expression data, aliases were obtained from Saccharomyces Genome Database (Cherry et al. 2012). This step again refines the interaction data.

Data availability

The codes, supplementary tables and figures, adjacency matrices of the networks constructed using longitudinal sporulation data drawn from SK1 and S288c strains, the corresponding gene indices and transcription factors are freely available online at figshare https://doi.org/10.6084/m9.figshare.3457508.v4.

Structural parameters

Several statistical measures are proposed to understand specific features of the network (Albert and Barabási 2002; Boccaletti et al. 2006). The number of connections possessed by a node is termed as its degree. The spread in the degrees is characterised by a distribution function $P(k)$, which gives the probability that a randomly selected node has precisely k edges. The degree distribution of a random graph is a Poisson distribution with a peak at $P(\langle k \rangle)$. However, in most large networks such as the World Wide Web, the Internet or the metabolic networks, the degree distribution significantly deviates from a Poisson distribution but has a power-law tail $P(k) \sim k^{-\gamma}$. The inherent tendency of social networks to form clusters representing circles of friends or acquaintances in which every member knows every other member is quantified by the clustering coefficient. We categorise the nodes as high and low degree nodes by arranging all the nodes in a network in descending order of degrees and keep assigning the nodes as high degree nodes until the next lower degree node differs by nearly 1.5-fold from the former in terms of the degree. The clustering coefficient of a node i denoted as C_i, is defined as the ratio of the number of links existing between the neighbours of the node to the possible number of links that could exist between the neighbours of that node (Newman et al. 2001) and is given by

$$C_i = \frac{2\sum_{j_2=1}^{k_i}\sum_{j_1=1}^{k_i}\left(A_{ij1}A_{j1j2}A_{j2i}\right)}{k_i\left(k_i-1\right)} \tag{3}$$

where i is the node of interest and j_1, and j_2 are any two neighbours of the node i and k_i is the degree of the node i. The average clustering coefficient of a network corresponding to a

particular condition $\langle C \rangle$ can be written as

$$\langle C \rangle = \frac{1}{N} \sum_{i=1}^{N} C_i \tag{4}$$

We define the betweenness centrality of node *i*, as the fraction of shortest paths between node pairs that pass through the said node of interest (Newman 2006).

$$x_i = \sum_{st} \frac{n_{st}^i}{g_{st}} \tag{5}$$

where n_i is the number of geodesic paths from *s* to *t* that passes through *i* and g_{st} is the total number of geodesic paths from *s* to *t*. All the nodes were plotted, and the top 5% of the nodes (genes) with high betweenness centrality but low degree were identified.

We quantify the degree-degree correlations of a network by considering the Pearson (degree-degree) correlation coefficient, given by Newman (2002)

$$r = \frac{\left[M^{-1} \sum_{i=1}^{M} j_i k_i \right] - \left[M^{-1} \sum_{i=1}^{M} \frac{1}{2} \left(j_i + k_i \right)^2 \right]}{\left[M^{-1} \sum_{i=1}^{M} \frac{1}{2} \left(j_i^2 + k_i^2 \right) \right] - \left[M^{-1} \sum_{i=1}^{M} \frac{1}{2} \left(j_i + k_i \right)^2 \right]} \tag{6}$$

where j_i, k_i are the degrees of nodes at both the ends of the ith connection and *M* represents the total connections in the network.

Link betweenness centrality is defined for an undirected link as

$$\beta_L = \sum_{v \in V_S} \sum_{w \in V/v} \sigma_{vw}(e) / \sigma_{vw} \tag{7}$$

where $\sigma_{vw}(e)$ is the number of shortest paths between *v* and *w* that contain *e*, and σ_{vw} is the total number of shortest paths between *v* and *w* (Onnela et al. 2007).

The overlap of the neighbourhood of two connected nodes *i* and *j* is defined as (Onnela et al. 2007)

$$O_{ij} = \frac{n_{ij}}{\left(k_i - 1\right) + \left(k_j - 1\right) - n_{ij}} \tag{8}$$

where n_{ij} is the number of neighbours common to both nodes i and j. Here k_i and k_j represent the degree of the ith and j^{th} nodes.

The hierarchy can be defined as the heterogeneous distribution of local reaching centrality of nodes in the network. The local reaching centrality, (C_R), of a node i is defined as (Mones et al. 2012)

$$C_R\left(i\right) = \frac{1}{N-1} \sum_{j:0<d\left(i,j\right)<\infty} \frac{1}{d\left(i,j\right)} \tag{9}$$

where $d(i,j)$ is the length of the shortest path between any pair of nodes i and j. The measure of hierarchy (h), termed as global reaching centrality is given by

$$h = \frac{\sum_{i \in V} \left[C_R^{\max} - C_R\left(i\right)\right]}{N-1} \tag{10}$$

ACKNOWLEDGEMENTS

SJ acknowledges Department of Science and Technology (DST), Govt. of India grant EMR/2014/000368 and Council of Scientific and Industrial Research (CSIR), Govt. of India grant 25(0205)/12/EMR-II. SG acknowledges support by Tata Institute of Fundamental Research during the initial phase. HS acknowledges support by Tata Institute of Fundamental Research during the initial phase and Indian Institute of Technology Madras.

AUTHOR CONTRIBUTIONS STATEMENT

SJ conceived the idea. SJ and HS designed and supervised the project. CS, RKV constructed the networks with the help of SG and analysed the structural properties. SG, RKV, CS, HS analysed the functional properties. CS, SG, HS wrote the initial draft, and all the authors contributed to the manuscript.

SUPPLEMENTARY INFORMATION

Supplementary information accompanies this paper.

COMPETING FINANCIAL INTERESTS STATEMENT

The authors declare no competing financial interests.

REFERENCES

Albert, R., and A.-L. Barabási, 2002 Statistical mechanics of complex networks. Rev Mod Phys 74: 47–97.

Barabási, A.-L., and Z. N. Oltvai, 2004 Network biology: understanding the cell's functional organization. Nat Rev Genet 5: 101–113.

Beese, S. E., T. Negishi, and D. E. Levin, 2009 Identification of positive regulators of the yeast Fps1 glycerol Channel. PLoS Genet 5: e1000738.

Ben-Ari, G., D. Zenvirth, A. Sherman, L. David, M. Klutstein et al., 2006 Four linked genes participate in controlling sporulation efficiency in budding yeast. PLoS Genet 2: e195.

Boccaletti, S., V. Latora, Y. Moreno, M. Chavez, and D. Hwang, 2006 Complex networks: Structure and dynamics. Phys Rep 424: 175–308.

Brem, R. B., G. Yvert, R. Clinton, and L. Kruglyak, 2002 Genetic dissection of transcriptional regulation in budding yeast. Science 296: 752–755.

Buganim, Y., D. A. Faddah, A. W. Cheng, E. Itskovich, S. Markoulaki et al., 2012 Single-cell expression analyses during cellular reprogramming reveal an early stochastic and a late hierarchic phase. Cell 150: 1209–1222.

Carter, H., M. Hofree, and T. Ideker, 2013 Genotype to phenotype via network analysis. Curr Opin Genet Dev 23: 611–621.

Cherry, J. M., E. L. Hong, C. Amundsen, R. Balakrishnan, G. Binkley et al., 2012 Saccharomyces Genome Database: the genomics resource of budding yeast. Nucl Acids Res 40: D700–5.

Chu, S., J. DeRisi, M. Eisen, J. Mulholland, D. Botstein et al., 1998 The transcriptional program of sporulation in budding yeast. Science 282: 699–705.

Cubillos, F. A., E. Billi, E. Zörgö, L. Parts, P. Fargier et al., 2011 Assessing the complex architecture of polygenic traits in diverged yeast populations. Mol Ecol 20: 1401–1413.

Davey, H. M., E. J. M. Cross, C. L. Davey, K. Gkargkas, D. Delneri et al., 2012 Genome-wide analysis of longevity in nutrient-deprived *Saccharomyces cerevisiae* reveals importance of recycling in maintaining cell viability. Environ Microbiol 14: 1249–1260.

Deutschbauer, A. M., and R. W. Davis, 2005 Quantitative trait loci mapped to single-nucleotide resolution in yeast. Nat Genet 37: 1333–1340.

Ding, J., X. Huang, L. Zhang, N. Zhao, D. Yang et al., 2009 Tolerance and stress response to ethanol in the yeast *Saccharomyces cerevisiae*. Appl Microbiol Biotechnol 85: 253–263.

Ding S., W. Wang, 2011 Recipes and mechanisms of cellular reprogramming: a case study on budding yeast *Saccharomyces cerevisiae*. BMC Syst Biol 5: 50.

Ehrenreich, I. M., N. Torabi, Y. Jia, J. Kent, S. Martis et al., 2010 Dissection of genetically complex traits with extremely large pools of yeast segregants. Nature 464: 1039–1042.

Eisen, M. B., P. T. Spellman, P. O. Brown, and D. Botstein, 1998 Cluster analysis and display of genome-wide expression patterns. Proc Natl Acad Sci USA 95: 14863–14868.

Gasch, A. P., B. A. Payseur, and J. E. Pool, 2016 The power of natural variation for model organism biology. Trends Genet 32: 147–154.

Gitter, A., M. Carmi, N. Barkai, and Z. Bar-Joseph, 2012 Linking the signaling cascades and dynamic regulatory networks controlling stress responses. Genome Res 23: 365–376.

Görner, W., E. Durchschlag, M. T. Martinez-Pastor, F. Estruch, G. Ammerer et al., 1998 Nuclear localization of the C2H2 zinc finger protein Msn2p is regulated by stress and protein kinase A activity. Genes Dev 12: 586–597.
Granovetter, M. S., 1973 The strength of weak ties. Am J Sociol 78: 1360–1380.
Gupta, I., S. Clauder-Münster, B. Klaus, A. I. Järvelin, R. S. Aiyar et al., 2014 Alternative polyadenylation diversifies post-transcriptional regulation by selective RNA-protein interactions. Mol Syst Biol 10: 719–719.
Gupta, S., A. Radhakrishnan, R. Nitin, P. Raharja-Liu, G. Lin et al., 2016 Meiotic interactors of a mitotic gene *TAO3* revealed by functional analysis of its rare variant. G3 (Bethesda) 6: 2255–2263.
Gupta, S., A. Radhakrishnan, P. Raharja-Liu, G. Lin, L. M. Steinmetz et al., 2015 Temporal expression profiling identifies pathways mediating effect of causal variant on phenotype. PLoS Genet 11: e1005195.
Hanlon, S. E., J. M. Rizzo, D. C. Tatomer, J. D. Lieb, and M. J. Buck, 2011 The stress response factors Yap6, Cin5, Phd1, and Skn7 direct targeting of the conserved co-repressor Tup1-Ssn6 in *S. cerevisiae*. PLoS ONE 6: e19060.
Honigberg, S. M., and K. Purnapatre, 2003 Signal pathway integration in the switch from the mitotic cell cycle to meiosis in yeast. J Cell Sci 116: 2137–2147.
Horak, C. E., N. M. Luscombe, J. Qian, P. Bertone, S. Piccirrillo et al., 2002 Complex transcriptional circuitry at the G1/S transition in *Saccharomyces cerevisiae*. Genes Dev 16: 3017–3033.
Huang, S., 1999 Gene expression profiling, genetic networks, and cellular states: an integrating concept for tumorigenesis and drug discovery. J Mol Med 77: 469–480.
Huber, W., A. von Heydebreck, H. Sültmann, A. Poustka, and M. Vingron, 2002 Variance stabilization applied to microarray data calibration and to the quantification of differential expression. Bioinformatics 18 Suppl 1: S96–104.
Jalan, S., C. Sarkar, A. Madhusudanan, and S. K. Dwivedi, 2014 Uncovering randomness and success in society. PLoS ONE 9: e88249.
Keeney, S., 2009 *Meiosis* Methods in Molecular Biology, Vol. 557, pages 365, Humana Press.
Kerr, M. K., and G. A. Churchill, 2001 Bootstrapping cluster analysis: assessing the reliability of conclusions from microarray experiments. Proc Natl Acad Sci USA 98: 8961–8965.
Lardenois, A., Y. Liu, T. Walther, F. Chalmel, B. Evrard et al., 2011 Execution of the meiotic noncoding RNA expression program and the onset of gametogenesis in yeast require the conserved exosome subunit Rrp6. Proc Natl Acad Sci USA 108: 1058–1063.
Leung, G. P., L. Lee, T. I. Schmidt, K. Shirahige, and M. S. Kobor, 2011 Rtt107 is required for recruitment of the SMC5/6 complex to DNA double strand breaks. J Biol Chem 286: 26250–26257.
Liti, G., D. M. Carter, A. M. Moses, J. Warringer, L. Parts et al., 2009 Population genomics of domestic and wild yeasts. Nature 458: 337–341.
Luscombe, N. M., M. M. Babu, H. Yu, M. Snyder, S. A. Teichmann et al., 2004 Genomic analysis of regulatory network dynamics reveals large topological changes. Nature 431: 308–312.
Marzluf, G. A., 1997 Genetic regulation of nitrogen metabolism in the fungi. Microbiol Mol Biol Rev 61: 17–32.
Maslov, S., and K. Sneppen, 2002 Specificity and stability in topology of protein networks. Science 296: 910–913.

McDonald, C. M., K. F. Cooper, and E. Winter, 2005 The Ama1-directed anaphase-promoting complex regulates the Smk1 mitogen-activated protein kinase during meiosis in yeast. Genetics 171: 901–911.
McNicholas, P. D., and T. B. Murphy, 2010 Model-based clustering of longitudinal data. Can J Stat 38: 153–168.
Merz, S., and B. Westermann, 2009 Genome-wide deletion mutant analysis reveals genes required for respiratory growth, mitochondrial genome maintenance and mitochondrial protein synthesis in *Saccharomyces cerevisiae*. Genome Biol 10: R95.
Mieczkowski, P. A., M. Dominska, M. J. Buck, J. L. Gerton, J. D. Lieb et al., 2006 Global analysis of the relationship between the binding of the Bas1p transcription factor and meiosis-specific double-strand DNA breaks in *Saccharomyces cerevisiae*. Mol Cell Biol 26: 1014–1027.
Mones, E., L. Vicsek, and T. Vicsek, 2012 Hierarchy measure for complex networks. PLoS ONE 7: e33799.
Murugesapillai, D., M. J. McCauley, R. Huo, M. H. Nelson Holte, A. Stepanyants et al., 2014 DNA bridging and looping by *HMO1* provides a mechanism for stabilizing nucleosome-free chromatin. Nucl Acids Res 42: 8996–9004.
Neiman, A., 2005 Ascospore formation in the yeast *Saccharomyces cerevisiae*. Microbiol Mol Biol Rev 69: 565–584.
Neiman, A. M., 2011 Sporulation in the budding yeast *Saccharomyces cerevisiae*. Genetics 189: 737–765.
Newman, M. E. J., 2002 Assortative mixing in networks. Phys Rev Lett 89: 208701.
Newman, M. E. J., 2001 Scientific collaboration networks. II. Shortest paths, weighted networks, and centrality. Phys Rev E 64: 016131.
Newman, M. E. J., 2006 The structure and function of complex networks. SIAM Rev 45: 167–256.
Newman, M. E. J., S. H. Strogatz, and D. J. Watts, 2001 Random graphs with arbitrary degree distributions and their applications. Phys Rev E 64: 026118.
Onnela, J.-P., J. Saramäki, J. Hyvönen, G. Szabó, M. A. de Menezes et al., 2007 Analysis of a large-scale weighted network of one-to-one human communication. New J Phys 9: 179–179.
Primig, M., R. M. Williams, E. A. Winzeler, G. G. Tevzadze, A. R. Conway et al., 2000 The core meiotic transcriptome in budding yeasts. Nat Genet 26: 415–423.
Rai, A., A. V. Menon, and S. Jalan, 2014 Randomness and preserved patterns in cancer network. Sci Rep 4: srep06368.
Ravasz, E., A. L. Somera, D. A. Mongru, Z. N. Oltvai, and A. L. Barabasi, 2002 Hierarchical Organization of Modularity in Metabolic Networks. Science 297: 1551–1555.
Rodriguez-Colman, M. J., G. Reverter-Branchat, M. A. Sorolla, J. Tamarit, J. Ros et al., 2010 The forkhead transcription factor Hcm1 promotes mitochondrial biogenesis and stress resistance in yeast. J Biol Chem 285: 37092–37101.
Roy, S., D. K. Bhattacharyya, and J. K. Kalita, 2013 CoBi: Pattern Based Co-Regulated Biclustering of Gene Expression Data. Pattern Recognit Lett 34: 1669–1678.
Sarkar, C., A. Yadav, and S. Jalan, 2016 Multilayer network decoding versatility and trust. Europhys Lett 113: 18007.
Sato, T., M. C. Lopez, S. Sugioka, Y. Jigami, H. V. Baker et al., 2000 The E-box DNA binding protein Sgc1p suppresses the *gcr2* mutation, which is involved in transcriptional activation of glycolytic genes in *Saccharomyces cerevisiae*. FEBS Letters 463: 307–311.
Shen, L., I. Chepelev, J. Liu, and W. Wang, 2010 Prediction of quantitative phenotypes based on genetic networks: a case study in yeast sporulation. BMC Syst Biol 4: 128.

Shinde, P., A. Yadav, A. Rai, and S. Jalan, 2015 Dissortativity and duplications in oral cancer. Eur Phys J B 88: 197.
Spellman, P. T., G. Sherlock, M. Q. Zhang, V. R. Iyer, K. Anders et al., 1998 Comprehensive identification of cell cycle-regulated genes of the yeast *Saccharomyces cerevisiae* by microarray hybridization. Mol Biol Cell 9: 3273–3297.
Su, S. S., and A. P. Mitchell, 1993 Identification of functionally related genes that stimulate early meiotic gene expression in yeast. Genetics 133: 67–77.
Szell, M., and S. Thurner, 2010 Measuring social dynamics in a massive multiplayer online game. Soc Networks 32: 313–329.
Teixeira, M. C., P. T. Monteiro, J. F. Guerreiro, J. P. Goncalves, N. P. Mira et al., 2013 The YEASTRACT database: an upgraded information system for the analysis of gene and genomic transcription regulation in *Saccharomyces cerevisiae*. Nucl Acids Res 42: D161–D166.
Wakefield, J. C., C. Zhou, and S. G. Self, 2003 *Modelling Gene Expression Data over Time: Curve Clustering with Informative Prior Distributions* (J. M. Bernardo, A. P. Dawid, J. O. Berger, M. West, D. Heckerman, et al., Eds.). Oxford University Press.
Wang, W., J. M. Cherry, Y. Nochomovitz, E. Jolly, D. Botstein et al., 2005 Inference of combinatorial regulation in yeast transcriptional networks: a case study of sporulation. Proc Natl Acad Sci USA 102: 1998–2003.
Warringer, J., E. Zörgö, F. A. Cubillos, A. Zia, A. Gjuvsland et al., 2011 Trait variation in yeast is defined by population history. PLoS Genet 7: e1002111.
Watts, D. J., and S. H. Strogatz, 1998 Collective dynamics of 'small-world' networks. Nature 393: 440–442.
Yook, S.-H., F. Radicchi, and H. Meyer-Ortmanns, 2005 Self-similar scale-free networks and disassortativity. Phys Rev E 72: 045105.

FIGURE LEGENDS

Fig. 1. (A) Schematic diagram of the sporulation process. Upon nutrient starvation, the yeast cell exits mitosis (cell division) and initiates meiosis and sporulation. This developmental process of meiosis and sporulation is divided into three phases: early, middle, and late, with each stage having distinct functions and crucial genes activated, adapted from Chu et al. (1998) (B) Heatmap of gene expression profiles of crucial meiosis and sporulation regulators in early, middle and late phases of sporulation. Early, middle, and late phases with their corresponding time points, demarcated by dashed lines, are defined based on the SK1 profile. For SK1, early, middle, and late regulators are indicated by orange, green and grey bars, respectively.

Fig. 2. Structural properties of the sporulation networks of SK1 across different time points. (A) Average degree $\langle k \rangle$; (B) average clustering coefficient $\langle C \rangle$; (C) Pearson degree-degree correlation coefficient or dis-assortativity (r); (D) global reaching centrality or hierarchy (h). The solid red line represents the average of 1000 randomised values of each structural

property with error bars representing standard deviation.

Fig. 3. Plots of average degree $\langle k \rangle$ as a function of betweenness centrality (β_C) in (A) SK1 (red) and (B) S288c (blue) networks. Each dot represents a gene and genes with the low degree but higher betweenness centrality in their respective time points are marked as black and named. Supplementary Tables S3 and S4 gives the complete list of genes.

Fig. 4. Structural properties of the sporulation networks of S288c across different time points. (A) Average degree $\langle k \rangle$; (B) average clustering coefficient $\langle C \rangle$; (C) Pearson degree-degree correlation coefficient or dis-assortativity (*r*); (D) global reaching centrality or hierarchy (*h*). The solid red line represents the average of 1000 randomised values of each structural property with error bars representing standard deviation. Supplementary Table S4 gives the complete list of genes.

TABLES

Table 1. Structural attributes of SK1 networks

N, N_c and N_{TF} respectively denote the network size, number of connections and number of transcription factors at a particular time point. Catalogue of high degree nodes (degree mentioned in brackets) and degree of *NDT80* gene are given for each time point of SK1 networks.

Time points	N	N_c	N_{TF}	**High degree nodes**	**Degree of *NDT80***
T_1	360	529	13	*BAS1* (283), *RIM101* (39)	-
T_2	791	2107	34	*BAS1* (480), *MSN4* (300), *KAR4* (132), *FKH1* (109), *AFT2* (75)	-
T_3	1096	4033	46	*ASH1* (612), *BAS1* (560), *MSN4* (363), *HMS1* (168)	47
T_4	1495	6540	63	*ACE2* (1250), *ASH1* (770), *MSN4* (428), *AFT*1 (369), *HMS1* (209)	69
T_5	1640	7720	73	*ASH1* (926), *MSN4* (499), *AFT*1 (437), *GCR1* (399), *HMO1* (248)	87
T_6	1900	8978	77	*ASH1* (1074), *MSN4* (580), *AFT*1 (486), *ISW2* (292)	101
T_7	1928	8012	75	*ASH1* (1118), *AFT*1 (504), *ISW2* (309)	102
T_8	1886	8281	76	*ASH1* (1094), *AFT*1 (484), *STE12* (432), *FHL1* (392), *ISW2* (299), *HMO1* (289)	95
T_9	2024	9213	74	*ACE2* (1700), *ASH1* (1039), *AFT*1 (461), *STE12* (421), *FHL1* (362), *ISW2* (266)	87
T_{10}	1880	8410	69	*ACE2* (1579), *ASH1* (955), *AFT*1 (435), *STE12* (402), *FHL1* (322), *ISW2* (243)	79

T_{11}	1711	7188	66	*ACE2* (1429), *ASH1* (870), *AFT*1 (428), *STE12* (386), *FHL1* (291), *HMS1* (222)	69
T_{12}	1745	7206	61	*ACE2* (1472), *ASH1* (886), *AFT*1 (481), *GAL4* (299), *HMS1* (246), *SIN4* (216), *INO4* (177)	77

Table 2. Structural attributes of S288c networks

N, N_c and N_{TF} respectively denote the network size, number of connections and number of transcription factors at a particular time point. Catalogue of high degree nodes (degree mentioned in brackets) and degree of *IME1* gene are given for each time point of S288c networks.

Time points	N	N_c	N_{TF}	**High degree nodes**	**Degree of *IME1***
T_1	434	605	15	*BAS1* (356), *HAP4* (92), *TUP1* (48)	10
T_2	706	1327	24	*BAS1* (523), *STE12* (158), *HAP4* (134), *CUP2* (117), *TUP1* (72)	14
T_3	691	1280	26	*BAS1* (524), *HAP4* (136), *CUP2* (114), *TUP1* (76), *PUT*3 (70)	13
T_4	569	1025	25	*BAS1* (442), *HAP4* (108), *RLM1* (93)	13
T_5	582	1049	24	*BAS1* (455), *HAP4* (112), *SWI5* (92)	15
T_6	415	733	23	*HMO1* (138), *HAP4* (125), *SWI5* (98)	16
T_7	409	686	18	*HMO1* (131), *HAP4* (124), *SWI5* (97)	19
T_8	513	1195	23	*MSN4* (261), *HSF1* (246), *HMO1* (129), *SWI5* (95), *INO4* (75)	21

Table 3. Pairs of interacting genes that have low overlap (O) and high link betweenness centrality (β_L) in SK1 networks

The interacting gene pairs in bold are mentioned in the main text. Their corresponding indices as given in Supplementary Fig. S4.

Time points	**Node pairs with low *O* high β_L**
T_1	*STP2-BAS1*, *HCM1-BAS1*, *BAS1-RIM101*
T_9	*ACE2-MGA1*, *ACE2-SIR2*, ***DAL81-ACE2***, ***CDC14-ACE2***
T_{10}	***CDC14-ACE2***, *ACE2-ARG82*, ***DAL81-ACE2***
T_{11}	***DAL81-ACE2***, *ACE2-MGA1*, ***CDC14-ACE2***
T_{12}	*YPL056C-ARG82*, ***DAL80-ACE2***, *RIF1-TPK1*

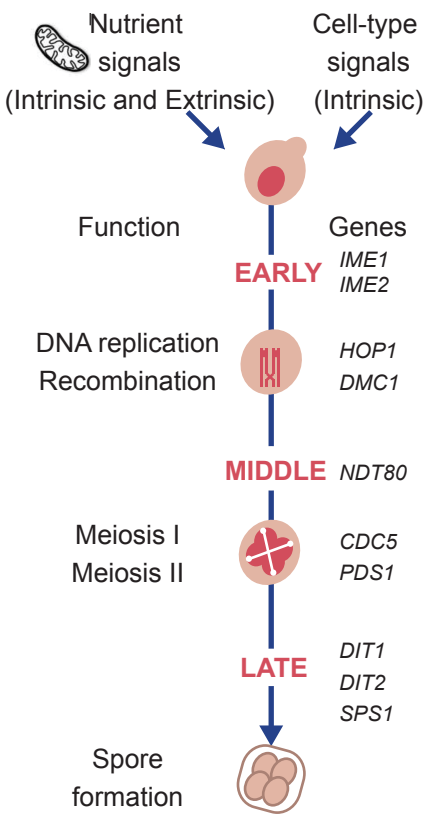
Table 4. Pairs of interacting genes that have low overlap (O) and high link betweenness centrality (β_L) in S288c networks

The interacting gene pairs in bold are mentioned in the main text. Their corresponding indices as given in Supplementary Figure S5.

Time points	**Node pairs with low *O* high β_L**
T_1	***BAS1-RTT107***, ***BAS1-TYE7***
T_2	***BAS1-RTT107***, ***BAS1-TYE7***
T_3	***BAS1-RTT107***, *BAS1-RPH1*, ***YAP6-BAS1***
T_4	***YAP6-BAS1***
T_6	***ASK10-HMO1***, *HMO1-XBP1*
T_7	***ASK10-HMO1***, *YAP6-HMO1*
T_8	*HSF1-IME1*, ***ASK10-HMO1***

Figure 1

A


Nutrient signals (Intrinsic and Extrinsic)
Cell-type signals (Intrinsic)
Function
Genes
EARLY
IME1
IME2
DNA replication
Recombination
HOP1
DMC1
MIDDLE
NDT80
Meiosis I
Meiosis II
CDC5
PDS1
LATE
DIT1
DIT2
SPS1
Spore formation

B

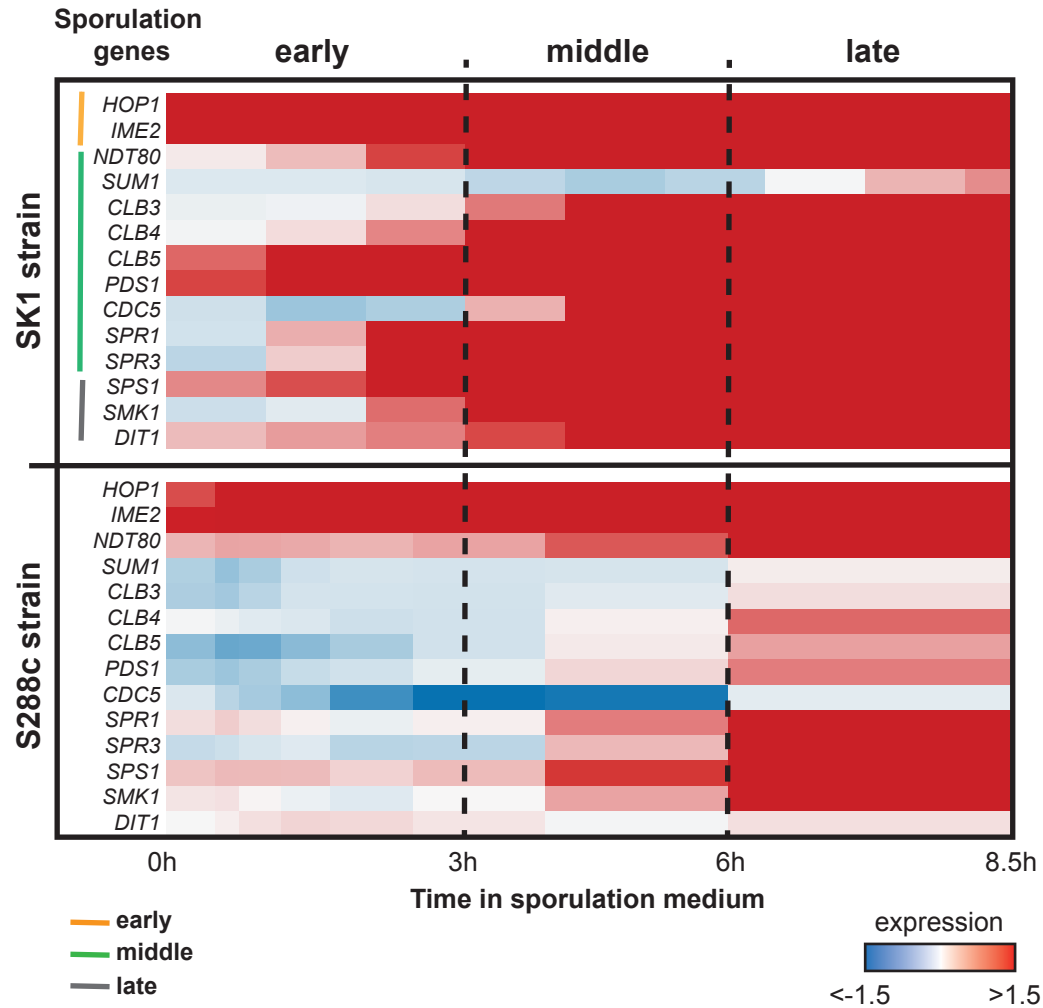
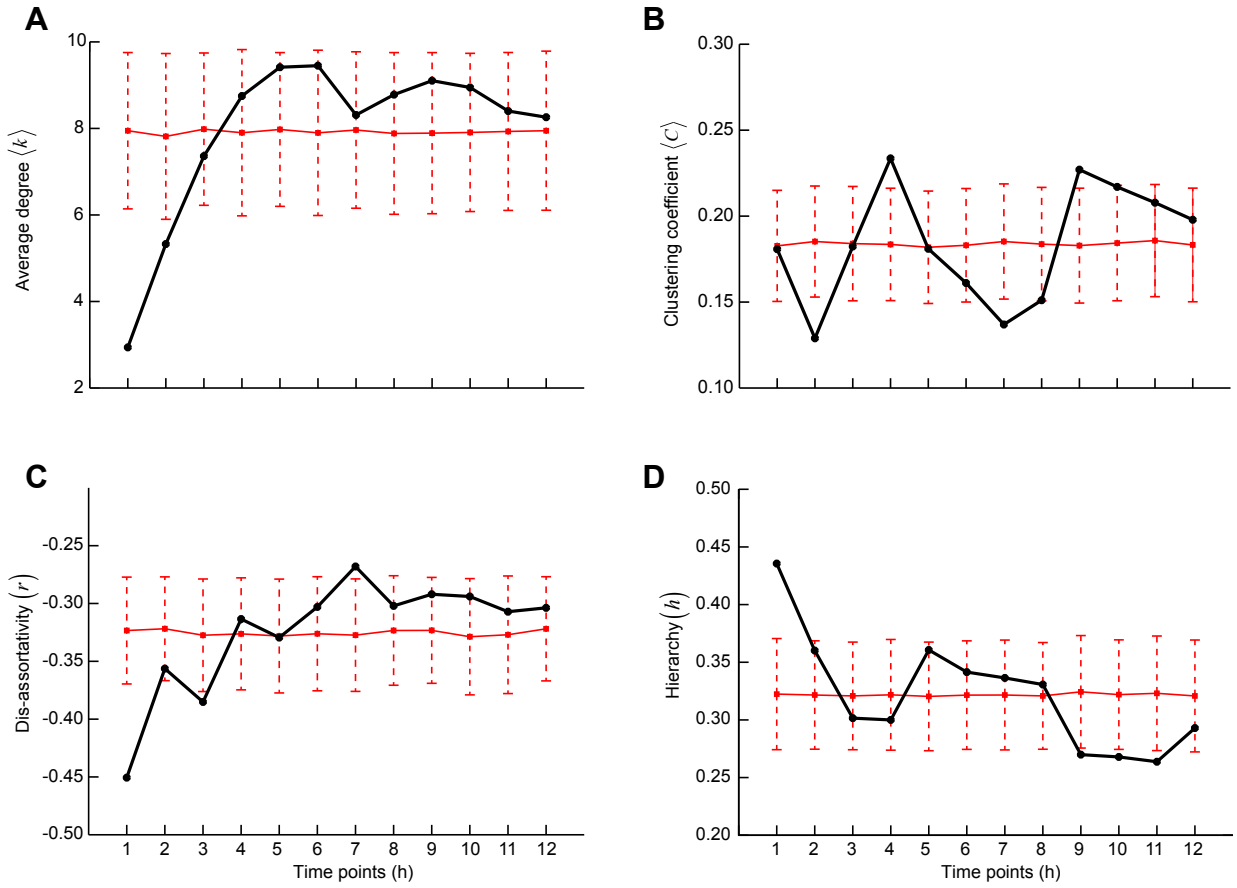

Sporulation genes
early
middle
late
SK1 strain
S288c strain
HOP1
IME2
NDT80
SUM1
CLB3
CLB4
CLB5
PDS1
CDC5
SPR1
SPR3
SPS1
SMK1
DIT1
0h
3h
6h
8.5h
Time in sporulation medium
early
middle
late
expression
<-1.5
>1.5

Figure 2


A
Average degree ⟨k⟩
10
8
6
4
2
B
Clustering coefficient ⟨c⟩
0.30
0.25
0.20
0.15
0.10
C
Dis-assortativity (r)
-0.25
-0.30
-0.35
-0.40
-0.45
-0.50
1 2 3 4 5 6 7 8 9 10 11 12
Time points (h)
D
Hierarchy (h)
0.50
0.45
0.40
0.35
0.30
0.25
0.20
1 2 3 4 5 6 7 8 9 10 11 12
Time points (h)

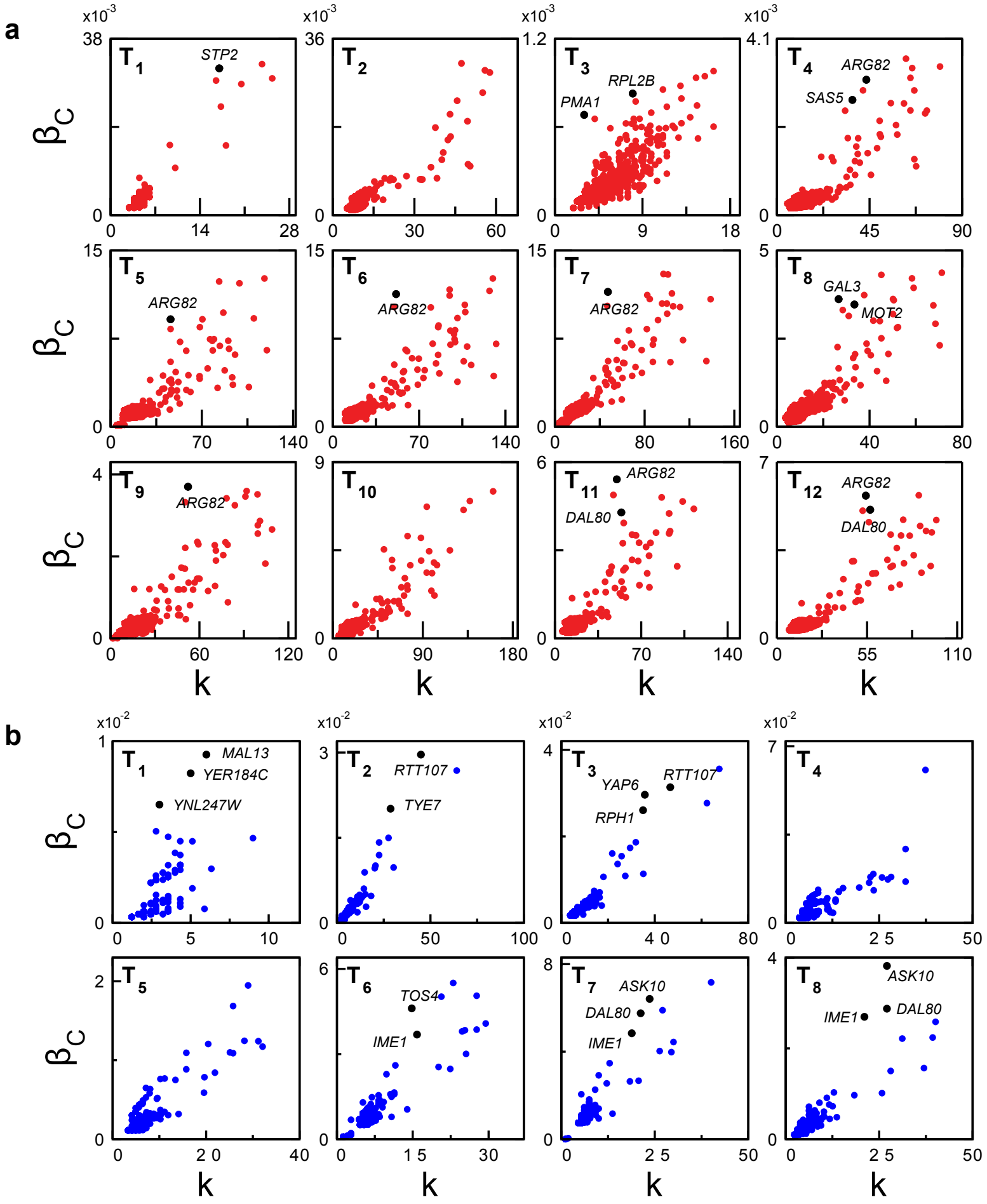
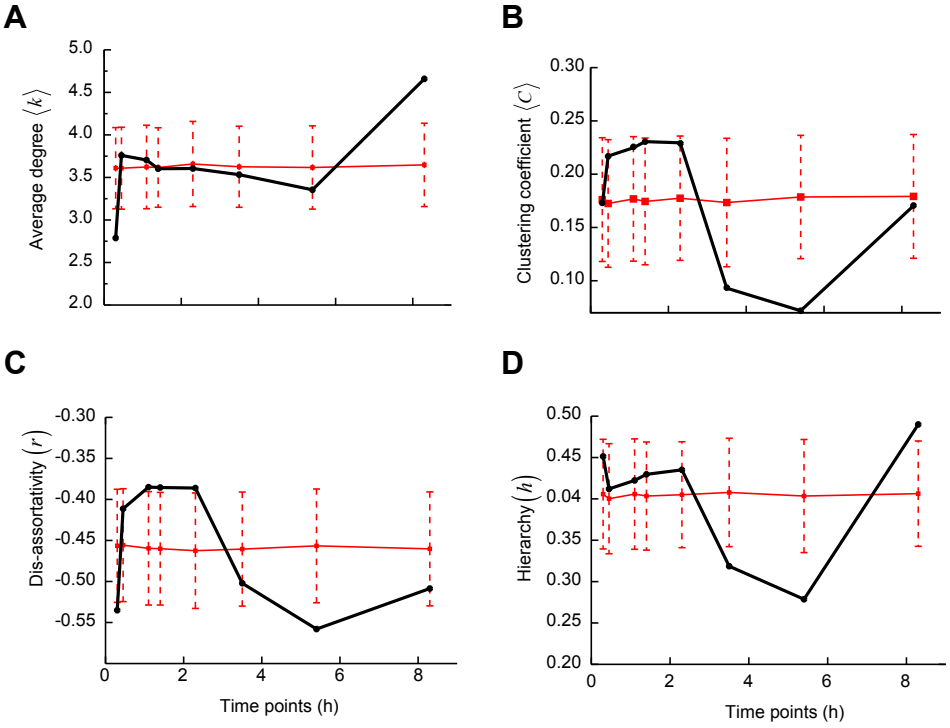

a
T1
STP2
T2
T3
RPL2B
PMA1
T4
ARG82
SAS5
T5
ARG82
T6
ARG82
T7
ARG82
T8
GAL3
MOT2
T9
ARG82
T10
T11
ARG82
DAL80
T12
ARG82
DAL80
βC
k
b
T1
MAL13
YER184C
YNL247W
T2
RTT107
TYE7
T3
YAP6
RTT107
RPH1
T4
T5
T6
TOS4
IME1
T7
ASK10
DAL80
IME1
T8
ASK10
IME1
DAL80
βC
k

Figure 4


A
Average degree ⟨k⟩
5.0
4.5
4.0
3.5
3.0
2.5
2.0
B
Clustering coefficient ⟨C⟩
0.30
0.25
0.20
0.15
0.10
C
Dis-assortativity (r)
-0.30
-0.35
-0.40
-0.45
-0.50
-0.55
0
2
4
6
8
Time points (h)
D
Hierarchy (h)
0.50
0.45
0.04
0.35
0.30
0.25
0.20
0
2
4
6
8
Time points (h)