

Networks and Circuits in Cell Regulation

Pasquale Palumbo¹, Gabriella Mavelli¹, Lorenzo Farina^{2§} and Lilia Alberghina^{3§}

1 Istituto di Analisi dei Sistemi ed Informatica "Antonio Ruberti", Consiglio Nazionale delle Ricerche, Viale Manzoni 30, 00185 Roma, Italy

2 Dipartimento di Informatica e Sistemistica "Antonio Ruberti", Università di Roma La Sapienza, Via Ariosto 25, 00185 Roma, Italy

3 Dipartimento di Biotecnologie e Bioscienze, Università di Milano-Bicocca, Piazza della Scienza 2, 20126 Milano, Italy

§ corresponding authors:

Lilia Alberghina lilia.alberghina@unimib.it

Dept. Biotechnology and Biosciences, University of Milano-Bicocca, piazza della Scienza 2, Milan, Italy;

Lorenzo Farina lorenzo.farina@uniroma1.it

Dipartimento di Informatica e Sistemistica "Antonio Ruberti", Università di Roma La Sapienza, Via Ariosto 25, 00185 Roma, Italy

SUMMARY

The complete understanding of the complex regulation network of eukaryotic cells is currently out of reach. The large amount of data emerging from high-throughput techniques requires the development of a new awareness of the conceptual and computational tools needed by cell biologists. Large-scale “omics” data are often represented as networks of interacting components but such representation is inherently static and, as such, cannot provide a realistic picture of the temporal dynamics of complex cellular functions. These difficulties suggest to move to a modelling strategy that explicitly takes into account both the wiring of the components and the task they perform. From an engineering perspective, this problem resembles that of “circuit analysis”. In this work we will concentrate on a particular but important biological circuit, the G1/S transition event in the yeast cell cycle and report, for the first time, a comparison between the network representation of a biological module and the corresponding circuit which accounts for its dynamical properties.

Introduction

It is widely recognized that only rarely cellular functions are determined by an individual gene product, more often they require dynamic interactions among a large number of molecular components (DNA, RNA, proteins and small molecules, etc.), modulated by internal and external cues (*Hartwell 1999, Lauffenburger 2000*). Although the concepts of “component”, “interaction” and “function” are largely subjective and depend upon the chosen aggregation level, a certain degree of reductionism is unavoidable being inherent to the process of scientific understanding (*Nurse 1997*). Nevertheless, a strict reductionist paradigm is no more acceptable (*Nurse 1997, Strohman 2000*) since biological functions are not determined by one level or component, but they are distributed, as emergent (or system-level) properties, over many levels and components. The problem then is posed on how “best” to describe these molecular networks and how fruitful a mathematical approach can be, in order to reach a reasonable computational predictive ability on the behaviour of the cell under changing genetic, metabolic and environmental conditions. It is plain that the choice of the “best” mathematical model must be considered with great care since models are essentially “tools for the mind” and not real entities.

Many high-throughput techniques (*e.g.* DNA microarrays, yeast two-hybrid and mass-spectrometry analyses) have been exploited to analyze macromolecular interactions in a number of organisms (*Uetz, 2001; Tyers, 2003*) and they have been used to generate the so-called large scale “interaction maps”. A very popular mathematical approach to study such interaction maps is certainly the formal concept of “graph”. A graph is composed of nodes denoting molecules (proteins, metabolites, ions, ...), and edges (arcs) representing interactions among them, or more generally, the relation linking two interacting nodes. The main interest in studying cellular processes as networks is given by the relevance of classifying the molecules that compose a network on the base of their connectivity, and not only on their biochemical activity. The foundations of graph theory analysis are very old, but only in the sixties of the past century that the “random graph” concept (*Erdos, 1960*) has been introduced. A random graph is such that the number of links associated to each node is defined by stochastic variables. Thus, the entire pattern of connections in a graph is described by a degree distribution, $P(k)$, which provides the number of nodes having k links. Two main types of distributions have been found in real networks: the Poisson and the power-law degree distributions (*Barabasi, 1999, 2004, Albert, 2000*) also called “scale-free”. The latter has the property that scaling by a constant, simply multiplies the original power-law relation by a constant. However, the biological interpretation of the scale-free distribution is currently quite controversial (*Wolf, 2002*). Other interesting features of real networks are the

“small-world” property, *i.e.* that any pair of nodes can be connected, on average, by a limited number of intermediates (Watts, 1998, 2003;Wagner, 2001;Nunes Amaral, 2000, Fell, 2000) and the presence of modules, or network clusters possessing a peculiar biological function. For example, in (Rasvatz 2002) it has been found that metabolic networks of many different organisms possess high clustering coefficients, and display a ‘hierarchical organization’, which is also present in protein interaction network (Jeong, 2001), made of small modules inside larger modules and so on.

In the yeast *Saccharomyces cerevisiae* protein-protein interaction (PPI) network up to 3000 interactions among about 1000 proteins have been derived from combined data (Uetz, 2000; Xenarios, 2000). In Figure 1A we reported the PPI network of the main proteins involved in the G1 to S transition during the yeast cell cycle. The interactions displayed comes from the high-confidence dataset provided by Tyers group (Reguly *et al* 2006). The map of the yeast interactome, being inherently static, does not provide a complete picture of cellular regulation. In fact, biological functions derive from the unfolding of dynamic features, from gene expression to protein interactions and catalytic events that, in fact, exhibit highly specific temporal profiles and spatial compartmentalizations. Interactions represented in such static maps do not necessarily occur simultaneously, thus the connectivity diagram alone is not sufficient to describe the dynamics of the network, but it is generally useful to reveal the underlying structural constraints of the system.

In conclusion, the application of computational methods to network modelling using “omics” data and the associated network representation, despite the great experimental technological effort and the sophisticated bioinformatics technology developed, cannot provide a realistic picture of the execution and control systems of complex cellular functions, nor it has been possible to predict their dynamics at changing environmental conditions.

These difficulties suggest to move to a modelling strategy that explicitly takes into account both the wiring of the components and the task they perform. From an engineering perspective, this problem resembles that of “circuit analysis”. A circuit, in the language of electronic engineering, is an interconnection of elementary devices purposely designed to perform a given task (the definition of a task therefore characterizes the circuit representation as compared to the network one). The function of the electronic circuit, whereas precisely characterized in terms of its components, is not due to each single component, but “emerges” as a collective behaviour from the wiring of interconnections. As a matter of fact, the design of a complex electronic circuit fully exploits the modular structure and the aim of the designer is to appropriately interconnect “modules” made by a certain number of circuits in order to satisfy pre-specified functional requirements. The final outcome is typically composed of a nested sequence of modules defining new high-level structure and function, and so on. A single module is therefore defined by a functional capacity (such as, for

example, a constant current source) and the circuit is given by the specific interconnections of components (resistors, capacitors and transistors). Two considerations are in order: first, it is important to keep in mind that different circuits may well produce the same functional unit and, second, that the work of the designer is based on the availability of very accurate mathematical descriptions. In fact, each individual component has a specific “constitutive equation” and therefore the input-output behaviour of each module can be obtained by solving a set of differential equations over the network interconnections. In sum, the understanding of complex electronic devices requires the identification of basic modules, the knowledge of how and why such modules are assembled and a mathematical description of each component to accurately simulate the circuit behaviour.

So far, the engineering point of view: what about “molecular circuits” and biological functions? The modular structure of an electronic circuit may provide an intriguing and possibly useful “metaphor” for the understanding of biological systems (Sauro 2004). In fact, a cell can be viewed as a system composed of many interconnected modules coordinately performing specific biological functions (metabolism, signalling, transcription, growth, cycle, autophagy, apoptosis, differentiation etc.), each module being characterized by a network of interacting molecules (Hartwell, 1999, Eisenberg, 2000). In this perspective, one has to determine the *structure* (i.e. the topology) and the *dynamics* (i.e. the circuit’s dynamic behavior) of the circuit. In summary, the circuit representation should allow to account for the biological functions as emergent properties of the system, better than the network representation. Moreover, the knowledge of biological constitutive equations (such as mass law equations) may be useful to find more detailed models able to provide a fine grain mathematical description on a smaller scale. As an example of a biological circuit, we mention the “osmostat” active during the yeast response to osmotic shock (Klipp *et al*, 2005), which is given by a MAP kinase signalling pathway, an activation of gene expression of a specific metabolic pathway, with the adaptation of cellular metabolism to produce an osmotic active end product (glycerol), together with the concurrent regulation of the exchange activity of the aquaglyceroporin, an osmolarity regulated glycerol channel. The quantitative feature of the “osmostat” is the value of the internal osmotic pressure that is maintained in homeostasis by its activity, just as the temperature of a room is maintained constant by a thermostat. It is also worth noting that the emergence of a cell fate/function from a biological circuit may also occur as the attractor of a dynamic system. For example, in (Huang *et al*. (2005)) the authors showed that trajectories of neutrophil differentiation converge to a common state from different directions of an high-dimensional gene expression state space. Moreover, also the emergence of a “checkpoint” can be obtained using a simple mathematical model consisting of mixed feedback loops (Sriram 2007).

In this work we will concentrate on a particular but important biological circuit, the G1/S transition event in the yeast cell cycle, aiming to compare the network and circuit representations of such module and discuss their relative contribution to the understanding of a complex cellular function.

Results

The G1 to S transition is a fundamental module of yeast cell cycle, since it allows the coordination of growth and cycle progression by requiring the attainment of a critical cell size to enter into S phase, and characterized by both the onset of budding and of DNA replication (see *Alberghina et al 2009a*, for a recent review). If we search for all known players of the G1 to S transition in the yeast interactome map (*Reguly, 2006*), the results are reported in Fig. 1A. The proteins of the G1 to S transition involved both in cell sizer mechanism (*Barberis et al, 2007*) and in the onset of DNA replication (*Difley et al, 2004; Tanaka et al 2007*) are found disperse in a hub of the yeast interactome centered around Cdc28, the cyclin-dependent protein kinase that powers cycle progression. Moreover, even considering the subnetwork of their interactions (Fig. 1B) using manually curated interaction data, it is clear that no peculiar mechanistic insight can be derived from visual inspection.

The mathematical model of the cell sizer mechanism active during the G1 to S transition in budding yeast has been reported by one of our laboratories (*Barberis et al, 2007*). The critical cell size to enter S phase has been shown to be an emergent property of the G1 to S network, modulated by growth rate, as experimentally observed (*Barberis et al, 2007*). Besides, the G1 to S network works as a *sizer plus timer circuit* (shown in Figure 2), structure confirmed later by other authors (*Di Talia et al, 2007*), thereby explaining the dependence of the critical cell size from growth rate. Given the novelty of circuit representation in biology the notation used for describing the sizer plus timer structure of Fig. 2 is very crude, since it has not reached yet the uniqueness and precision of notation achieved for electronic circuits.

It is worth stating that different emergent properties may be experimentally generated by the sizer plus timer circuit following perturbations of the genetic or metabolic conditions of yeast cells. These perturbations have the common effect to modulate the strength of binding for the various interactors of the circuit, not only by altering the expression level of the various involved proteins or their degradation rate, but also by the action of mono and multisite phosphorylation/dephosphorylation reactions (reviewed in *Alberghina et al, 2009, b*).

In the following we will analyze, by simulation of the mathematical model described by Barberis et al (2007), the effect of changing the binding affinity of the three relevant pairs of interactors of the circuit: Far1/Cln3, Whi5/SBF-MBF, Sic1/Clb5-Cdc28, as well as the rate of degradation of Sic1, which is dependent upon the interaction of Sic1 with its degradation complex SCF.

The yeast SCF complex is composed by subunits Skp1, Cdc53 and Rbx1, which interact with an E2 enzyme, usually Cdc34 (*Pickart C.M., 2001*). The substrate specificity is determined by a large family of adaptors, the F-box proteins, each of which targets a limited set of substrates. The SCF^{Cdc4} complex (containing the F-box protein Cdc4) degrades several important regulatory proteins, among which the Cdk inhibitors Sic1 and Far1. In several of the tested perturbations a relevant role is played by Sic1, whose impairment of degradation, together with a block of growth, brings to a cycle arrest in Sic1 as emergent property of the network (reviewed in *Alberghina et al, 2009, b*). It has been shown that both osmotic stress and TOR pathway inhibition induce a phosphorylation of Cdc4 in Thr 173, that hinders Sic1 binding to the degradation complex SCF (*Escote et al, 2004; Zinzalla et al, 2007*). Also Cdc34 activity can be modulated by phosphorylation: Ck2 phosphorylation of Ser130 and Ser 167 in the active site of Cdc34 is required for its optimal ubiquitin-binding and conjugating activity (*Cocchetti et al, 2008*).

In conclusion, the interactions of Sic1, an intrinsically disordered protein (*Brocca et al, 2009*) with various constituents of SCF, may affect its degradation rate and therefore the release of free active Clb5,6.Cdk, that promotes the onset of DNA replication (*Difley, 2004*).

The emergent properties of the cell-sizer circuit presented in Fig. 2 are the setting of the critical cell size (P_s) and of the closely connected length of the timer (T) that is the duration of G1 phase of daughter cells (*Di Talia et al, 2007*). The interest of this circuit is that both emergent properties can be measured in a quantitative way and may be compared easily and directly with experimental readouts. Figs. 3 and 4 report the changes of P_s values as a function, over several orders of magnitude, of the binding activities for three couples of interactors: Far1/Cln3, Whi5/SBF.MBF and Sic1/Clb5,6.Cdk1. Computations are based on the mathematical model reported in [Barberis et al (2007)] (see Suppl. Material for details).

At a fast growth rate (Fig. 3A) the P_s value is substantially low and constant, when the Sic1 affinity constant increases (even of two order of magnitude). The changes of the other affinities have negligible effect, except for the case of very high Far1 and Whi5 affinities (both increased of two order of magnitude), conditions in which a small further increase of P_s is observed. On the other hand, when Sic1 affinity is decreased, P_s substantially increases and, also in this case, the

other two binding affinities do not produce substantial variations of P_s , apart from the case in which both Far1 and Whi5 affinities increased by two orders of magnitude.

At a low growth rate (Fig. 3B), we have roughly the same pattern of relations among the three affinities variations and the corresponding changes in the P_s values, this time concerning a reduced range of variations: indeed, when Sic1 affinity increases, we actually observe no changes in the low values, regardless to the variations of the other two binding affinities; substantial variations (even if referred to a smaller extent) occur only when Sic1 affinity decreases, and also in this case further variations of the other two binding affinities do not provide significant variations in P_s . In conclusion the setting of P_s is strongly modulated by the growth rate and by the Sic1 affinity constant, a much smaller role being played by the interactions of Whi5 and Far1.

When the rate of degradation of Sic1 is reduced by 2 orders of magnitude (Fig. 4A), at a high rate of growth, the setting of P_s is substantially unaffected (compare Fig. 4A with Fig. 3A). Instead, if the rate of degradation of Sic1 is increased of 2 orders of magnitude with respect to the wild type (Fig. 4B), the P_s setting is substantially increased, showing again a stronger sensitivity to Sic1 affinity (especially when it is decreased below its wild type value) with respect to the other two affinities.

Similar results are obtained for the setting of the timer duration (Figs. S-1A and S-1B), which seems quite insensitive to the changes of the Whi5 and Far1 affinities (with the small exception of very low values of Far1 affinity for the fast and for some of the low growth rates), whilst the Sic1 affinity strongly controls the duration of the timer. When the degradation of Sic1 is reduced of 2 orders of magnitude in the fast growth rate there are no appreciable changes in timer duration (Fig. S-2A), whilst in the case of a degradation of Sic1 increased of 2 orders of magnitude in the fast growth rate the variability of the timer is reduced to an almost uniform pattern at changing values of the three affinities (Fig. S-2B).

Discussion

The understanding of complex cellular functions (cell cycle, death, differentiation, senescence, transformation, etc.) is the challenge of modern biology. High-throughput technologies are able to provide a wealth of data on genome, transcriptome, proteome, interactome, metabolome, etc., on many cellular types both in normal and in altered or pathological conditions. As a consequence, there is a urgent need to find out the most profitable way for organizing such large amount of data to extract useful knowledge about the quantitative dynamics of complex biological functions, under different internal and external conditions.

The “network metaphor” (Palumbo et al. 2006) has attracted many researchers in recent years. An example of a network model arising from large scale data, is the probabilistic Bayesian approach (Jansen, 2003). In this model, nodes represent molecules and edges (interactions) the likelihood that one molecule affects another one. This technique has proved to be useful, for example, to increase the accuracy in the prediction of phosphorylation events by treating the probability as an approximation for the proximity between a kinase and its potential substrate (Linding, 2007). In a recent study (Chuang, 2007), by applying a protein network-based approach, genes known to be involved in breast cancer – not showing any significant differential expression – have been shown to play a central role in the protein-protein interactions network topology by interconnecting genes that are differentially expressed in patients with metastasis thereby suggesting a potential usefulness of network approach to identify new markers for metastatic cancer. Recently, another “network approach” revealed that biological systems contain characteristic motifs that do not occur as frequently as in random networks (Alon, 2007). Moreover, modules, or network clusters, have been identified in biological metabolic networks (Rasvatz, 2002). Lastly, we mention the study of the relationship between essentiality and topological properties for the *Saccharomyces cerevisiae* protein-protein interaction network (Jeong, 2001) and, for the metabolic network, the finding that mutation lethality is related to the absence of an alternative path which can connect the pair of nodes after the deletion of their link (enzyme) (Palumbo et al., 2005). Despite the success achieved by this approach, the network approach has many unavoidable limitations. The most important certainly being the lack of quantitative information about the dynamic of the process. To put it simply, the wiring of a network does not provide information on the flows actually present in the system.

This paper reports for the first time a comparison between the network representation of a biological module and the corresponding circuit which accounts for its dynamical properties. More precisely, we studied a module composed by the biochemical events occurring during the G1 to S transition in a yeast cell cycle, and the quantitative aspect of our analysis focused on the setting of the critical cell size and the duration of the timer/G1 phase in daughter cells. The most relevant results of our computational analysis of the G1 to S (sizer plus timer) circuit can be summarized as follows:

- The same circuit topology may yield many different functional properties (as monitored in our case by the computed values of P_s and T) in response to changes of the strength of interactions between different elements of the circuit.
- There is not a simple linear relationship between the value of any given affinity constant for a given pair of interactors and the emergence of the functional property.

By contrast, the value of P_s emerges from the collective co-ordinated action of the various elements of the entire circuit. However, some elements of the circuit, such as Sic1 or the rate of growth, are shown to possess a more relevant effect than that of other players on the establishment of the functional property (P_s).

Since systems biology may be defined as the “science that discovers the principles underlying the emergence of the functional properties of living organisms from interactions between macromolecules” (Westerhoff and Alberghina, 2005), the circuit representation appears to offer to systems biology a more incisive and profitable approach than simple networks. Moreover, our investigation makes clear that circuits, like the G1 to S cycle module or the “osmostat”, are made by the interaction of several different biochemical pathways: from transcription to protein degradation, from protein catalytic activity (especially in this case phosphorylation / dephosphorylation) to nuclear/cytoplasmic import or export, from regulated protein synthesis to control of plasma membrane exchange activities.

Finally, the most interesting insight coming from the results collected in this paper may offer a new appreciation on the evolution of biological systems. Natural selection is taken to modulate the evolutionary fitness of an organism by monitoring the suitability of its performance to the challenges posed by the environment. Performance is closely linked to function and in this paper we have shown that function is generated as a collective emergent property of an ensemble of gene products, in which the actual behaviour of the system has a non-linear dependence on the strength of interactions found in the set of gene products organized in a specific architecture. Many different values of affinity between different interactors may generate the same quantitative value of a function. Several robustness devices have been described (Kitano, 2004), but no indication has yet been reported on the possibility that the same circuit structure may alter a robustness behaviour, like the one reported in this paper. A complex biological process is expected to derive from the integration of a number of circuits, each one providing a specific functional property: for instance, in cell cycle we have shown a circuit for the setting of the critical cell size, which may work together with other circuits determining different relevant features to be tested by natural selection (for example: accuracy of DNA replication, coherence between signalling and cell cycle progression etc.).

In conclusion, circuits are shown in this paper to offer great advantages for a systems biology analysis of complex cellular functions. The recursive approach typical of systems biology will enrich the definition of the molecular events accounted for by each circuit and therefore yield a

more and more accurate predictive ability of the circuit representation under an increasing large number of circumstances.

Experimental procedures

Model description

The mathematical model of the G_1/S transition here adopted consists of a set of 34 Ordinary Differential Equations (ODE), detailing about production and degradation of mRNA and proteins as well as the formation of dimeric and trimeric protein complexes which are involved in and rule the G_1/S transition of a single cell, (Barberis et al., PLoS, 2007). It accounts for both nuclear and cytoplasmic compartments, whose volumes are state variables: the nuclear volume is kept constant while the cytoplasmic volume exponentially increases with time. All the other state variables are compound concentrations related to mRNAs, transcription factors, cyclins, cyclin-dependent-kinases, inhibitors and different complexes. The model comprises 54 parameters: some of them are estimated from data, some others are set according to proper biological reasons. One of the parameters is the growth rate of the cell, by varying which different framework can be considered (a fast/slow growth rate may be referred to a glucose/ethanol environment). Other crucial parameters, are the affinity constants of three couples of interactors, such as Far1/Cln3, Whi5/SBF.MBF and Sic1/Clb5,6.Cdk1, as well as the clearance rate of Sic1: they will be referred to as k_{30} , k_{34} , k_{32} , k_{18} respectively, according to the original notation adopted in (Barberis et al. 2007), which has been reported for the ease of the reader in the supplementary material together with the wild type setting of the system parameters (including both model parameters and the nontrivial initial conditions). New simulations are provided in the present paper, referring to a different setting of the above mentioned parameters.

Once all the system parameters are set, a single cell model is fixed, by running which we obtain the time evolutions of all the state variables. From this stream of simulated data, the following items may be computed:

- **Timer size:** it is the time length between:
 - the time instant T_1 , according to which, in the nucleus, the active kinase Cdk1-Cln3 overcomes in concentration its inactive form Cdk1-Cln3-Far1, thus starting the chain of biochemical events eventually leading the cell to exit the G_1 phase and enter the S phase;
 - and the time instant T_2 , according to which the nuclear concentration of the active kinase Cdk1-Clb5,6 reaches half of its maximal value (such a value is comparable

with the time instant according to which half of the replication origins are activated and is associated to the onset of DNA replication, the transition time between G_1 and S phase).

- **Critical cell size:** it is the cell size (i.e. the sum of the nuclear and cytoplasmic volume) computed at time T_2 , and represents the protein content at the onset of DNA replication.

Both the timer size (T_2-T_1) and the critical cell size (P_s) are readily obtained by a single simulation, being related to the model state variable evolutions.

In order to simulate an N cells population, a normal distribution is considered for each of the model parameters (54 model parameters and 7, among 34, nontrivial initial conditions): the mean values are given by the wild type values which have been previously estimated/assessed; the standard deviation is 0.287 times each nominal value, according to (Barberis et al. 2007). Thus by taking a realization from the parameter distributions, we consider a single cell from the population: for any realization (say, for any single cell simulation) we compute the timer size T_2-T_1 and the critical cell size P_s . These values are, finally, averaged over a set of 100 cells population.

We have performed different sets of simulations, each referred to a given choice of the meta-parameters (i.e. the mean values and standard deviations of the normal distributions related to the system parameters). Two main frameworks refer to a *fast growth rate* ($k_{\text{growth}} = 0.0051 \text{ min}^{-1}$, as in the glucose environment of (Barberis et al. 2007)) and to a *slow growth rate* ($k_{\text{growth}} = 0.0023 \text{ min}^{-1}$, as in the ethanol environment of (Barberis et al. 2007)). Within these two frameworks, different cell populations have been run according to different settings of the mean values for three of the above mentioned parameters k_{30} , k_{34} , k_{32} : besides their wild type configuration, two orders of magnitude greater and smaller have been considered (i.e. 10-, 100-, 0.1-, 0.01-fold the wild type). This way we have $5 \times 5 \times 5 = 125$ different cell populations for each triple (k_{30} , k_{34} , k_{32}) of the affinities, each of which provides an average value for the timer size and the critical cell size in both the fast and slow growth rate frameworks. The average value of the Sic1 degradation rate, k_{18} , is kept at his wild type value for this set of simulations.

A second set of simulation has been carried on according to the only fast growth rate environment. This time different frameworks are characterized by two different settings of the mean value of Sic1 degradation (parameter k_{18}): we have considered a case of strong overexpression (100-fold the wild type) and of strong inhibition (0.01-fold the wild type). Within these two frameworks, different cell populations have been run according to different settings of the mean values for parameters k_{30} , k_{34} , k_{32} , the same chosen for the previous *in silico* experiments. Thus we have again $5 \times 5 \times 5 = 125$ different cell populations for each triple (k_{30} , k_{34} , k_{32}) of the affinities, each of which

provides an average value for the timer size and the critical cell size in both the overexpressed and inhibited cases for k_{18} .

The following criterion has been adopted in order to picture the results, which are collected in 4 sets of 125 average values for both the experimental settings: timer size and cell size for both fast and slow growth rate frameworks. Each picture is a three-dimensional grid in a logarithmic scale with respect to the three affinities k_{30} , k_{34} , k_{32} . A sphere referring to the timer/cell size is positioned correspondingly to a triple of values of k_{30} , k_{34} , k_{32} (their wild type is central in each figure). The color of the spheres gives us information about the value: it is assigned according to the RGB scale as a triple of values [r g b] coming as an affine combination of pure red [1 0 0], pure green [0 1 0] and pure blue [0 0 1]. Lowest values are referred to a pure green; highest values are referred to a pure blue.

References

- [] Hartwell LH, Hopfield JJ, Leibler S & Murray AW (1999) From molecular to modular cell biology. *Nature* **402**, 47-52.
- [] Lauffenburger DA (2000) Cell signalling pathways as control modules: Complexity for simplicity? *PNAS* **97**, 5031-5033.
- [] Nurse P (1997) Reductionism: The ends of understanding. *Nature* **387**, 657-658.
- [] Strohman RC (2000) Genetic determinism as a failing paradigm in biology and medicine. In *Promoting Human Wellness: Frontiers for Research, Practice and Policy* (Jamner MS & Stokol D, eds), 302-322. University of California Press.
- [] Uetz P (2001) Two-hybrid arrays. *Current Opinion in Chemical Biology* **6**, 57-62.
- [] Tyers M & Mann M (2003) From genomics to proteomics. *Nature* **422**, 193-197.
- [] Erdos P & Rényi A (1960) On the evolution of random graphs *Publ Math Inst Hungar Acad Sci* **5**, 17-61.
- [] Barabasi A-L & Albert R (1999) Emergence of scaling in random networks. *Science* **286**, 509-512.
- [] Barabasi A-L & Oltvai ZN (2004) Network biology: understanding the cell's functional organization. *Nat Rev Genet* **5**, 101-113.
- [] Watts DJ & Strogatz SH (1998) Collective dynamics of "small-world" networks. *Nature* **393**, 440-442.
- [] Watts DJ (2003) *Small Worlds: the dynamics of networks between order and randomness*. Princeton University Press, Princeton, New Jersey, USA.

- [] Wagner A & Fell DA (2001) The small world inside large metabolic networks. *Proc R Soc Ser B* **268**, 1803-1810.
- [] Nunes Amaral LA, Scala A, Barthélemy M & Stanley HE (2000) Classes of small-world networks. *Proc Natl Acad Sci USA* **97**, 11149-11152.
- [] Fell DA & Wagner A (2000) The small world of metabolism. *Nat Biotechnol* **18**, 1121-1122.
- [] Rasvatz E, Somera AL, Mongru DA, Oltvai ZN & Barabasi A-L (2002) Hierarchical organization of modularity in metabolic networks. *Science* **297**, 1551-1555.
- [] Jeong H, Mason SP, Barabasi A-L & Oltvai ZN (2001) Lethality and centrality in protein networks. *Nature* **411**, 41-42.
- [] Albert R, Jeong H & Barabasi A-L (2000) Error and attack tolerance of complex networks. *Nature* **406**, 378-382.
- [] Wolf YI, Karev G & Koonin EV (2002) Scale-free networks in biology: new insights into the fundamentals of evolution? *Bioessays* **24**, 105-109.
- [] Uetz P, Giot L, Cagney G, Mansfield TA, Judson RS, Knight JR, Lockshon D, Narayan V, Srinivasan M, Pochart P (2000) A comprehensive analysis of protein–protein interactions in *Saccharomyces cerevisiae*. *Nature* **403**, 623–627.
- [] Xenarios I, Rice DW, Salwinski L, Baron MK, Marcotte EM & Eisenberg D (2000) DIP: the database of interacting proteins. *Nucleic Acids Res* **28**, 289-291.
- [] Eisenberg D, Marcotte EM, Xenarios I & Yeates TO (2000) Protein function in the post-genomic era. *Nature* **405**, 823- 826.
- [] Alberghina L, Coccetti P & Orlandi I (2009a) Systems biology of the cell cycle of *Saccharomyces cerevisiae*: from network mining to system-level properties. *Biotechnol Adv* **27**, 960-978.
- [] Reguly T, Breitkreutz A, Boucher L, Breitkreutz B-J, Hon GC, Myers CL, Parsons A, Friesen H, Oughtred R, Tong A, Stark C, Ho Y, Botstein D, Andrews B, Boone C, Troyanskaya OG, Ideker T, Dolinski K, Batada NN & Tyers M (2006) Comprehensive curation and analysis of global interaction networks in *Saccharomyces cerevisiae*. *J of Biology* **5**, 11.1- 11.28.
- [] Barberis M, Klipp E, Vanoni M & Alberghina L (2007) Cell size at S phase initiation: an emergent property of the G1/S network. *PLoS Comp Bio* **3**, 649-666.
- [] Brümmer A, Zinzalla V, Salazar C, Alberghina L & Höfer T (2009) Mathematical modelling of DNA replication in budding yeast identifies multi-site phosphorylation as a coherence and robustness mechanism. *To appear*.
- [] Di Talia S, Skotheim JM, Bean JM, Siggia ED, & Cross FR (2007) The effects of molecular noise and size control on variability in the budding yeast cell cycle. *Nature* **448**, 947-951.

- [] Alberghina L, Höfer T & Vanoni M (2009b) Molecular networks and system-level properties. *J of Biotechnol* in press, doi:10.1016/j.jbiotec. 2009.07.009.
- [] Pickart CM (2001) Mechanisms underlying ubiquitination. *Annu Rev Biochem* **70**, 503-533.
- [] Escote X, Zapater M, Clotet J & Poses F (2004) Hog1 mediates cell cycle arrest in G1 phase by the dual targeting of Sic1. *Nat Cell Biol* **6**, 997-1002.
- [] Zinzalla V, Graziola M, Mastriani A, Vanoni M & Alberghina L (2007) Rapamicyn mediated G1 arrest involves regulation of the Cdk inhibitor Sic1 in *Saccharomyces cerevisiae*. *Mol Microbiol* **63**, 1482-1494.
- [] Coccetti P, Tripodi F, Tedeschi G, Nonnis S, Merin O, Fantinato S, Citrulli C, Vanoni M & [] Alberghina L (2008) The Cdk2 phosphorylation of catalytic domain of Cdc 34 modulates its activity at the G1 to S transition in *Saccharomyces cerevisiae*. *Cell Cycle* **7**, 1391-1401.
- [] Brocca S, Samalikova M, Uversky VN, Lotti M, Vanoni M, Alberghina L & Grandori R (2009) Order propensity of an intrinsically disordered protein, the cyclin-dependent-kinase inhibitor Sic1. *Proteins* **76**, 731–746.
- [] Huang S, Eichler G, Bar-Yam Y & Ingber DE (2005) Cell fates as high-dimensional attractor states of a complex gene regulatory network. *Physical Review Letters* **94**, 128701
- [] Sriram K, Bernot G & Kepes F (2007) A minimal mathematical model combining several regulatory cycles from the budding yeast cell cycle. *IET Syst. Biol.*, **1**, 326-341
- [] Jansen R, Yu H, Greenbaum D, Kluger Y, Krogan NJ, Chung S, Emili A, Snyder M, Greenblatt JF & Gerstein M (2003) A Bayesian Networks Approach for Predicting Protein-Protein Interactions from Genomic Data. *Science*, **302**, 449-453
- [] Linding R, Jensen LJ, Ostheimer GJ, van Vugt MA, Jørgensen C, Miron IM, Diella F, Colwill K, Taylor L, Elder K, Metalnikov P, Nguyen V, Pasculescu A, Jin J, Park JG, Samson LD, Woodgett JR, Russell RB, Bork P, Yaffe MB & Pawson T (2007) Systematic discovery of in vivo phosphorylation networks. *Cell*, **29**, 1415-1426
- [] Chuang HY, Lee E, Liu YT, Lee D & Ideker T (2007) Network-based classification of breast cancer metastasis. *Mol Syst Biol.*, **3**, 140
- [] Alon U (2007) Network motifs: theory and experimental approaches. *Nat. Rev. Genet.*, **8**, 450-461
- [] Palumbo MC, Colosimo A, Giuliani A & Farina L (2005) Functional essentiality from topology features in metabolic networks: a case study in yeast. *FEBS Letters*, **579**, 4642-4646
- [] Klipp E, Nordlander B, Kruger R, Gennemark P & Hohmann S (2005) Integrative model of the response of yeast to osmotic shock. *Nat. Biotechnol.*, **23**, 975-982.

- [] Palumbo MC, Farina L, Colosimo A, Tun K, Dhar PK & Giuliani A (2006) Networks everywhere? Some general implications of an emergent metaphor. *Current Bioinformatics*, **1**, 219-234
- [] Kitano H (2004) Biological robustness. *Nature Reviews Genet.*, **5**, 826-837.
- [] Westerhoff HV & Alberghina L (2005) Systems biology: did we know it all along? in L. Alberghina and HV Westerhoff (Eds.) “*Systems biology: Definitions and perspectives*”, Springer Verlag, Berlin-Heidelberg 2005, 3-9.
- [] May RM (2004) Uses and abuses of mathematics in biology, *Science*, **303**, 790-793
- [] Difley JF (2004) Regulation of early events in chromosome replication. *Curr. Biol* **14**, R778-R786
- [] Tanaka S, Tak YS, Araki H (2007) The role of Cdk in the initiation step of DNA replication in eukaryotes. *Cell Division* **2**, 16.
- [] Alberghina L, Rossi RL, Querin L, Wanke V, Vanoni M (2004) A cell sizer network involving Cln3 and Far1 controls entrance into S phase in the mitotic cycle of budding yeast. *J Cell Biol* **167**(3):433-43
- [] Sauro HM & Kholodenko BM (2004) Quantitative analysis of signaling networks. *Progress in Biophysics & Molecular Biology* **86**, 5–43

FIGURES

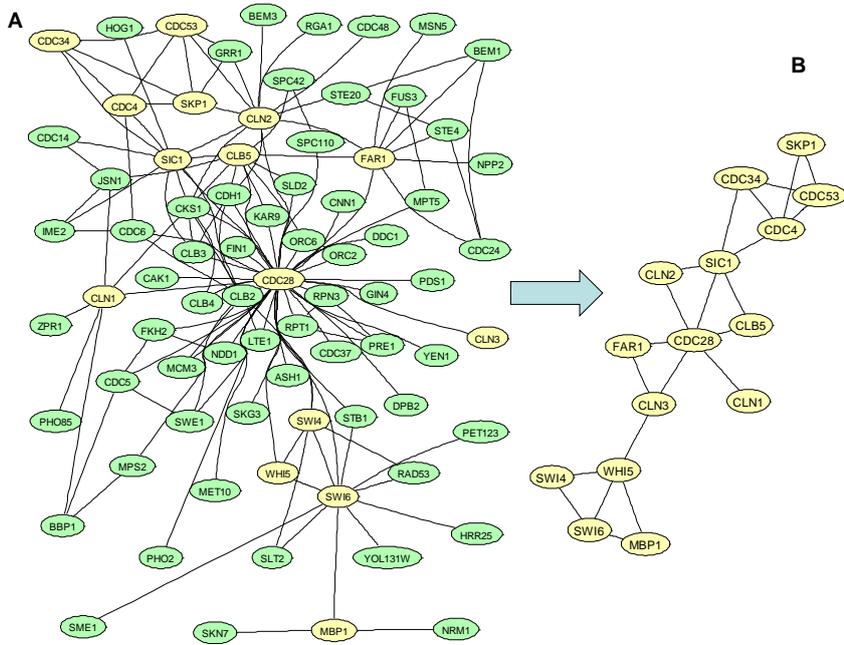


FIG. 1

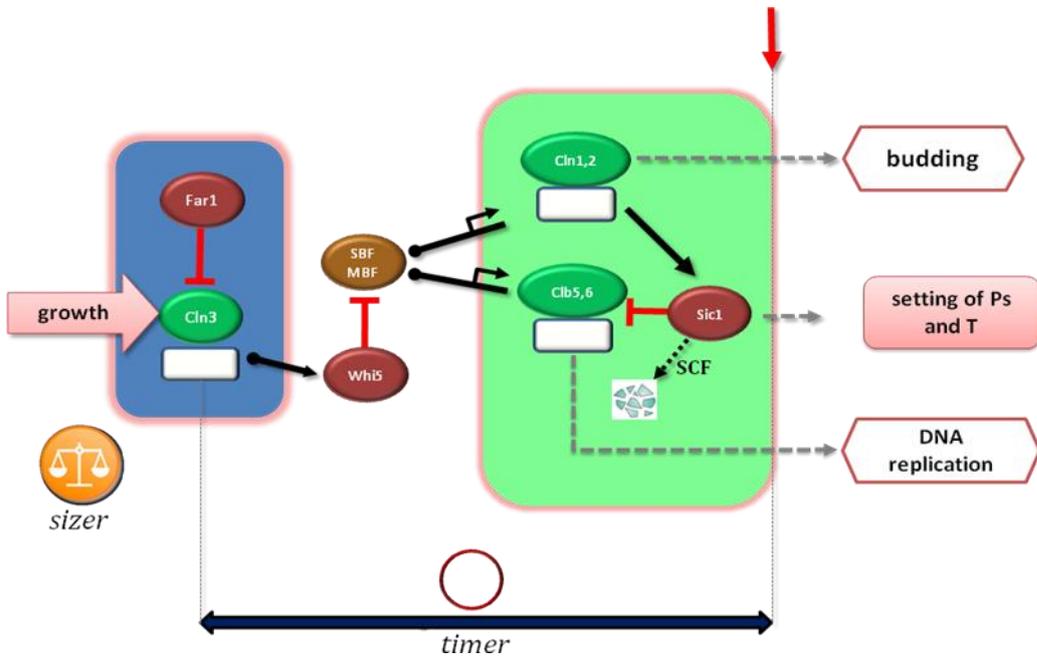


FIG. 2

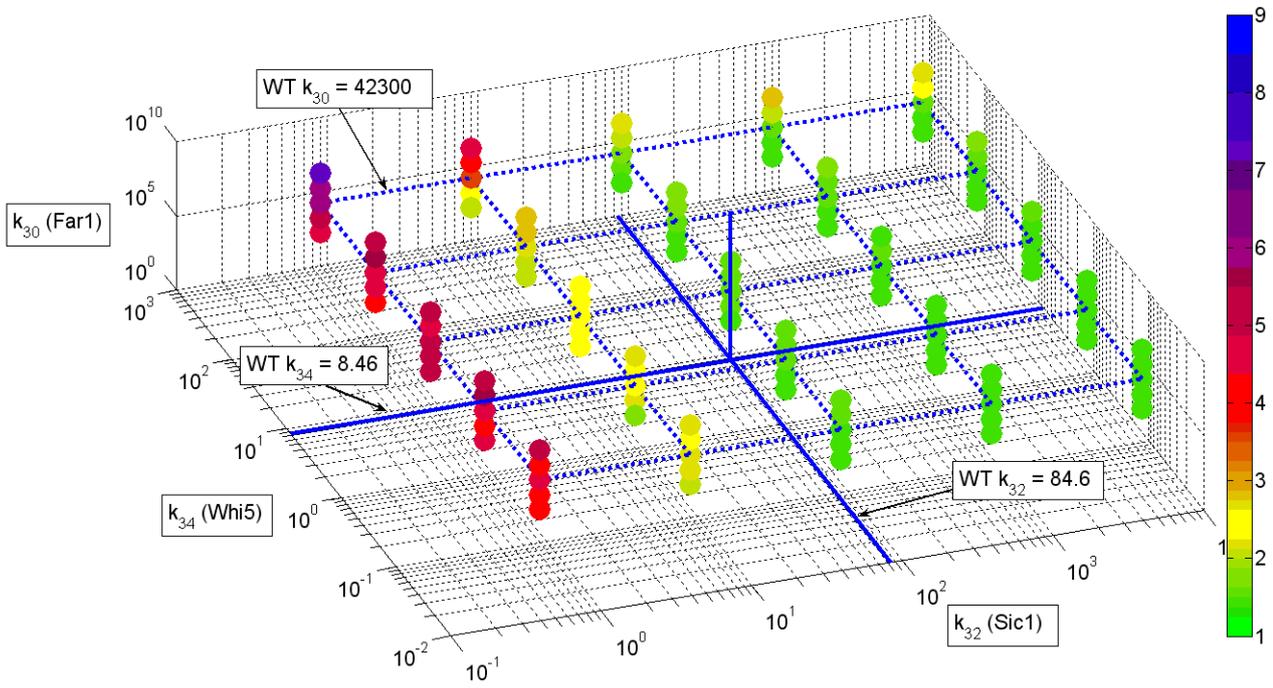


FIG. 3A

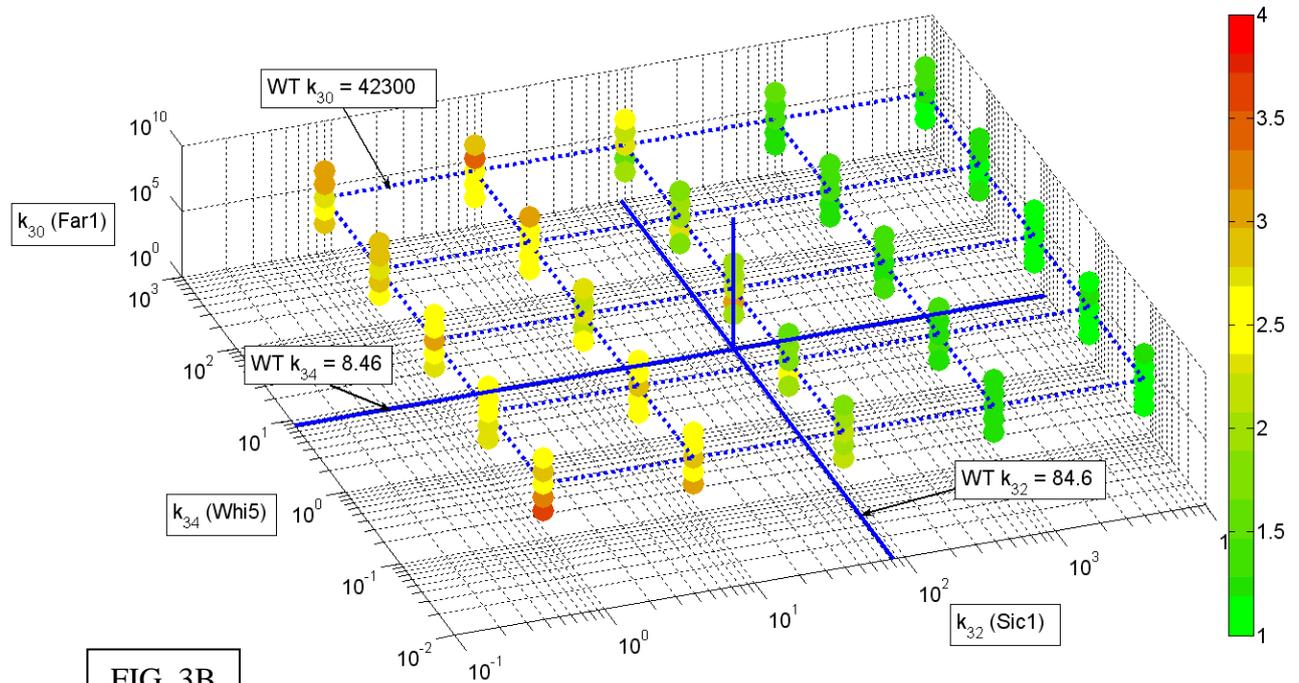


FIG. 3B

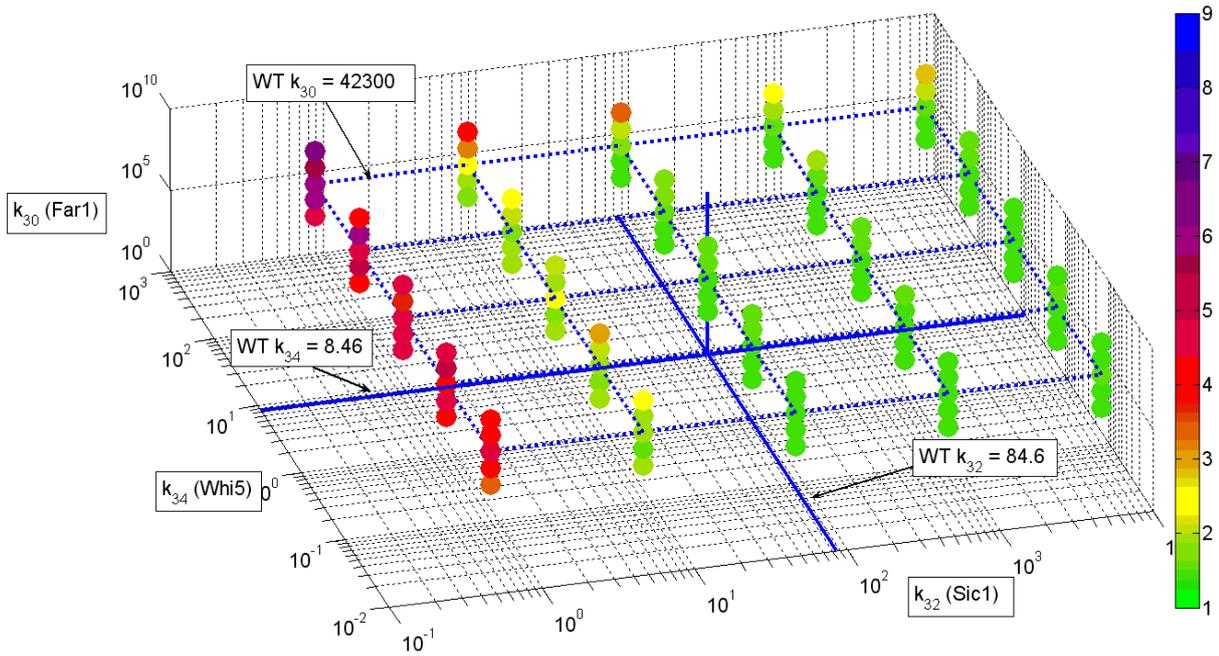


FIG. 4A

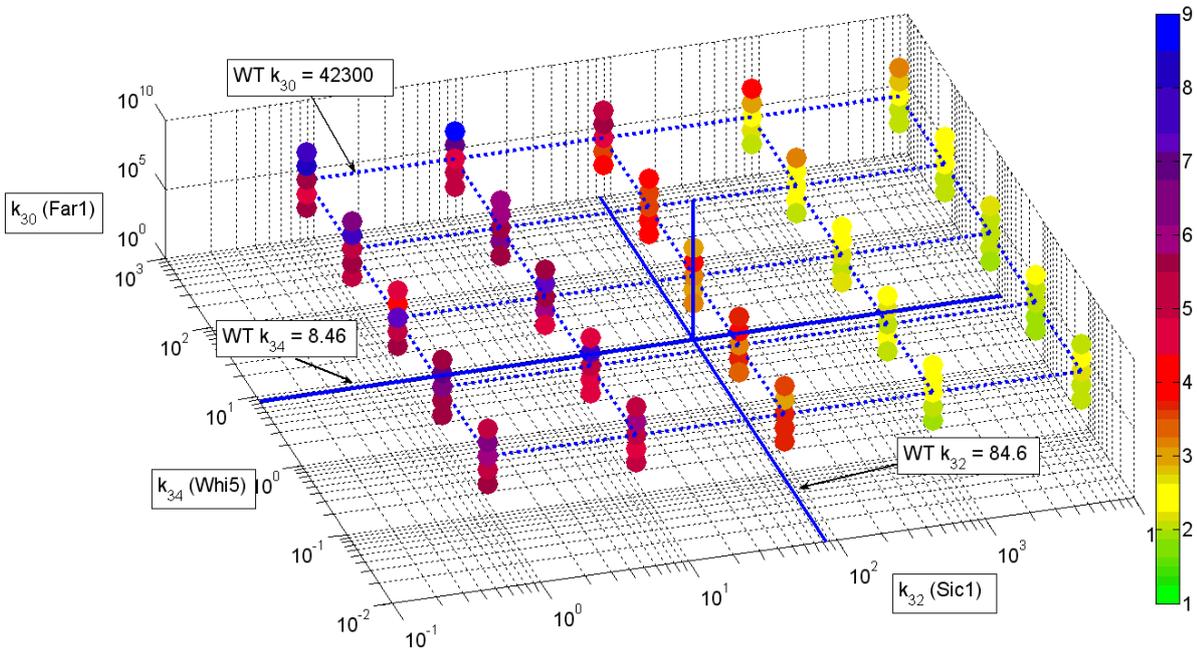


FIG. 4B

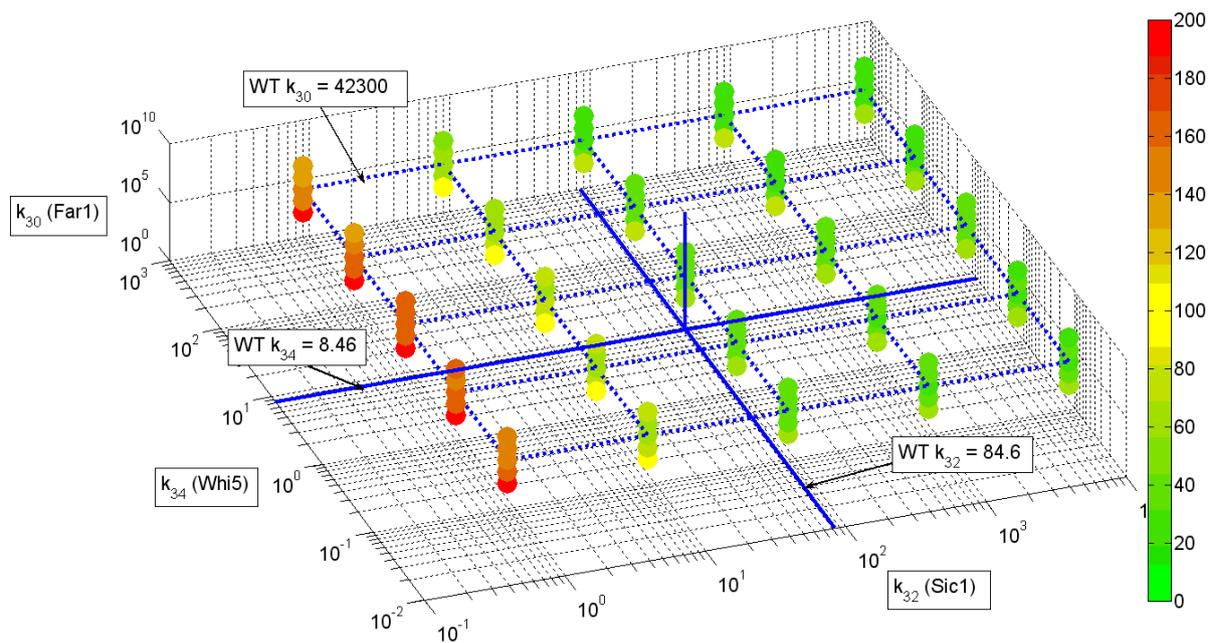


FIG. S-1 A

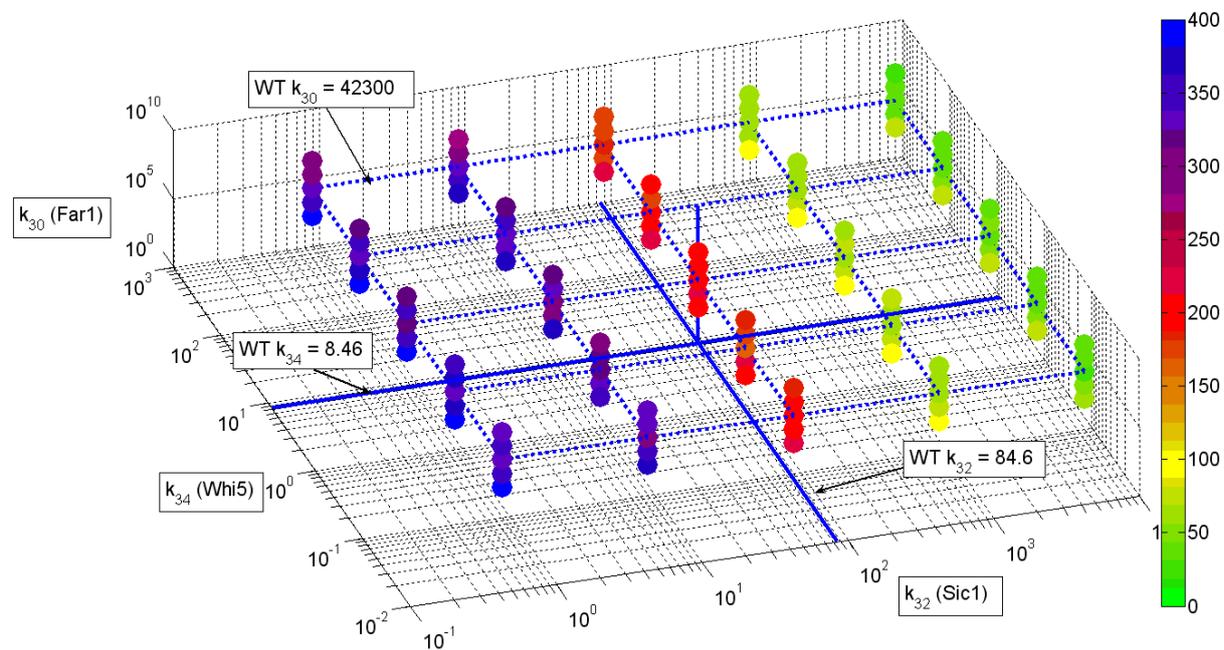


FIG. S-1 B

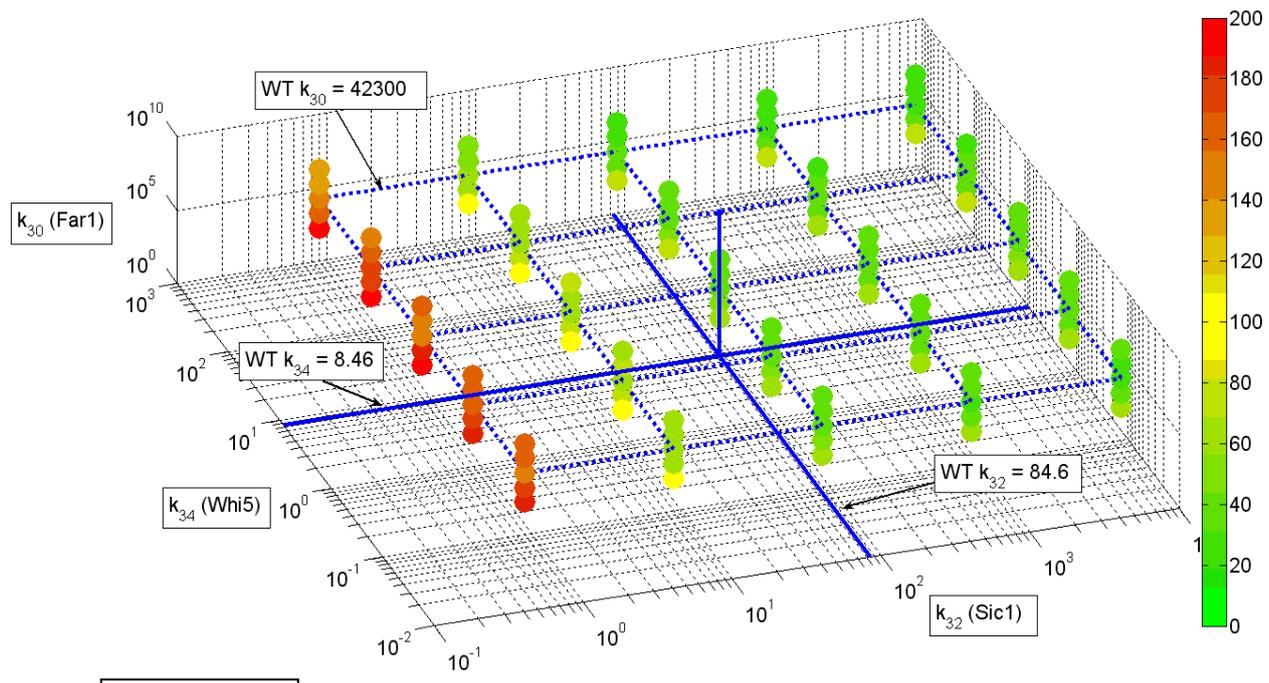


FIG. S-2 A

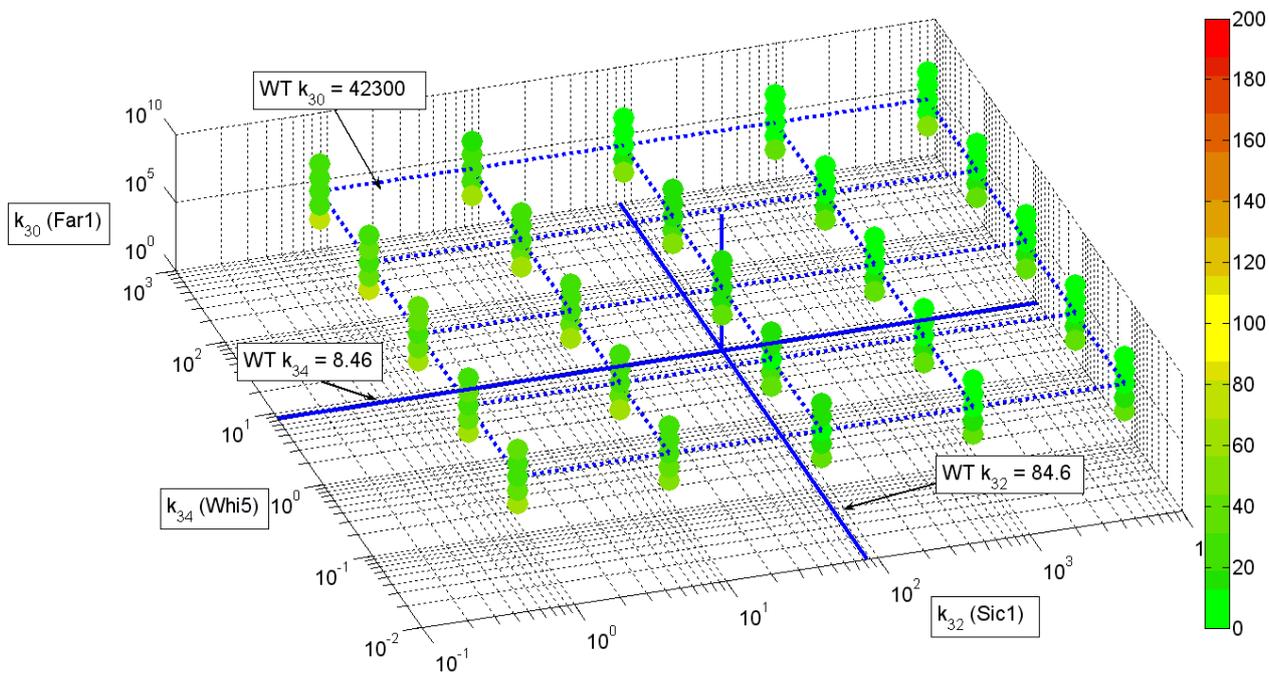


FIG. S-2 B

Figure legends

Fig 1 - Protein-protein interactions network of the proteins playing a major role in the G1 to S transition. A) Extended network (green and yellow) including all the proteins which interact with a core network (yellow). The interaction data were obtained from Reguly *et al.* high-confidence data set. B) The G1/S core network using literature curated data (Barberis *et al.* 2007)

Fig. 2 - Schematic representation of the cell size-dependent G1/S transition control circuit in budding yeast. The emergent behaviour of the G1/S circuit can be described by the presence of a sizer and a timer. The cell sizer mechanism is set by two sequential growth-modulated thresholds. The first threshold is executed when Cln3 exceeds Far1 and is made irreversible by Far1 degradation, whereas the second threshold is due to Sic1 degradation exerted by SCF. The levels of Far1, Whi5 and Sic1 are inherited from the preceding mitotic exit and received by the newborn cell together with a cell mass quantity which is different for daughter and parent cells. The timer is also controlled by the transcription factors SBF and MBF, whose full activity is inhibited by whi5. Active Cdk1-Cln3 phosphorylates Whi5, promoting its dissociation from MBF/SBF and thus activating them. Free MBF/SBF promotes transcription of Cln1,2 and Clb5,6 and commits the cell to a new round of DNA replication and budding.

Fig. 3A - Critical size (P_s) characterization according to different values of the binding affinities of Far1, Whi5 and Sic1 with their respective interactors (k_{30} , k_{34} , k_{32} , respectively), at a *fast* growth rate. Colors associated to the spheres provide a measure of the critical size P_s , according to the colorbar: $P_s = 1$ corresponds to [0 1 0] in an RGB scale (pure Green); $P_s = 2.5$ corresponds to [1 1 0] in an RGB scale (pure Yellow); intermediate values of P_s falling in [1, 2.5] are affine combinations of Green and Yellow; $P_s = 4$ corresponds to [1 0 0] in an RGB scale (pure Red); intermediate values of P_s falling in [2.5, 4] are affine combinations of Yellow and Red; $P_s = 9$ corresponds to [0 0 1] in an RGB scale (pure Blue); intermediate values of P_s falling in [4, 9] are affine combinations of Red and Blue. Values are referred to the average values over 100 *in silico* simulations.

Fig. 3B - Critical size (P_s) characterization according to different values of the binding affinities of Far1, Whi5 and Sic1 with their respective interactors (k_{30} , k_{34} , k_{32} , respectively), at a *slow* growth

rate. The picture is referred to the same colorbar of Fig.3A. Values are referred to the average values over 100 *in silico* simulations.

Fig. 4A - Critical size (P_s) characterization according to different values of the binding affinities of Far1, Whi5 and Sic1 with their respective interactors (k_{30} , k_{34} , k_{32} , respectively), at a fast growth rate, with the Sic1 degradation rate *lowered* of two orders of magnitude. The picture is referred to the same colorbar of Fig.3A. Values are referred to the average values over 100 *in silico* simulations.

Fig. 4B - Critical size (P_s) characterization according to different values of the binding affinities of Far1, Whi5 and Sic1 with their respective interactors (k_{30} , k_{34} , k_{32} , respectively), at a fast growth rate, with the Sic1 degradation rate *increased* of two orders of magnitude. The picture is referred to the same colorbar of Fig.3A. Values are referred to the average values over 100 *in silico* simulations.