Charting the genetic interaction map of a cell
Michael Costanzo¹, Anastasia Baryshnikova¹, Chad L Myers², Brenda Andrews¹ and Charles Boone¹

Genome sequencing projects have revealed a massive catalog of genes and astounding genetic diversity in a variety of organisms. We are now faced with the formidable challenge of assigning functions to thousands of genes, and how to use this information to understand how genes interact and coordinate cell function. Studies indicate that the majority of eukaryotic genes are dispensable, highlighting the extensive buffering of genomes against genetic and environmental perturbations. Such robustness poses a significant challenge to those seeking to understand the wiring diagram of the cell. Genome-scale screens for genetic interactions are an effective means to chart the network that underlies this functional redundancy. A complete atlas of genetic interactions offers the potential to assign functions to most genes identified by whole genome sequencing projects and to delineate a functional wiring diagram of the cell. Perhaps more importantly, mapping genetic networks on a large-scale will shed light on the general principles and rules governing genetic networks and provide valuable information regarding the important but elusive relationship between genotype and phenotype.

Introduction

The relationship between genotype and phenotype is thought to be governed by complex networks of genetic interactions [1]. A systematic approach to map these networks is required to elucidate the genetic interactions underlying disease and develop strategies for therapeutic intervention [2]. The human genome, however, is incredibly complex with an individual’s genome estimated to contain on the order of ~4 million genetic variants and polymorphisms [3]. Given this degree of complexity, determining how critical alleles and polymorphisms combine to manifest a phenotype is a daunting task [4]. A systematic approach involves the mapping of genetic interaction networks in genetically tractable model organisms amenable to large-scale gene deletion or gene knockout technology [5]. Preliminary evidence suggests that at least some genetic interactions and genetic network properties emerging from model organism studies are conserved in higher organisms [6,7,8,9,10,11]. Thus, mapping genetic interaction networks in simple systems should provide an invaluable resource and serve as an important reference to facilitate experimental and comparative analyses of genetic interactions in complex organisms, including humans.

Here, we review progress made to map the genetic interaction network for the budding yeast, *Saccharomyces cerevisiae*. We discuss the importance of a complete genetic interaction network as a wiring diagram for discovering gene and pathway function and as an atlas for predicting analogous networks in more complex systems.

Defining genetic interactions

A genetic interaction refers to an unexpected phenotype not easily explained by combining the effects of individual genetic variants [12]. For example, in the case where cell fitness is the phenotype of choice, a digenic interaction is identified when a double mutant shows a significant deviation in fitness compared to the expected fitness associated with the combination of the two single mutant phenotypes. Thus, quantitative measurement of genetic interactions based upon fitness requires measurements of the single mutant phenotypes, an estimate of the expected double mutant phenotype, and a measurement of the observed double mutant phenotype. For yeast fitness, the expected double mutant phenotype can be modeled as a multiplicative combination of the single mutant phenotypes [13,14–16] (Figure 1).

Negative genetic interactions refer to a more severe fitness defect than expected, with the extreme case being synthetic lethality. Synthetic lethal interactions, which occur when both single mutants are viable but the double mutant is lethal, are of particular interest because they often identify genes that impinge on a common, essential biological function [17,18] (Figure 1). On the other hand, positive interactions refer to double mutants with a less severe fitness defect than expected, and can be further subclassified into a variety of categories [15,16,19,20].
A graphical representation of how genetic interactions are inferred from a quantitative phenotype, such as fitness. (a) Negative genetic interactions. The expected fitness of the resultant AB double mutant based on a multiplicative model is 0.35 (0.7 × 0.5). Negative deviations from the expected fitness are scored as either synthetic sick or synthetic lethal interactions. (b) Symmetric positive interactions. The measured fitness of the AB double mutant (0.5) is greater than the multiplicative expectation (0.25) indicating a positive genetic interaction. This interaction is classified as symmetric because the two single mutants (A and B) and the resultant double mutant (AB) exhibit an equivalent fitness defect (0.5) relative to wild-type (1.0). Symmetric interactions of this kind are enriched among members of the same nonessential protein complex. (c) Asymmetric positive interactions. In this scenario, single mutants and double mutants differ in fitness. Positive deviations from expectation along with single mutant fitness comparisons allow the classification of asymmetric positive interactions into different subcategories, including masking and suppression [13**,19]. Modified from [5].
The symmetric class describes a type of positive interaction whereby the phenotypes associated with the single mutants and resultant double mutant are quantitatively indistinguishable. Conversely, the asymmetric class consists of those interactions in which the strength of the phenotypic effect varies between single and double mutants.

Importantly, different types of positive genetic interaction categories are associated with different biological interpretations. For example, genetic interactions between genes encoding proteins of the same nonessential protein complex tend to be symmetric [13**,19] (Figure 1). While symmetric interactions are indicative of complex membership, asymmetric interactions have the potential to define gene order within specific biochemical pathways [13**,19]. Thus, when measured accurately genetic interactions offer the potential to infer biochemical relationships between gene products and elucidate how different pathways and complexes relate to one another to modulate cellular functions.

**Large-scale mapping of genetic interactions in S. cerevisiae**

*S. cerevisiae* serves as a powerful model system for dissecting the fundamental properties of eukaryotic cells at a molecular level. Indeed it shares many of the basic cell division and growth functions of human cells and numerous genes are conserved from yeast to human. Systematic deletion analysis demonstrated that the majority of the ~6000 budding yeast genes are individually dispensable, with only a relatively small subset (~20%) required for viability [21,22]. These findings highlight the extensive buffering of genomes against genetic perturbations [4]. Genome-scale screens for genetic interactions that impact the fitness of a cell or organism are an effective means to chart the genetic network that underlies this functional redundancy [18,23,24,25**].

The development of synthetic genetic array (SGA) methodology first enabled the systematic mapping of genetic interactions [17]. SGA is an automated method that combines either arrays of nonessential gene deletion mutants [21,22] or conditional alleles of essential genes with robotic manipulations for high-throughput construction of haploid yeast double mutants and the scoring of genetic interactions [17]. Application of SGA created a large-scale genetic interaction map of a cell revealing fundamental properties of genetic networks and illustrating the effectiveness of genetic interactions for organizing genes into specific biological pathways and complexes [18,25**].

In a complementary approach, diploid synthetic lethal analysis by microarray (dSLAM) takes advantage of the unique DNA sequences — molecular barcodes — associated with each mutant strain in the yeast deletion collection to map synthetic lethal interactions by measuring the relative abundance of double mutants in a mixed population [26]. dSLAM has been used to map synthetic genetic interactions between genes involved in DNA integrity and histone modification [24,27]. A third method for genetic interaction discovery, called genetic interaction mapping (GIM), was used to examine interactions between genes involved in mRNA processing [23]. GIM represents a hybrid of SGA and dSLAM technologies because, in a manner reminiscent of SGA, double mutants are generated by mating and sporulation. However, similar to dSLAM, all steps are performed in a pooled format and interactions are identified by comparing barcode microarray hybridization intensities between double mutants and a reference population [23].

**Mapping quantitative genetic interaction networks**

Early genetic interaction studies were based on the binary assessment of cellular fitness (sick/lethal versus no fitness defect) [17,18]. Although synthetic genetic relationships of this kind are informative, quantitative analysis enables identification of more subtle interactions and construction of higher resolution genetic networks encompassing both negative and positive interactions. For example, liquid growth profiling was used to accurately measure genetic interactions between a subset of genes involved in DNA replication and repair [13**]. In addition to identifying positive and negative genetic interactions, positive interactions were also differentiated into five distinct subclasses associated with different biological interpretations [13**]. In another example, fitness was measured from fluorescence-labeled populations of wild-type cells mixed with either single or double mutant yeast strains to map a quantitative genetic interaction network for genes encoding components of the 26S proteasome [20*]. These various assays have also been applied to quantify genetic interactions between duplicated genes [28–31].

Phenotypes other than fitness have also been quantified to measure genetic interactions. These include theoretical analysis of biomass yield [15] and quantitative invasive growth assays [19], used to examine the genetic networks underlying yeast metabolism and filamentous growth, respectively. Gene expression has also been used as a phenotypic readout to map genetic interactions among signal transduction pathway components in other organisms [32,33]. The studies mentioned above highlight the utility of quantitative genetic interaction analysis for functional analysis of pathways and protein complexes in fine detail. However, while they provide high-resolution interaction measurements, these methodologies are not easily amenable to genome-scale studies.

The S-score was designed for the analysis of SGA data derived from relatively small subsets of the yeast deletion
4 Analytical biotechnology

mutant collection [14] and it has been applied toward functionally biased groups of genes to define networks underlying specific biological processes [34–37]. While providing an estimate of the confidence with which genetic interactions can be assigned, the S-score does not reflect either single or double mutant fitness, which are critical measurements for detailed interpretation of genetic interactions [20*].

A recent study combined SGA with a new genome-scale quantitative scoring methodology to examine ~5.4 million gene pairs covering ~30% of the S. cerevisiae genome [25**,38]. This large-scale endeavor measured single and double mutant yeast fitness to uncover ~170,000 genetic interactions (both negative and positive) and provide the first glimpse of a quantitative, genome-scale genetic interaction network for a eukaryotic cell. Consistent with the degree distribution of other biological networks [39], the majority of genes are sparsely connected in the genetic interaction network while a small number have many interactions and serve as network ‘hubs’. While most genetic interactions occur between genes involved in the same biological process [18,25**], network hubs tend to be pleiotropic and interact with many functionally diverse sets of genes [25**]. Importantly, genes annotated to chromatin/transcription, secretion and membrane trafficking showed a significant number of genetic interactions with numerous different processes indicating that genes involved in these functions are important for mediating cross-process connections in the genetic network [25**].

Interpreting negative and positive genetic interactions

Studies over the past several years have examined the relationship between genetic and physical interaction datasets. In general, genetic and protein–protein interactions are largely orthogonal reflecting the ability of negative genetic interactions, such as synthetic lethality, to connect genes lying in different, functionally redundant pathways, whereas physical interactions identify gene products functioning within the same pathway or protein complex [18,40] (Figure 2a,b). However, early analysis of smaller scale genetic interaction networks showed that, unlike negative interactions, which do not overlap physical interactions, positive genetic interactions can connect genes encoding members of the same protein complex or pathway [13*,34,35]. These observations led to a generalized conclusion suggesting that positive genetic interactions connect members of the same protein complex or pathway while negative interactions occur between biological pathways and protein complexes (e.g. see [34,35,41,42]) (Figure 2b).

However, several observations suggest that the relationship between genetic and physical networks is more complex than previously appreciated [25**,38]. First, genes encoding members of the same essential pathway or complex tend to be connected by negative rather than positive interactions [38,43]. This small subset of negative interactions shows significant overlap with protein–protein interactions reflecting the so-called within-pathway synthetic lethal genetic interactions [44] (Figure 2c,d). Second, while previous small-scale studies (e.g. see [34,35,42]) implied that physical interactions correspond better with positive as opposed to negative interactions, systematic comparison of genome-scale datasets revealed the extent of overlap with protein–protein interactions to be similar for both positive and negative genetic interactions [25**]. Thus, although positive interactions tend to connect members of the same protein complex, the vast majority does not overlap with physical interactions, indicating that positive interactions usually connect genes in different pathways and define functional relationships between the pathways (Figure 2d). Indeed, extraction of a positive genetic interaction network that connects complexes from a global dataset reveals a network of loss-of-function suppression in which the fitness defect associated with the loss-of-function of one allele is suppressed by the loss-of-function of another allele [38]. Consistent with these observations, in silico studies in yeast and E. coli revealed that positive interactions often connect functionally distinct metabolic pathways [45].

Global genetic interaction profiles are a rich source of functional information

A genetic interaction profile — the set of genetic interactions for a particular gene — provides a rich phenotypic signature that reflects gene function. In other words, genes belonging to the same pathway or protein complex normally share similar genetic interaction profiles. As a result, grouping genes according to their pattern of genetic interaction, using various clustering algorithms (e.g. see [46]), is a simple and effective way to define pathways and complexes and, therefore, these methods provide a powerful tool for predicting gene function precisely (Figure 2b and d) [18]. The wealth of functional information encoded in genetic interaction profiles suggests that an unbiased, genome-wide survey of genetic interactions should, in principle, assign most genes to a specific pathway and define the functional relationships between pathways, thereby establishing a detailed functional map of the cell.

The ability of genetic interaction profiles to identify coherent functional clusters was exploited to construct a global network that partitions genes with similar interaction patterns together (Figure 3) [25**]. Because interactions are derived from phenotypic measurements, a network based on genetic interaction profile similarity uncovered broad relationships between diverse biological processes illustrating the inherent functional organization of the cell. Genes displaying tightly correlated profiles
A representation of the molecular mechanisms underlying genetic interactions. **(a)** Between pathway genetic interactions describe negative interactions that can arise from disruption of parallel nonessential pathways that converge on a common essential biological process. **(b)** (top panel) Simple relationship between genetic and physical interactions. Negative interactions (red lines) tend to occur between nonessential pathways and complexes while genes belonging to the same nonessential pathway or complex can be connected by positive genetic interactions (green lines). Dashed lines denote physical interactions. (bottom panel) In addition to a network view, genetic interactions can also be visualized as a clustered matrix where genes sharing similar patterns of negative (red) and positive (green) interactions are grouped together. Genetic interaction profiles are functionally informative and grouping genes according to their pattern of genetic interaction identifies pathway and complex membership. **(c)** Within pathway negative genetic interactions may occur when mutations combine to decrease the activity of the same essential pathway or complex. **(d)** Increasingly complex patterns of negative (red lines) and positive (green lines) are revealed as we continue to identify genetic interactions and move toward completion of a genome-wide genetic interaction map. For example, negative interactions connect components of the same essential protein complex.
formed readily discernable clusters corresponding to distinct biological processes and the relative distance between distinct clusters appeared to reflect shared functions (Figure 3a).

The genetic interaction network is structured such that interrogation of the global genetic map enables dissection of broad biological processes into distinct yet interdependent gene cohorts (Figure 3b). For example, in one region of the global network, distinct gene clusters involved in various processes such as DNA replication, recombination and repair, microtubule biogenesis, RNA processing and RNA decay are readily distinguishable (Figure 3b). In another region of the global network, the related processes complex and positive interactions occur between genes belonging to different protein complexes. Circular and diamond shaped nodes represent nonessential and essential genes, respectively. Dashed lines indicate physical interactions. Similar to (b), clustering of genetic interaction profiles reveals functional relationships between genes. Modified from [5].

(Figure 2 Legend Continued) complex and positive interactions occur between genes belonging to different protein complexes. Circular and diamond shaped nodes represent nonessential and essential genes, respectively. Dashed lines indicate physical interactions. Similar to (b), clustering of genetic interaction profiles reveals functional relationships between genes. Modified from [5].

Please cite this article in press as: Costanzo M, et al. Charting the genetic interaction map of a cell, Curr Opin Biotechnol (2010), doi: 10.1016/j.copbio.2010.11.001
of endoplasmic reticulum (ER)/Golgi traffic, endosome/vacuole protein sorting, cell polarity, morphogenesis, cell wall integrity, protein folding, glycosylation, and ER-dependent protein degradation also clustered into well delineated groups (Figure 3c).

At a more detailed level within the genetic network we can view specific gene clusters that are linked by similar patterns of negative (red) and positive (green) genetic interactions (Figure 3d) [25**]. At this level of scrutiny, genes exhibit a highly modular organization and, consistent with theoretical studies [15], modules correspond to discrete biological pathways and/or protein complexes that are connected almost exclusively by a single type of genetic interaction, either negative or positive (Figure 3d). The modular nature of the genetic network enables accurate functional predictions for poorly characterized genes [25**]. For example, ECM30, UBP15, and PAR32 had not been linked to Gap1 sorting previously; however, they all share a similar genetic interaction profile as Gap1 sorting pathway genes, showing negative genetic interactions with genes involved in glutamate biosynthesis and homoserine/chorismate/serine biosynthesis but positive interactions with genes encoding HOPS and CORVET protein complexes (Figure 3d). The positive interactions linking the HOPS/CORVET complexes and the Gap1 sorting pathway provide a clear example of the between pathway connections often observed for positive interactions (Figure 3d). The functional link between ECM30, UBP15, PAR32 and the Gap1 sorting pathway, anticipated by genetic interactions, was validated by mislocalization and reduced activity of the Gap1 permease phenotypes observed in ecm30Δ, ubp15Δ, and par32Δ deletion mutants. These findings suggest that Ecm30, Ubp15, and Par32 are indeed novel members of the Gap1 sorting pathway [25**].

Thus, a genome-wide genetic interaction network provides a multiscale view of functional connectivity within a cell (Figure 3). On a global level, the genetic network reveals inter-dependencies of general cellular processes. At its most detailed level, genetic interactions define membership of specific pathways and protein complexes and identify genetic relationships between different pathways, generating a functional wiring diagram of the cell.

**Genetic network conservation and importance of a reference genetic network**

Large-scale genetic interaction mapping techniques have been developed for other unicellular organisms [10*,11*,47*,48*] and were recently applied to examine the conservation of genetic interactions between two distantly related yeast species, S. cerevisiae and Schizosaccharomyces pombe [10*,11*]. Despite hundreds of millions of years of evolutionary separation, on the order of ~30% of the genetic interactions tested were found to be conserved between the two yeast species [10*,11*]. Moreover, genetic interactions between budding yeast orthologs have been recapitulated in worms [6,9**,51] and mammalian cells [7*]. Indeed, with the development of RNAi-based methods for gene knockdown studies in mammalian cells [5], we anticipate that global mapping of digenic interactions will enable detailed genetic network analysis in various types of mammalian cells. For example, genome-wide RNAi screens have identified genetic interactions involving the RAS oncogene [49*,50*].

Although the initial findings suggest that genetic interactions can be conserved from yeast to higher organisms, the extent to which individual genetic interactions are conserved across evolutionary time remains unclear [8**,51,52]. Despite this uncertainty, discovery of a central and thus highly pleiotropic role for chromatin-related and transcription-related genes in both the S. cerevisiae [25**] and Caenorhabditis elegans genetic networks [8**] provides evidence suggesting that network structure and topology may be conserved across organisms. These hubs in the worm network function as general buffers of phenotypic variation because they are capable of enhancing the phenotypic consequences associated with mutations in numerous different genes [8**]. This finding emphasizes the importance of identifying genes capable of multiple genetic interactions since genetic network hubs may act as general modifiers of genetic diseases in humans [8**]. Given what appears to be a general functional conservation of network hubs [8**,25**, the yeast genetic network may serve as a template to guide experimental and computational analyses as well as predict genetic interaction hubs in complex organisms where genome-wide combinatorial perturbation analysis is more technically challenging.

**Expanding the phenotypic spectrum of genetic interaction networks**

Most large-scale studies have focused on fitness as the primary phenotype to identify genetic interactions (e.g. see [13**,18,23,25**,26]). In theory, all phenotypes are measurable and amenable to genetic interaction analyses. SGA methodology provides an efficient and systematic means for combining mutations and can be readily applied to identify additional genetic interactions that do not result in overt fitness defects. For example, reporter-gene constructs can be incorporated into the SGA methodology to monitor specific transcriptional responses in the ~5000 deletion mutant backgrounds [55,56] and also used as an alternative to fitness for uncovering genetic interactions [57*]. Furthermore, combining SGA technology with different cytological reporters and high content screening (HCS) methodologies identifies mutant combinations that lack obvious growth defects but elicit subtle yet unexpected cell biological phenotypes [58,59]. Integrating SGA technology with
diverse and quantitative phenotypic assays will lead to construction of high-resolution networks that provide comprehensive genome coverage that may integrate temporal and environmental influences to accurately reflect global cellular functions.

**References and recommended reading**

Papers of particular interest, published within the period of review, have been highlighted as:

- of special interest
- of outstanding interest


8. This study recapitulated a yeast genetic interaction in mammalian cells.


10. This study reports the first large-scale genetic interaction network for *C. elegans* and identified chromatin regulators as important genetic modifiers.


12. This study directly assesses the conservation of specific synthetic lethal interactions between *S. cerevisiae* and *C. elegans* orthologs.


16. See annotation to Ref. [10].


19. This study illustrates how quantitative genetic interactions can be used to infer gene order in a biological pathway.


27. This study describes a new method to measure fitness from competitive growth of fluorescence-labeled yeast strains.


33. This study describes the first quantitative and genome-scale genetic interaction network for a eukaryotic cell.


Charting the genetic interaction map of a cell Costanzo et al.


This study and Ref. [48] report the first E. coli genetic interaction network.


This study and Ref. [50] report genetic interactions involving the Ras oncogene in mammalian cancer cell lines.


