UCSF
UC San Francisco Previously Published Works

Title
Genetic interaction analysis of point mutations enables interrogation of gene function at a residue-level resolution

Permalink
https://escholarship.org/uc/item/8dd6440h

Journal
BioEssays, 36(7)

ISSN
0265-9247

Authors
Braberg, Hannes
Moehle, Erica A
Shales, Michael
et al.

Publication Date
2014-07-01

DOI
10.1002/bies.201400044

Peer reviewed

eScholarship.org
Powered by the California Digital Library
University of California

NIH Public Access
Author Manuscript
Bioessays. Author manuscript; available in PMC 2015 July 01.

Published in final edited form as:
Bioessays. 2014 July ; 36(7): 706–713. doi:10.1002/bies.201400044.

Genetic interaction analysis of point mutations enables interrogation of gene function at a residue-level resolution:

Exploring the applications of high-resolution genetic interaction mapping of point mutations

Hannes Braberg[1),2)], **Erica A. Moehle**[3)], **Michael Shales**[1),2)], **Christine Guthrie**[3)], and **Nevan J. Krogan**[1),2),4),*]

[1)]Department of Cellular and Molecular Pharmacology, University of California, San Francisco, CA, USA

[2)]California Institute for Quantitative Biosciences, QB3, San Francisco, CA, USA

[3)]Department of Biochemistry and Biophysics, University of California, San Francisco, CA, USA

[4)]J. David Gladstone Institutes, San Francisco, CA, USA

Abstract

We have achieved a residue-level resolution of genetic interaction mapping – a technique that measures how the function of one gene is affected by the alteration of a second gene – by analyzing point mutations. Here, we describe how to interpret point mutant genetic interactions, and outline key applications for the approach, including interrogation of protein interaction interfaces and active sites, and examination of post-translational modifications. Genetic interaction analysis has proven effective for characterizing cellular processes; however, to date, systematic high-throughput genetic interaction screens have relied on gene deletions or knockdowns, which limits the resolution of gene function analysis and poses problems for multifunctional genes. Our point mutant approach addresses these issues, and further provides a tool for in vivo structure-function analysis that complements traditional biophysical methods. We also discuss the potential for genetic interaction mapping of point mutations in human cells and its application to personalized medicine.

Keywords

E-MAP; genetic interactions; high-throughput; point mutant; RNA polymerase II; structure-function; yeast

Introduction

A central challenge in the post-genomic era has been to functionally annotate the genetic features identified in the genome sequencing efforts. Budding yeast has long been a model

© 2014 WILEY Periodicals, Inc.

*Corresponding author: Nevan J. Krogan, nevan.krogan@ucsf.edu.

NIH-PA Author Manuscript NIH-PA Author Manuscript NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

organism for genetics, and since the systematic identification of its genes [1], a large effort has been made to determine how these genes function in the biology of the organism. The first comprehensive screen to this end used high-throughput reverse genetics to determine the effects of single gene deletions on cell growth on a genome-wide scale [2]. The resulting dataset provided insights into the importance of individual genes, but did not address the interplay between them. Indeed, to map these functional connections and determine the roles that genes play in pathways requires investigation of combinations of gene disruptions. To this end, large-scale efforts were undertaken to map genetic interactions, which describe how the function of a given gene is affected by the presence or absence of a second gene [3-13]. Genetic interactions have proven highly effective for determining gene functions and identifying groups of genes that encode proteins in the same pathway or complex. Negative genetic interactions (synthetic sick/lethal interactions, SSL) arise when two mutations together cause a stronger growth defect than expected based on the growth phenotypes of the individual single mutations (Fig. 1A). These are often observed for genes that encode proteins that act in separate pathways carrying out the same function (Fig. 1B). Conversely, we define positive genetic interactions as occurring between pairs of mutations where the double mutant is healthier than expected, based on the growth defects conferred by the two single mutants individually (Fig. 1A). Positive interactions often arise between factors that act in the same non-essential pathway and/or belong to the same non-essential complex (Fig. 1B), as has been shown in several organisms, including budding yeast [3], fission yeast [14] and mammalian cells [15].

In the early 2000s, two approaches were developed to identify genetic interactions on a large scale in budding yeast: synthetic genetic arrays (SGA) [5, 6] and diploid based synthetic lethality analysis on microarrays (dSLAM) [4, 16]. However, the readouts were limited to negative interactions and thus only surveyed a subset of the genetic interaction spectrum. To address this limitation, an extension of the SGA approach was developed that allowed for quantitative measurement of both negative and positive genetic interactions. This technique, termed epistatic miniarray profile (E-MAP) [10, 17, 18], enabled the quantitative collection of genetic interaction profiles, which describe how a given mutant is affected by a large number of secondary mutations. Since these genetic profiles report on phenotypes in numerous different mutant backgrounds, they provide highly specific readouts that allow precise identification of sets of genes that have similar effects on cell physiology. E-MAPs have proven invaluable for interrogating a large number of cellular processes, including the early secretory pathway, chromosome function, kinase signaling, RNA processing, plasma membrane function, and lipid biology [3, 9-13].

To date, the vast majority of genetic interaction screens have surveyed deletions of non-essential genes or hypomorphic alleles (or “knockdowns”) of essential genes. While effective for functional interrogation on a whole-gene level, many proteins carry out multiple functions and are difficult to assess, as the observed phenotype of a deletion or knockdown is caused by disruption of all the protein’s functions. Moreover, for proteins that carry out a single function, the whole-gene approach fails to provide information about the active domains or regions. Some larger screens have included different alleles of the same gene, and a few such allelic variants have exhibited different genetic profiles [3, 8]. However, these alleles mainly served as substitutes for deletions of essential genes (e.g.

NIH-PA Author Manuscript NIH-PA Author Manuscript NIH-PA Author Manuscript

temperature-sensitive alleles), and were not designed to interrogate the function of specific regions or residues of the encoded proteins [3, 8]. To address the limitations of whole-gene disruptions, we have enhanced the precision of the E-MAP approach by generating a point mutant E-MAP (pE-MAP). It was unclear at the outset whether individual point mutations would generate robust enough genetic profiles to provide functionally meaningful information. Fortunately, our initial pE-MAP increased the resolution of genetic interaction profiling and facilitated the functional interrogation of specific protein domains, interfaces, and even single amino acid residues [19].

In our primary study, we used the pE-MAP approach to functionally dissect RNA polymerase II (RNAPII), a 12-subunit protein machine responsible for all protein-coding gene expression [19]. We generated a series of 53 single point mutations distributed across the 5 subunits that are essential and unique to RNAPII (not shared with RNAPI or RNAPIII) [20], and screened these against a library of 1,200 deletion and hypomorphic allele strains, selected to represent all major biological processes. The resulting pE-MAP comprised approximately 60,000 interactions that provided insights into RNAPII structure and function at a residue-level resolution. This map enabled us to assign functions to domains of RNAPII, and revealed connections between specific regions of the enzyme and other protein complexes. Furthermore, the resolving power of the genetic profiles distinguished between two sets of RNAPII active site mutants, which we subsequently validated as having distinct phenotypes: one set increased transcription rate and exhibited an upstream shift in transcription start site selection, and the other decreased transcription rate and shifted transcription start downstream. Because of the tight coordination between transcription and mRNA processing that has been observed [21, 22], we further analyzed these two mutant categories on splicing-sensitive microarrays, and discovered a negative correlation between splicing efficiency and RNAPII transcription rate. Finally, we used patterns of genetic interaction scores as a tool for identifying novel positivelyand negatively-acting transcription factors [19].

The pE-MAP technique will be of immense value for future structural and functional interrogations of macromolecular machines, and could ultimately be used to map the effects of point mutations on disease states and drug responses inhuman cells [23]. We here describe how to interpret point mutant genetic interactions, and outline what we believe will be key applications for the method.

Interpretation of genetic interactions is complex due to their high information content

Pairs of point mutations can exhibit genetic interactions that differ from the standard expectation based on deletion interactions

Interpretation of genetic interactions is complex by virtue of the measured readout. In contrast to other assays that directly measure a specific quantity – such as presence of a physical interaction between two proteins – genetic interactions are quantified based on basic cellular features, such as growth rate. Herein, lies the power of the genetic interaction: by measuring how combinations of gene alterations affect fundamental in vivo features, we

NIH-PA Author Manuscript NIH-PA Author Manuscript NIH-PA Author Manuscript

acquire knowledge about how the genes function together in a system-wide context. To better interpret these interactions, it is helpful to outline typical relationships that would give rise to positive or negative interactions. For pairs of gene deletions, a common interpretation of a negative interaction is that the removed genes act in parallel pathways that share a common function (Fig. 1B). Conversely, positive interactions typically arise for pairs of genes that act in the same serial pathway, frequently for genes that encode proteins that are members of the same physical complex (Fig. 1B).

The underlying assumption of these models is that the deletion of a gene fully disrupts the pathway or complex in which its gene product operates. However, point mutations are less prone to cause complete disruption of function, and this can give rise to more complex interpretations of certain types of genetic interactions. For instance, consider two subunits of a non-essential complex, both of which are necessary for complex formation. Deletion of either subunit will fully disrupt complex formation, and deletion of both subunits will thus give rise to a positive genetic interaction (epistasis). Introduction of a moderately destabilizing point mutation in either subunit will only partially impede complex function, and a positive interaction will be observed in the case of deletion of the other subunit. However, in contrast, two point mutations in different subunits are likely to exhibit a negative interaction, resulting from a combined severe destabilization of complex formation. An analogous situation may arise for point mutations in pairs of proteins that act in the same pathway (Fig. 1C). In particular, we expect strong negative interactions when either of the point mutations alone weakens a common interaction interface and the double mutant completely disrupts the interface (Fig. 1D). Importantly, the models discussed here are not universal, but merely examples of common relationships that may lead to the specified interaction types.

Genetic interactions are often more informative when analyzed in sets than individually

While a single genetic interaction is highly informative, the complete genetic interaction profiles are often even more revealing [3, 7]. These report on the mutant phenotype in a large number of secondary mutant backgrounds, which results in a twofold advantage over single interactions: first, the large information content provides a much more detailed and comprehensive comparison than a single interaction; and second, the signal-to-noise ratio increases significantly with the multiple data points. Comparison of the complete genetic interaction profiles between pairs of point mutants thus provides a robust measure for functional similarity between them (Fig. 1E) [19]. It has further proven fruitful to define modules of similar genes or mutations and compare the aggregate interactions or profile similarities between different modules (Fig. 1F) [12, 14, 19, 24, 25]. On a whole-gene level, such modules can be defined by functional annotations such as Gene Ontology terms [26], or by co-complex memberships [27-29]. Point mutations in a single protein can instead be grouped by protein domain or sub-domain. We have also found that clustering of mutations based on the similarity of their genetic profiles can be highly effective for classifying functional modules [14]. The grouping of mutants into modules provides a more intuitive representation of these large datasets, and allows for application of statistical tests to assign significance to connections. We fruitfully exploited these benefits to generate a map linking individual RNAPII residues to defined protein complexes based on similarity of genetic

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

profiles. We identified connections to not only the transcriptional apparatus but to distinct complexes, including the kinetochore and the Rpd3C(L) histone deacetylase complex [19].

Finally, interpretation of genetic interaction patterns can be significantly aided by utilizing previously characterized features. For example, a strong positive interaction between a charge-altering point mutation and the deletion of a different protein could suggest a physical interaction between the mutated residue and the protein. Structural context would inform whether the residue is surface-accessible for binding, thereby helping to interpret the interaction.

pE-MAPs allow for interrogation of diverse features of macromolecular assemblies

Screening of mutations in select protein regions facilitates mapping of protein interaction interfaces and active sites

In our initial application of pE-MAP technology, we chose to interrogate a protein machine with known structure and generally understood function. The RNAPII structure has been determined at different states, several binding partners have been identified, and its active site has been mechanistically characterized [30-38]. By leveraging previously characterized structural and functional features, we could readily interpret the genetic interaction patterns of individual mutants. In the future, these interpretations could be instructive to guide both the choice of residues to mutate and the types of mutations to create. For example, residues involved in physical interactions are likely to be exposed on the protein surface, and mutating an interaction interface residue can adversely affect binding affinity strongly [39]. Thus, for a protein with known structure, screening a panel of surface mutations likely to be disruptive for binding (e.g. by preventing a hydrogen bond or a charge-charge interaction) could be a targeted way to identify potential binding partners in vivo (Fig. 2A).

Similarly, knowledge of active site locations can be utilized to tailor mutations that are likely to alter the activity of the protein being examined (Fig. 2B). An example of this is the series of transcription rate-altering point mutants (RNAPII active site mutations) that interacted with genes that encode transcription factors. As deletion of a positive transcription factor slows down transcription, we expected positive genetic interactions with fast active site mutants and negative interactions with slow mutants. By screening for this pattern, we could thereby predict positive transcription factors and found that Sub1 is a positive transcription factor that plays a role in start site selection. Similarly, negative transcription factors could be predicted by searching for the opposite pattern [19]. Some catalytic machines that would be interesting to study in this way are the proteasome, the ribosome, and different DNA polymerases.

Selective design of the mutational space allows for characterization of functional roles of post-translational modifications

Prior knowledge about properties of individual residues can also be utilized to design mutations that alter these in specific and predictable manners. In this respect, the realm of reversible post-translational modifications (PTMs) would be particularly interesting to

NIH-PA Author Manuscript NIH-PA Author Manuscript NIH-PA Author Manuscript

explore. PTMs, such as protein phosphorylation, ubiquitination and acetylation, play important roles in controlling the functions and/or cellular localizations of proteins. Over the last few years, improvements in mass-spectrometry have made it possible to discover thousands of PTM sites in a single study [40]. Recently, a collection of 200,000 PTM sites in 11 eukaryotes was compiled, along with predictions of the functional relevance of these different sites and modifications based on evolutionary conservation [41]. pE-MAP screening of mutations that either prevent or mimic these PTMs should prove instructive for determining the functions that they serve (Fig. 2C). In a related fashion, the Boeke lab has generated a vast library of systematic point mutations of histones H3 and H4 in budding yeast [42]. The histones comprise the core components of nucleosomes, which maintain the structure of compacted DNA and play roles in transcription control and DNA metabolism. Histone function is in large part regulated by PTMs of specific residues and interactions with other proteins [43]. The Boeke library contains alanine substitutions for all residues, as well as specific substitutions of modifiable residues with the aim of mimicking constitutively modified or unmodified states [42]. A pE-MAP of the alanine substitutions should shed light on interactions with other proteins, and screening the substitutions of modifiable residues would help categorize the functional roles of the PTMs. Other proteins whose functions are governed by post-translational modifications should be fruitful targets for pE-MAP screens.

In principle, any macromolecular assembly can be interrogated via pE-MAPs, and the mutational space can be designed to address any functional aspects of interest. Notably, the method is not limited to proteins, but could be applied to RNA machines as well. The ribosome and the spliceosome are made up of both non-coding RNAs and numerous proteins, and point mutations could be introduced in either of these components for interrogation via pE-MAP screening.

pE-MAPs elucidate structure-function relationships in vivo

In our screen of RNAPII, we found a high correlation between the genetic profile similarity of pairs of mutations and the spatial distance between the mutated residues. This relation held true both for residue pairs residing in the same RNAPII subunit, and for those in different subunits [19]. In other words, the functional similarity between residue pairs gleaned from a pE-MAP is higher for residues that are close to each other in space than for those that are far apart. The pE-MAP profiles could thus, in principle, be used to gain structural information about a protein with unknown structure. This would prove particularly valuable for e.g. membrane proteins, which are difficult to characterize using traditional structural biology methods [44] (Fig. 3A). While pE-MAP profiles alone are not sufficient to carry out de novo structure predictions, they should be effective when constrained by data from experiments such as chemical crosslinking. Over the last few years, great progress has been made in modeling structures of macromolecular assemblies using approaches that integrate data types from several different sources. A combination of comparative protein structure models, electron microscopy and proteomics data was used to structurally characterize the 26S proteasome [45], and a model of the nuclear pore complex (NPC) was generated using numerous data types, including stoichiometry data from protein quantification, proximities from subcomplex purifications and overall shape from electron

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

microscopy (EM) [46, 47]. Recently, a software package named The Integrative Modeling Platform (IMP) was specifically developed for modeling macromolecular assemblies by integrating different types of data [48]. Genetic interaction profiling could be particularly valuable for this type of modeling, as the readouts are highly independent from traditional structural biology data types, such as X-ray crystallography, NMR, chemical crosslinking or EM. Importantly, the pE-MAP data are collected in the natural state of a living cell, and can thus be especially effective for eliminating artifacts introduced in biophysical assays, such as crystal contacts [49, 50] (Fig. 3B).

The complementarity of pE-MAP data to that from traditional structural biology techniques could also be harnessed for charting different conformational states of proteins. Protein structure is dynamic in nature, and while the different states are often similar, several machines, such as the heat shock protein chaperones [51], function via major conformational rearrangements of their domains. Point mutant genetic profiles could here be of value to detect residue pairs that are far apart in a previously characterized conformation but are in close proximity in a different, unknown, conformation. In a similar vein, the pE-MAP approach should be suitable for identifying residues that are functionally dependent in a fashion not directly evident from a characterized structure. A particularly fruitful subject to investigate in this fashion is allosteric regulation, where perturbation of a certain protein site can activate or inhibit an active site located in a different region of the protein structure [52] (Fig. 3C). The potential to identify allosteric sites is especially appealing because the targeting of allosteric ligand binding sites on G protein-coupled receptors is an emerging field in drug discovery [53]. Finally, in the event of complete lack of structural information, the genetic interaction profiles can be used to define functional domains of a protein. In this fashion, an illustrative map of a protein could be devised based on the functional characteristics of its residues rather than their spatial locations (Fig. 3D). Such a map will be partially indicative of structure, but more importantly will allow for a representation of the processes in which different domains are involved.

pE-MAP analysis in human cells has potential as a powerful tool for personalized medicine

The vast majority of genetic interaction screens have been carried out in model organisms, such as yeast [3, 8, 10, 14, 25], bacteria [54, 55], worms [56, 57], and flies [58]. However, recently, two methods have been developed in mammalian cells which mirror early yeast genetic interaction mapping by simultaneously reducing the expression of pairs of genes by RNA interference (RNAi) [15, 59]. Genetic interaction mapping in mammalian cells opens up a realm of possibilities for charting alleles that are associated with disease states or that affect drug responses [15, 59]. While RNAi knockdowns will prove effective for surveying the effect of whole-gene depletions in this respect, the vast majority of genetic variation in the human population is on the level of single-nucleotide polymorphisms (SNPs) [60]. Some point mutations can lead to dysfunctional protein products or lowered expression levels, and these could be well approximated by RNAi knockdowns. However, a point mutation in a gene can also alter the function of the encoded protein, and in this scenario, a simple knockdown cannot accurately portray the natural variation of SNPs in the human population.

NIH-PA Author Manuscript NIH-PA Author Manuscript NIH-PA Author Manuscript

With genome editing technology, it has become possible to introduce specific point mutations into investigator-specified locations in the human genome [61, 62], and with the repurposing of CRISPR/Cas9 for use in cell culture [63, 64], it is now possible to create systematic point mutations in human cells in a high-throughput fashion [65, 66]. Incorporation of this methodology into the mammalian genetic interaction approaches should in the near future allow for the generation of human pE-MAPs. This ability will be valuable for assessing how disease states vary in mutational backgrounds designed to represent the variability found in human populations. With increasingly cost-effective genome sequencing, the ability to build databases of alleles enriched in human populations and disease conditions is soaring [60, 67, 68]. Human pE-MAPs can provide instrumental knowledge for meeting the grand challenge of understanding how and why these alleles affect the risk and progression of disease. The heterogeneous genetic makeup of the human population has also been shown to affect how different individuals respond to pharmaceuticals [69]. By adding drugs to mammalian pE-MAP screens, we will be able to chart efficacy and side effects in different genetic backgrounds, providing instrumental information for personalized medicine. Finally, recent advances in high-content microscopy allow the mammalian E-MAP approaches to be extended to simultaneously study several cellular features, using organelle specific stains or fluorescent reporters [15, 70]. The ability to measure how nearly any cellular readout is affected by drugs or mutations in a tailor-made genetic background will prove uniquely powerful for understanding how different populations react to disease and drugs.

Conclusions and outlook

The quest of charting the functions of genes has become a major scientific challenge in the post-genomic era. Genetic interaction mapping has proven highly effective to this end, and has typically been employed on a whole-gene level [7]. However, several aspects of gene function analysis require interrogation at a higher resolution, and the pE-MAP approach for quantitatively measuring genetic interactions between point mutants helps address this issue [19]. We have here outlined what we believe will be key applications of pE-MAP, including its use in concert with biophysical methods to further our understanding of in vivo structure-function relationships. While these applications will often rely on the design of specific mutations to test a hypothesis, the method can also be applied to examine the effects of naturally occurring genetic variants. Accordingly, the extension of pE-MAP into human cells will allow us to study the effect of genetic variation on disease and drug responses [15, 59]. We thus believe that the pE-MAP method will play an important role in understanding basic cellular machinery and could ultimately be invaluable for advances in personalized medicine.

Acknowledgments

The authors wish to thank members of the Krogan and Guthrie labs. This work was supported by grants from QB3 at UCSF and the NIH (R01GM084448, R01GM084279, P50GM081879, and R01GM098101 to N.J.K., and R01GM021119 to C.G.). E.A.M. was supported by an NSF graduate research fellowship. C.G. is an ACS Research Professor of Molecular Genetics. N.J.K. is a Searle Scholar and a Keck Young Investigator.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

References

1. Goffeau A, Barrell BG, Bussey H, Davis RW, et al. Life with 6000 genes. Science. 1996; 274:63–7.
2. Giaever G, Chu AM, Ni L, Connelly C, et al. Functional profiling of the Saccharomyces cerevisiae genome. Nature. 2002; 418:387–91. [PubMed: 12140549]
3. Collins SR, Miller KM, Maas NL, Roguev A, et al. Functional dissection of protein complexes involved in yeast chromosome biology using a genetic interaction map. Nature. 2007; 446:806–10. [PubMed: 17314980]
4. Pan X, Yuan DS, Xiang D, Wang X, et al. A robust toolkit for functional profiling of the yeast genome. Mol Cell. 2004; 16:487–96. [PubMed: 15525520]
5. Tong AH, Evangelista M, Parsons AB, Xu H, et al. Systematic genetic analysis with ordered arrays of yeast deletion mutants. Science. 2001; 294:2364–8. [PubMed: 11743205]
6. Tong AH, Lesage G, Bader GD, Ding H, et al. Global mapping of the yeast genetic interaction network. Science. 2004; 303:808–13. [PubMed: 14764870]
7. Beltrao P, Cagney G, Krogan NJ. Quantitative genetic interactions reveal biological modularity. Cell. 2010; 141:739–45. [PubMed: 20510918]
8. Costanzo M, Baryshnikova A, Bellay J, Kim Y, et al. The genetic landscape of a cell. Science. 2010; 327:425–31. [PubMed: 20093466]
9. Fiedler D, Braberg H, Mehta M, Chechik G, et al. Functional organization of the *S cerevisiae* phosphorylation network. Cell. 2009; 136:952–63. [PubMed: 19269370]
10. Schuldiner M, Collins SR, Thompson NJ, Denic V, et al. Exploration of the function and organization of the yeast early secretory pathway through an epistatic miniarray profile. Cell. 2005; 123:507–19. [PubMed: 16269340]
11. Surma MA, Klose C, Peng D, Shales M, et al. A lipid E-MAP identifies Ubx2 as a critical regulator of lipid saturation and lipid bilayer stress. Mol Cell. 2013; 51:519–30. [PubMed: 23891562]
12. Wilmes GM, Bergkessel M, Bandyopadhyay S, Shales M, et al. A genetic interaction map of RNA-processing factors reveals links between Sem1/Dss1-containing complexes and mRNA export and splicing. Mol Cell. 2008; 32:735–46. [PubMed: 19061648]
13. Aguilar PS, Frohlich F, Rehman M, Shales M, et al. A plasma-membrane E-MAP reveals links of the eisosome with sphingolipid metabolism and endosomal trafficking. Nat Struct Mol Biol. 2010; 17:901–8. [PubMed: 20526336]
14. Ryan CJ, Roguev A, Patrick K, Xu J, et al. Hierarchical modularity and the evolution of genetic interactomes across species. Mol Cell. 2012; 46:691–704. [PubMed: 22681890]
15. Roguev A, Talbot D, Negri GL, Shales M, et al. Quantitative genetic-interaction mapping in mammalian cells. Nat Methods. 2013; 10:432–7. [PubMed: 23407553]
16. Pan X, Ye P, Yuan DS, Wang X, et al. A DNA integrity network in the yeast *Saccharomyces cerevisiae*. Cell. 2006; 124:1069–81. [PubMed: 16487579]
17. Collins SR, Schuldiner M, Krogan NJ, Weissman JS. A strategy for extracting and analyzing large-scale quantitative epistatic interaction data. Genome Biol. 2006; 7:R63. [PubMed: 16859555]
18. Schuldiner M, Collins SR, Weissman JS, Krogan NJ. Quantitative genetic analysis in *Saccharomyces cerevisiae* using epistatic miniarray profiles (E-MAPs) and its application to chromatin functions. Methods. 2006; 40:344–52. [PubMed: 17101447]
19. Braberg H, Jin H, Moehle EA, Chan YA, et al. From structure to systems: high-resolution, quantitative genetic analysis of RNA polymerase II. Cell. 2013; 154:775–88. [PubMed: 23932120]
20. Archambault J, Friesen JD. Genetics of eukaryotic RNA polymerases I II, and III. Microbiol Rev. 1993; 57:703–24. [PubMed: 8246845]
21. Bentley DL. Rules of engagement: co-transcriptional recruitment of pre-mRNA processing factors. Curr Opin Cell Biol. 2005; 17:251–6. [PubMed: 15901493]
22. de la Mata M, Munoz MJ, Allo M, Fededa JP, et al. RNA polymerase II elongation at the crossroads of transcription and alternative splicing. Genet Res Int. 2011; 2011:309865. [PubMed: 22567350]

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

23. Ryan CJ, Cimermancic P, Szpiech ZA, Sali A, et al. High-resolution network biology: connecting sequence with function. Nat Rev Genet. 2013; 14:865–79. [PubMed: 24197012]
24. Bandyopadhyay S, Kelley R, Krogan NJ, Ideker T. Functional maps of protein complexes from quantitative genetic interaction data. PLoS Comput Biol. 2008; 4:e1000065. [PubMed: 18421374]
25. Roguev A, Bandyopadhyay S, Zofall M, Zhang K, et al. Conservation and rewiring of functional modules revealed by an epistasis map in fission yeast. Science. 2008; 322:405–10. [PubMed: 18818364]
26. Ashburner M, Ball CA, Blake JA, Botstein D, et al. Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. Nat Genet. 2000; 25:25–9. [PubMed: 10802651]
27. Collins SR, Kemmeren P, Zhao XC, Greenblatt JF, et al. Toward a comprehensive atlas of the physical interactome of *Saccharomyces cerevisiae*. Mol Cell Proteomics. 2007; 6:439–50. [PubMed: 17200106]
28. Pu S, Wong J, Turner B, Cho E, et al. Up-to-date catalogues of yeast protein complexes. Nucleic Acids Res. 2009; 37:825–31. [PubMed: 19095691]
29. Benschop JJ, Brabers N, van Leenen D, Bakker LV, et al. A consensus of core protein complex compositions for *Saccharomyces cerevisiae*. Mol Cell. 2010; 38:916–28. [PubMed: 20620961]
30. Chen HT, Warfield L, Hahn S. The positions of TFIIF and TFIIE in the RNA polymerase II transcription preinitiation complex. Nat Struct Mol Biol. 2007; 14:696–703. [PubMed: 17632521]
31. Chen ZA, Jawhari A, Fischer L, Buchen C, et al. Architecture of the RNA polymerase II-TFIIF complex revealed by cross-linking and mass spectrometry. EMBO J. 2010; 29:717–26. [PubMed: 20094031]
32. Cramer P, Bushnell DA, Kornberg RD. Structural basis of transcription: RNA polymerase II at 2.8 angstrom resolution. Science. 2001; 292:1863–76. [PubMed: 11313498]
33. Gnatt AL, Cramer P, Fu J, Bushnell DA, et al. Structural basis of transcription: an RNA polymerase II elongation complex at 3.3 A resolution. Science. 2001; 292:1876–82. [PubMed: 11313499]
34. Kaplan CD. Basic mechanisms of RNA polymerase II activity and alteration of gene expression in *Saccharomyces cerevisiae*. Biochim Biophys Acta. 2013; 1829:39–54. [PubMed: 23022618]
35. Wang D, Bushnell DA, Westover KD, Kaplan CD, et al. Structural basis of transcription: role of the trigger loop in substrate specificity and catalysis. Cell. 2006; 127:941–54. [PubMed: 17129781]
36. Liu X, Bushnell DA, Wang D, Calero G, et al. Structure of an RNA polymerase II-TFIIB complex and the transcription initiation mechanism. Science. 2010; 327:206–9. [PubMed: 19965383]
37. Robinson PJ, Bushnell DA, Trnka MJ, Burlingame AL, et al. Structure of the mediator head module bound to the carboxy-terminal domain of RNA polymerase II. Proc Natl Acad Sci USA. 2012; 109:17931–5. [PubMed: 23071300]
38. Wang D, Bushnell DA, Huang X, Westover KD, et al. Structural basis of transcription: backtracked RNA polymerase II at 3.4 angstrom resolution. Science. 2009; 324:1203–6. [PubMed: 19478184]
39. DeLano WL. Unraveling hot spots in binding interfaces: progress and challenges. Curr Opin Struct Biol. 2002; 12:14–20. [PubMed: 11839484]
40. Choudhary C, Mann M. Decoding signalling networks by mass spectrometry-based proteomics. Nat Rev Mol Cell Biol. 2010; 11:427–39. [PubMed: 20461098]
41. Beltrao P, Albanese V, Kenner LR, Swaney DL, et al. Systematic functional prioritization of protein posttranslational modifications. Cell. 2012; 150:413–25. [PubMed: 22817900]
42. Dai J, Hyland EM, Yuan DS, Huang H, et al. Probing nucleosome function: a highly versatile library of synthetic histone H3 and H4 mutants. Cell. 2008; 134:1066–78. [PubMed: 18805098]
43. Kouzarides T. Chromatin modifications and their function. Cell. 2007; 128:693–705. [PubMed: 17320507]
44. Bill RM, Henderson PJ, Iwata S, Kunji ER, et al. Overcoming barriers to membrane protein structure determination. Nat Biotechnol. 2011; 29:335–40. [PubMed: 21478852]

NIH-PA Author Manuscript NIH-PA Author Manuscript NIH-PA Author Manuscript

45. Lasker K, Forster F, Bohn S, Walzthoeni T, et al. Molecular architecture of the 26S proteasome holocomplex determined by an integrative approach. Proc Natl Acad Sci USA. 2012; 109:1380–7. [PubMed: 22307589]

46. Alber F, Dokudovskaya S, Veenhoff LM, Zhang W, et al. Determining the architectures of macromolecular assemblies. Nature. 2007; 450:683–94. [PubMed: 18046405]

47. Alber F, Dokudovskaya S, Veenhoff LM, Zhang W, et al. The molecular architecture of the nuclear pore complex. Nature. 2007; 450:695–701. [PubMed: 18046406]

48. Russel D, Lasker K, Webb B, Velazquez-Muriel J, et al. Putting the pieces together: integrative modeling platform software for structure determination of macromolecular assemblies. PLoS Biol. 2012; 10:e1001244. [PubMed: 22272186]

49. Dasgupta S, Iyer GH, Bryant SH, Lawrence CE, et al. Extent and nature of contacts between protein molecules in crystal lattices and between subunits of protein oligomers. Proteins. 1997; 28:494–514. [PubMed: 9261866]

50. Janin J. Specific versus non-specific contacts in protein crystals. Nat Struct Biol. 1997; 4:973–4. [PubMed: 9406542]

51. Saibil H. Chaperone machines for protein folding unfolding and disaggregation. Nat Rev Mol Cell Biol. 2013; 14:630–42. [PubMed: 24026055]

52. Monod J, Wyman J, Changeux JP. On the nature of allosteric transitions: a plausible model. J Mol Biol. 1965; 12:88–118. [PubMed: 14343300]

53. Wootten D, Christopoulos A, Sexton PM. Emerging paradigms in GPCR allostery: implications for drug discovery. Nat Rev Drug Discov. 2013; 12:630–44. [PubMed: 23903222]

54. Butland G, Babu M, Diaz-Mejia JJ, Bohdana F, et al. eSGA: *E. coli* synthetic genetic array analysis. Nat Methods. 2008; 5:789–95. [PubMed: 18677321]

55. Typas A, Nichols RJ, Siegele DA, Shales M, et al. High-throughput, quantitative analyses of genetic interactions in *E. coli*. Nat Methods. 2008; 5:781–7. [PubMed: 19160513]

56. Byrne AB, Weirauch MT, Wong V, Koeva M, et al. A global analysis of genetic interactions in *Caenorhabditis elegans*. J Biol. 2007; 6:8. [PubMed: 17897480]

57. Lehner B, Crombie C, Tischler J, Fortunato A, et al. Systematic mapping of genetic interactions in *Caenorhabditis elegans* identifies common modifiers of diverse signaling pathways. Nat Genet. 2006; 38:896–903. [PubMed: 16845399]

58. Horn T, Sandmann T, Fischer B, Axelsson E, et al. Mapping of signaling networks through synthetic genetic interaction analysis by RNAi. Nat Methods. 2011; 8:341–6. [PubMed: 21378980]

59. Bassik MC, Kampmann M, Lebbink RJ, Wang S, et al. A systematic mammalian genetic interaction map reveals pathways underlying ricin susceptibility. Cell. 2013; 152:909–22. [PubMed: 23394947]

60. Genomes Project C, Abecasis GR, Auton A, Brooks LD, et al. An integrated map of genetic variation from 1,092 human genomes. Nature. 2012; 491:56–65. [PubMed: 23128226]

61. Carroll D. Genome engineering with zinc-finger nucleases. Genetics. 2011; 188:773–82. [PubMed: 21828278]

62. Urnov FD, Miller JC, Lee YL, Beausejour CM, et al. Highly efficient endogenous human gene correction using designed zinc-finger nucleases. Nature. 2005; 435:646–51. [PubMed: 15806097]

63. Mali P, Yang L, Esvelt KM, Aach J, et al. RNA-guided human genome engineering via Cas9. Science. 2013; 339:823–6. [PubMed: 23287722]

64. Gaj T, Gersbach CA, Barbas CF III. ZFN, TALEN, and CRISPR/ Cas-based methods for genome engineering. Trends Biotechnol. 2013; 31:397–405. [PubMed: 23664777]

65. Wang T, Wei JJ, Sabatini DM, Lander ES. Genetic screens in human cells using the CRISPR-Cas9 system. Science. 2014; 343:80–4. [PubMed: 24336569]

66. Zhou Y, Zhu S, Cai C, Yuan P, et al. High-throughput screening of a CRISPR/Cas9 library for functional genomics in human cells. Nature. 2014 in press. 10.1038/nature13166

67. Forbes SA, Bindal N, Bamford S, Cole C, et al. COSMIC: mining complete cancer genomes in the catalogue of somatic mutations in cancer. Nucleic Acids Res. 2011; 39:D945–50. [PubMed: 20952405]

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

68. Stenson PD, Mort M, Ball EV, Shaw K, et al. The human gene mutation database: building a comprehensive mutation repository for clinical and molecular genetics diagnostic testing and personalized genomic medicine. Hum Genet. 2014; 133:1–9. [PubMed: 24077912]

69. Wang L, McLeod HL, Weinshilboum RM. Genomics and drug response. New Engl J Med. 2011; 364:1144–53. [PubMed: 21428770]

70. Loo LH, Wu LF, Altschuler SJ. Image-based multivariate profiling of drug responses from single cells. Nat Methods. 2007; 4:445–53. [PubMed: 17401369]

Abbreviations

E-MAP	epistatic miniarray profile
pE-MAP	point mutant E-MAP
PTM	post-translational modification
RNAPII	RNA polymerase II

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

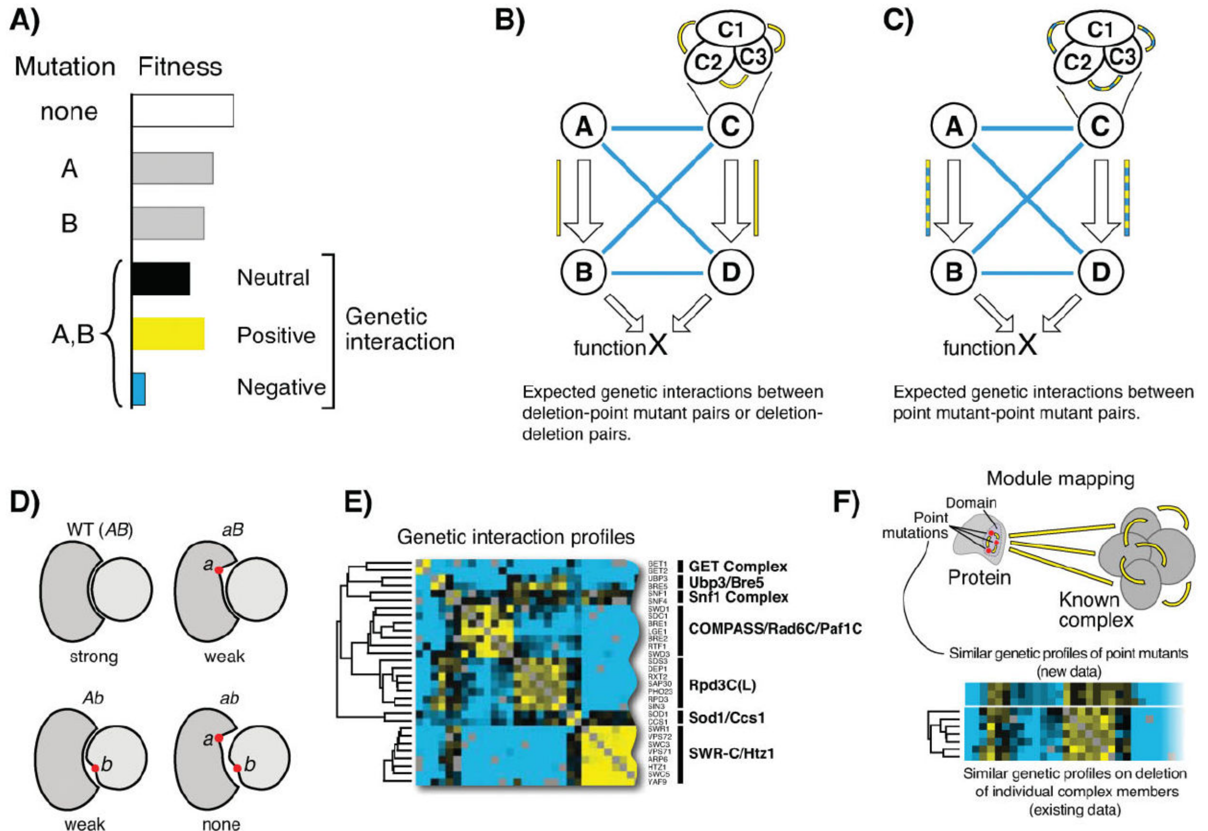

Figure 1.

Interpretation of genetic interactions. **A:** A genetic interaction arises when the fitness of a double mutant deviates from that expected from the two single mutants. Positive genetic interactions are observed for pairs of mutations where the double mutant exhibits better growth than expected from the single mutants, while negative interactions arise when the double mutant is sicker than expected (synthetic sick) or dead (synthetic lethal). **B:** For deletion-deletion or point mutant-deletion pairs, negative interactions commonly arise between genes that encode proteins that act in parallel pathways. Conversely, positive interactions are observed between genes whose products function in the same pathway or belong to the same complex. **C:** Certain relationships that exhibit positive interactions between deletion-deletion or deletion-point mutant pairs can instead give rise to negative interactions for point mutant-point mutant pairs. Specifically, genes that encode proteins that function in the same pathway or complex may exhibit either negative or positive interactions, depending on the severity of the individual point mutations. **D:** A strong negative genetic interaction is observed for a pair of point mutations that individually (aB or Ab) weaken a common interaction interface and together (ab) fully disrupt the interface. **E:** A genetic interaction profile reports on a mutant phenotype in a large number of secondary mutant backgrounds. Hierarchical clustering of genetic profiles categorizes mutants based on functional similarity. A subset of a previously published E-MAP [3] is shown to illustrate how members of the same complex cluster together based on their genetic profiles. The similarity tree on the left describes the hierarchical organization of the profiles, and genes and complexes are listed on the right. **F:** The grouping of related mutations into modules can

strengthen and simplify the interpretations of genetic interactions. In this example, three point mutations in the same protein domain are grouped into one module, while the deletions of different subunits of a known protein complex form another module. Members of the same module are expected to interact genetically with each other and exhibit similar genetic profiles. The protein domain is functionally linked to the complex if several members of each module exhibit strong genetic interactions (top) or high profile similarities (bottom) with the other module.

NIH-PA Author Manuscript NIH-PA Author Manuscript NIH-PA Author Manuscript

NIH-PA Author Manuscript
NIH-PA Author Manuscript
NIH-PA Author Manuscript

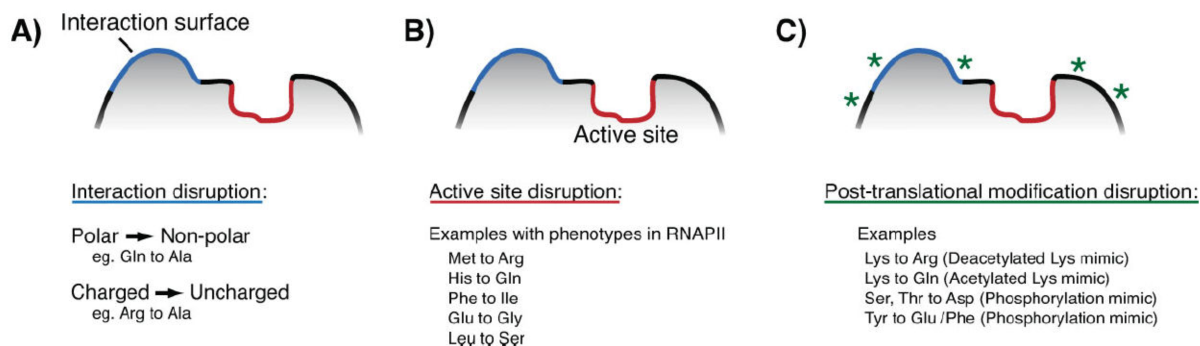

Figure 2.

Design of point mutation space. **A:** To explore protein interaction interfaces, point mutations should be designed to disrupt potential interactions. For example, polar amino acid residues can be mutated to non-polar in order to prevent hydrogen bonding. Similarly, an uncharged residue can be substituted for a charged residue to disrupt a charge-charge interaction. **B:** Active site mutations can be introduced to directly alter the activity of a studied protein. Mutation choices in an active site will vary depending on the specific mechanism to be altered. **C:** The functions of post-translational modification sites can be interrogated by generating mutations that mimic constitutively modified or unmodified states.

NIH-PA Author Manuscript
NIH-PA Author Manuscript
NIH-PA Author Manuscript

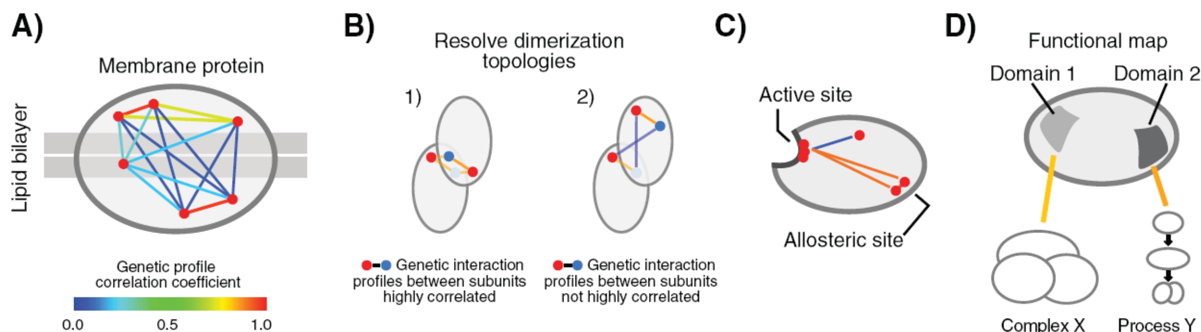

Figure 3.

In vivo interrogation of protein structure-function relationships. **A:** Residues that are close to each other in space exhibit more similar genetic interaction profiles than residues that are far apart. A membrane protein with an unknown or poorly characterized structure is depicted. Genetic profile similarities can be used to generate a map of spatial relationships between the residues, facilitating the construction of a coarse model of the protein. **B:** X-ray crystallography of a dimer structure has given rise to two possible binding interfaces, one of which is biologically relevant and one of which is artificial as a result of crystal packing. Genetic profile similarities can help identify the biologically relevant interface, as mutations of the residues that are close in vivo will exhibit more similar genetic profiles than those that are close due to crystal contacts. **C:** High similarity of genetic profiles between residues in vastly different regions of the same protein may be a strong indicator of an allosteric relationship. **D:** As an alternative approach to a structural model of a protein, a functional map can be assembled based solely on genetic relationships with members of characterized complexes, pathways or processes. Line colors indicate genetic profile similarities.