BioEssays

www.bioessays-journal.com

Genetics' Piece of the PI: Inferring the Origin of Complex Traits and Diseases from Proteome-Wide Protein–Protein Interaction Dynamics

Louis Gauthier, Bram Stynen, Adrian W. R. Serohijos, and Stephen W. Michnick**

How do common and rare genetic polymorphisms contribute to quantitative traits or disease risk and progression? Multiple human traits have been extensively characterized at the genomic level, revealing their complex genetic architecture. However, it is difficult to resolve the mechanisms by which specific variants contribute to a phenotype. Recently, analyses of variant effects on molecular traits have uncovered intermediate mechanisms that link sequence variation to phenotypic changes. Yet, these methods only capture a fraction of genetic contributions to phenotype. Here, in reviewing the field, it is proposed that complex traits can be understood by characterizing the dynamics of biochemical networks within living cells, and that the effects of genetic variation can be captured on these networks by using protein–protein interaction (PPI) methodologies. This synergy between PPI methodologies and the genetics of complex traits opens new avenues to investigate the molecular etiology of human diseases and to facilitate their prevention or treatment.

1. Introduction

Understanding how genomic variation translates to phenotypic variation is a central goal of biology and medicine (**Figure 1**). Over the past decade, genome-wide association studies (GWAS) have been essential to characterize the global landscape of human variation and its corresponding associations to complex traits and diseases. Recent increases in cohort sizes and statistical power have enabled the discovery of more variants with weak effects, thus explaining a larger fraction of the phenotypic variance of complex traits. This has led to the realization that the genetic architecture of most complex traits is highly polygenic, such that the average portion of phenotypic variance explained by any single variant is small.[1] In the context of diseases though, we might expect causal variants to group into a few pathways that drive pathology, even though genome-wide signals are most often scattered across multiple loci. For many of those signals, the connection to disease is not immediately obvious. Furthermore, GWAS have repeatedly shown that nearly 90% of associations between genetic variants and complex traits are found in non-coding regions of the genome,[2] implying a major contribution of regulatory and broadly non-coding events in shaping phenotypes. As a result, explaining the molecular mechanisms by which most variants contribute to phenotype remains a highly active area of research. Unlocking the full potential of the variant-trait association landscape would yield major advances in our understanding of disease etiology and toward the identification of molecular targets for therapeutic intervention. Thus, given the challenges in predicting the molecular effects of variants in general, and regulatory variants in particular, a natural strategy has been to empirically measure the molecular effects of variants in large pools of individuals. Of these strategies, expression quantitative trait loci (eQTL) have been useful in dissecting the molecular contributions of variants to complex traits. However, a large number of variants remain unexplained by these methods.

Recently, the notion of the "proteotype" has been coined to describe the complete proteomic state of a cell at a given time.[3] Concurrently, probing the consequences of genetic variation on protein abundances (pQTL) is proving to be effective in establishing functional links between GWAS-identified risk variants and disease outcomes.[4,5] Furthermore, it has recently been shown[6] that genetic and environmental differences have determinant effects on higher-order proteomic features such as the stoichiometry of protein complexes, and that these are likely to explain important phenotypic variation. Indeed, the ultimate consequences of genetic and environmental perturbations are reflected in the properties of proteins, including their steady-state abundances, post-translational modifications, and subcellular localization. Importantly, the outcome of all of these effects would be seen at the level of protein–protein interaction (PPI) networks. We have developed methods to monitor the dynamics of PPIs across the entire proteome in living cells and to

L. Gauthier, Dr. B. Stynen[+], Prof. A. W. R. Serohijos, Prof. S. W. Michnick
Departement de Biochimie
Université de Montréal
2900 Édouard-Montpetit, Montréal, Quebec H3T 1J4, Canada
E-mail: adrian.serohijos@umontreal.ca;
stephen.michnick@umontreal.ca

L. Gauthier, Dr. B. Stynen, Prof. A. W. R. Serohijos, Prof. S. W. Michnick
Centre Robert-Cedergren en Bioinformatique et Génomique
Université de Montréal
2900 Édouard-Montpetit, Montréal, Quebec H3T 1J4, Canada

The ORCID identification number(s) for the author(s) of this article can be found under https://doi.org/10.1002/bies.201900169

[+]Present address: Qarad BVBA, Pas 257, 2440 Geel, Belgium

DOI: 10.1002/bies.201900169

© 2019 WILEY Periodicals, Inc.

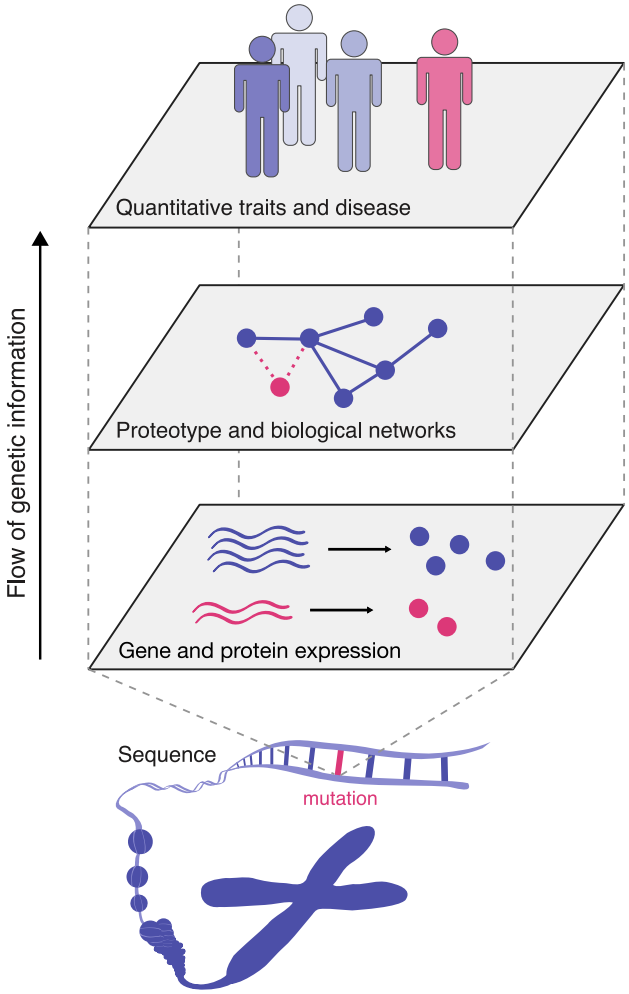

Figure 1. A central goal in medicine is to explain the effects of genetic variation on complex traits and diseases. Mutations associated with particular phenotypes affect intermediate molecular traits, such as gene expression and protein abundance. At a higher level of biological complexity, variation can also modify the biochemical state of the cell by perturbing protein–protein interactions and functional networks.

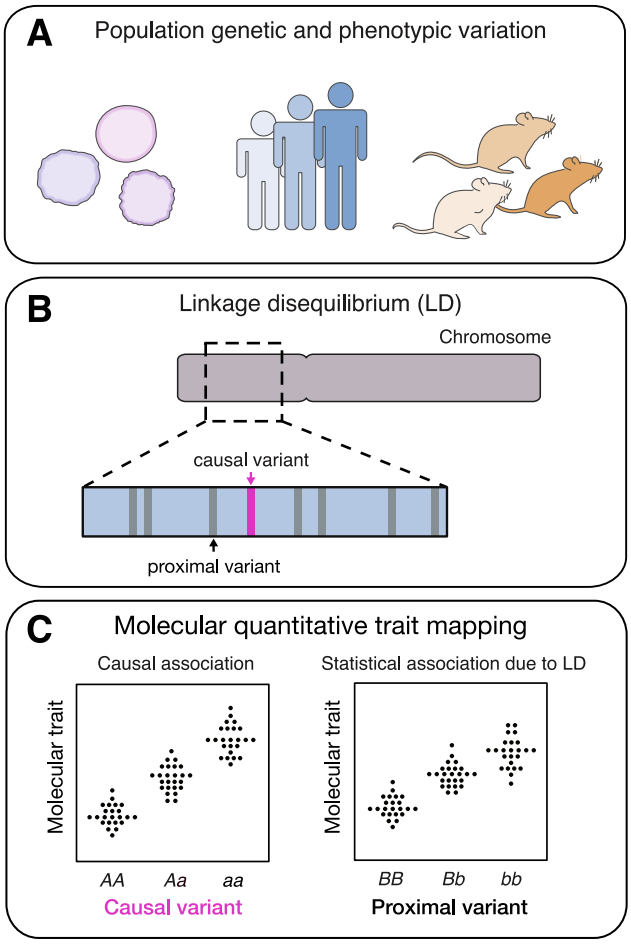

Figure 2. Genetic factors contribute to phenotypic variation observed among individuals in almost every aspect of biological systems. A) Natural genetic variation underlies a variable fraction of the observed phenotypic variation among populations. B) Multiple segregating markers with neutral phenotypic effects can be in linkage disequilibrium with causal variants. C) The non-random segregation of markers due to physical linkage can confound the interpretation of genome-wide associations and quantitative trait mapping.

connect the effects of drug perturbations to molecular mechanisms of action and to disease-associated genes.[7,8] These methods and related approaches should produce what we call protein interaction (pi)QTLs, which will provide the path from genetic variation to perturbations of the proteotype and cellular state underlying complex traits.

2. Complex and Quantitative Traits Are Characterized by Their Genetic Architecture

Individual phenotypes within populations vary in almost every aspect of an organism's states (**Figure 2**A), including morphology, behavior, physiology, development, and susceptibility to diseases.[9] One of the primary goals of genetics is to derive predictive information from individual genotypes, particularly in the context of human diseases. Several rare diseases exhibit monogenic or oligogenic inheritance, where the genetic causes of the phenotype can be attributed to variants at a single or a few loci. Although these diseases often manifest heterogeneity in their clinical expressivity, genetic etiology, and overall prevalence,[10] in many cases, geneticists have been able to create comprehensive maps of the underlying gene(s) and variant(s).[11–13] In turn, these discoveries have informed the pathophysiology of their respective diseases and led to major advances in molecular diagnosis and therapy.[14] However, many of the most common human pathologies, such as type 2 diabetes,[15] rheumatoid arthritis,[16] and schizophrenia[17] do not follow such inheritance patterns. Rather, these complex, or polygenic diseases, as well as many other human quantitative traits are the product of multiple genetic and environmental contributors. In genomics, the genetic architecture of a trait or disease is the landscape of all of its genetic contributions and their characteristics.[1] This comprises the number of genetic variants affecting a trait, their allelic

© 2019 WILEY Periodicals, Inc.

frequencies, the magnitude of their effects, and their interactions with each other (epistasis) and the environment. The characterization of the genetic architecture of a complex disease is central to understanding its pathophysiology and may reveal non-trivial therapeutic opportunities. For example, while variants may show small effect sizes and altogether minor contributions to disease, the genes or pathways they perturb might constitute druggable targets, as exemplified by the genes *PCSK9* in coronary artery disease[18] and *RANKL* in osteoporosis.[19] Furthermore, it has been observed that drugs directed at targets with human genetic support have a higher likelihood of therapeutic success than those without genetic support.[20,21]

To date, hundreds of GWAS have been broadly successful in identifying thousands of statistical associations of loci with diseases and traits,[22] revealing their genetic architectures. For example, diseases such as colorectal cancer exhibit complex genetic architectures, in which hundreds of common loci contribute to overall risk.[23] In many cases, loci have been replicated in separate cohorts,[24] suggesting a bona fide association. In spite of these successes, several hurdles remain on the road from association to function. Due to the large number of genome-wide tag-SNPs and strong linkage disequilibrium (LD), the non-random association between neighboring chromosomal loci (Figure 2B,C), we expect a non-trivial fraction of GWAS hits to be in LD with the causal variant(s). Consequently, these hits might not directly inform our understanding of the biology behind the trait or disease, and empirical validation is often required to narrow down a set of candidate SNPs to a single functional unit. The following section reviews some of the current approaches used for functional validation and interpretation of genomic variation.

3. Understanding the Phenotypic Effects of Variants in Genome-Wide Association Studies from Measurable Molecular Traits

3.1. Functional Interpretation is a Fundamental Step in Genomic Analyses

The functional interpretation of genetic variants is a key step of genetic analysis and is necessary for the development of diagnostics and therapeutics. Due to the massive number of SNP-trait associations identified by GWAS, disentangling the molecular effects of genetic variants constitutes a highly active area of research. To this end, both computational and experimental methods have been pursued.

Computationally, a popular approach to quantitating genomic variant effects is their annotation and prediction. Fundamentally, the prediction of a variant's qualitative and quantitative effects depends on accurate predictions of its molecular effects. For coding variants, where the use of protein structure and sequence conservation can strengthen predictions, some tools have successfully been applied to guide functional interpretation.[25,26] Nevertheless, proper validation of these tools remains challenging due to the scarcity of gold standard experimental data.[27] For non-coding variants, predictions are generally weaker owing to the difficulty in predicting molecular effects. Recent machine learning approaches have shown promising results,[28,29] although their clinical relevance has yet to be demonstrated in practice. Ultimately, a majority of GWAS signals are found in non-coding regions, stressing the need for systematic methods that can identify likely mechanisms and target genes.

3.2. Gene Expression Levels as Measurable Phenotypic Proxies

Experimentally, early GWAS results, together with the demonstration that gene expression levels constitute heritable traits,[30,31] have prompted the integration of intermediate molecular traits in association studies (Figure 2C), leading to the discovery and annotation of eQTLs. Notably, the Genotype-Tissue Expression (GTEx) project has been an important resource in understanding the effects of genetic variation on human gene expression and regulation.[32] For a comprehensive description of the impact of eQTL studies on human genetics, we direct the reader to excellent reviews on the subject.[22,33,34] Importantly, the complexity and the wealth of eQTL effects on different genes in different tissues and cell types pose a challenge in generating specific functional hypotheses from eQTL data alone.[35] In the context of monogenic and oligogenic disease, the pairing of transcriptome sequencing with tissue-specific eQTL annotations has successfully been used as a diagnostic tool in cases where whole-exome sequencing proved inconclusive.[36,37] However, a large fraction of patients remain undiagnosed by either of these methods,[38] suggesting the existence of molecular fingerprints of disease that cannot be captured at the transcript level.

3.3. Proteomic Measurements Can Significantly Improve Variant Interpretation

Despite the extensive characterization of variant effects at the level of gene expression, numerous cellular functions are exerted and mediated at the level of proteins (**Figure 3**B). Importantly, it is generally observed that transcript levels alone are not sufficient to accurately predict protein abundances.[39] In humans, a transcriptome- and proteome-wide screen of 29 healthy tissues reported a median per-gene RNA–protein correlation of ≈0.35.[40] This modest correlation is potentially due to the regulatory and buffering mechanisms at the level of translation, protein quaternary organization, and degradation.[41,42] Indeed, it has been observed that eQTLs often have attenuated or absent effects at the protein level.[43] Relatedly, Cenik and colleagues found that polymorphisms can mediate gene expression differences at the level of translation and identified variants responsible for regulating translation efficiency.[44] Importantly, genes that exhibit inter-individual variability in expression, but not in ribosome occupancy, are less likely to show corresponding differences at the protein level. Their analyses suggest that the functional impact of RNA-level differences needs to be carefully considered to establish causal relationships to phenotype. As a whole, these observations have led to the assumption that a direct examination of the proteome could provide additional biological insights beyond gene expression.[45]

© 2019 WILEY Periodicals, Inc.

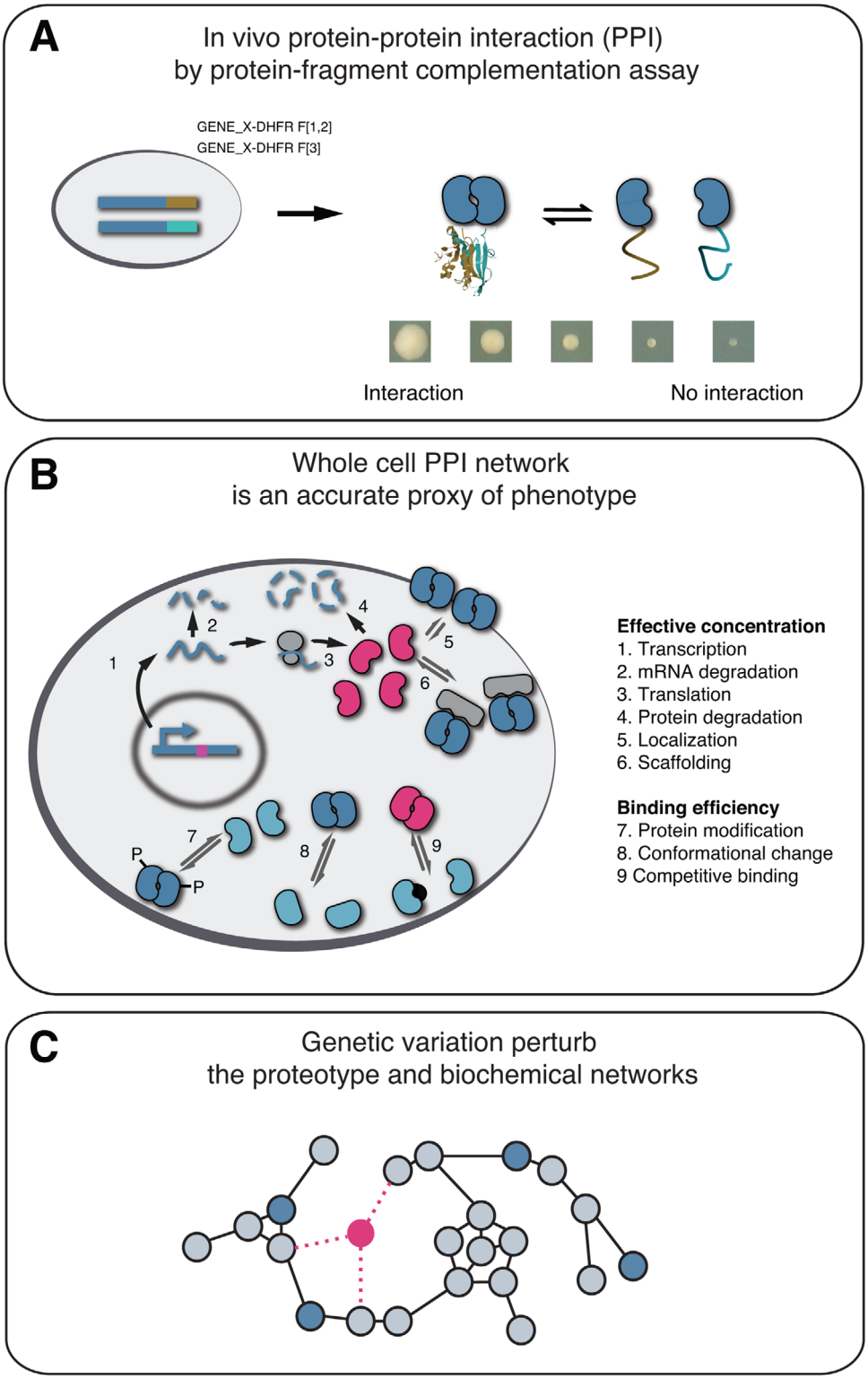

Figure 3. Protein-fragment complementation assays can capture the effects of genetic and environmental perturbations on protein–protein interactions and proteotype. A) The reconstitution of a murine dihydrofolate reductase upon interaction of tagged proteins quantitatively correlates with growth in the presence of methotrexate and is determined by effective concentration and binding efficiency. B) The degree of association of protein-interaction network units results from a combination of biochemical and cellular factors. C) Genetic variation can perturb local and global interaction networks and result in large changes in proteotype and phenotype.

© 2019 WILEY Periodicals, Inc.

Advances in proteomics have since opened up the possibility to investigate the association of genomic variants with protein abundances.[46] Studies using aptamer-,[5,21,47] immunoassay-,[4,48,49] or mass-spectrometry-based[50–52] proteomics have shown how changes in protein abundance can be associated with variants both in cis and in trans. Notably, such proteome-wide screens of pQTL have been conducted in yeast,[53,54] mice,[51,55] maize,[56] human cell lines,[43,57] and primary cell lines.[47] The recent developments in pQTL discovery and mapping suggest that these approaches can illuminate the role of common genetic variants lacking mechanistic links to disease.[4] Interestingly, a recent proteomic characterization of induced pluripotent stem cell (iPSC) lines showed that pQTLs were more likely than eQTLs to tag disease-linked GWAS variants.[58] Together with the existence of many *cis*-pQTLs without co-localizing *cis*-eQTLs,[43,51] these results reflect the higher-order information contained in post-translational mechanisms that determine the biochemical state of the cell, and likely reflect pervasive genetic effects on processes other than transcription, such as transcript stability, translation efficiency and protein degradation. Importantly, *trans*-pQTLs often show pleiotropy, mediating the abundances of multiple proteins.[21,54] Finally, the overlap of different molecular QTLs does not necessarily imply that the same genetic variant underlies both traits, since there may be distinct causal variants for each trait that are in strong linkage disequilibrium.[59]

3.4. Other Molecular Traits That Can Further Advance Our Understanding of Variant Effects

Beyond the investigations of genetic control on gene expression and protein abundance, there has been an emerging interest in tying genetic variation to other molecular phenotypes (molQTLs). Briefly, recent works have identified variants linked to DNA methylation (meQTL),[60] histone modifications and chromatin accessibility,[61] splicing,[62,63] microRNA expression,[64] cytokine,[65] and antibody[66] abundance, and circulating metabolite levels.[67,68] Crucially, the ongoing accumulation of genetic-molecular associations is empowering the discovery of distinct molQTLs sharing common causal variants. For example, Pierce et al. have identified hundreds of eQTL-meQTL pairs showing evidence of a shared biological mechanism, demonstrating instances of human variants with co-regulated effects on expression and methylation.[69] Leveraging the co-localization of molQTLs and the functional relationships that exist between molecular phenotypes can enhance our understanding of the contribution of specific variants to complex traits and diseases. In particular, the Enhancing GTEx (eGTEx) project[70] has been created to complement the expression phenotypes of GTEx with other molQTLs. While this project is bound to yield important contributions in bridging the gaps between genotype and disease, we note that the molecular phenotypes selected therein are limited in scope: protein–protein interactions constitute a salient blind spot. In the next section, we outline the rationale for genome-wide mapping of protein–protein interaction quantitative trait loci, or piQTL.

4. The Protein Interactome Is a Direct Reporter of Phenotype and Argues for Protein-Interaction (pi)QTL Mapping

The idea of a "proteotype" has recently emerged[3,6] to describe a particular instance of a proteome, or the "acute" state of the cell. To fully describe this particular state, one would need to know specific protein abundances, their localization and turnover, the presence and distribution of post-translational modifications and the overall topology/connectivity of the protein–protein interaction network. Importantly, measurement of the proteotype reflects the cellular response to perturbations, whether genetic or environmental. Genetic associations with intermediate proteotypic traits such as abundance and localization are crucial in establishing functional links between GWAS-identified risk variants and disease outcomes.[71] Proteotyping thus constitutes a promising approach to detect the subtle effects of genetic variation on protein and cellular fate. While state-of-the-art methods in proteomics can provide accurate measurements for single proteotypic features, a comprehensive, all-in-one assessment of the proteotype in relation to genomic variation is still lacking. Here, we review recent evidence that highlights how genetic variation may influence proteotypic outcomes.

4.1. Regulatory and Synonymous Mutations Affect Proteome Dynamics

One way by which genetic variation can affect cellular functions is through qualitative alteration of the transcriptome. Indeed, a large number of human polymorphisms occur in splice sites.[72] The complete landscape of single mutations in a human exon[73] has shown that splicing is highly sensitive to mutation: over 90% of positions and 60% of all SNPs altered splicing, regardless of their incidence on amino acid sequence. While the contribution of alternative splicing (AS) events to global proteomic diversity remains unclear, it was shown that AS can increase proteotypic complexity by disrupting the landscape of primary PPIs, in which alternative isoforms often interact in distinct functional modules, in different tissues.[74,75] The partial or complete loss of interaction domains in allele-specific protein isoforms may underlie important perturbations of biological networks that are relevant to phenotype.[76] In addition, Liu and colleagues[77] have shown that AS-driven intron retention events can significantly decrease protein abundances. Importantly, mis-splicing may be due to mutations in *cis*-regulatory elements or in *trans*-acting protein factors.[78] Genomic variants may also affect transcript processing by altering the binding sites of sequence-specific RNA-binding proteins, by disrupting RNA structures required for processing or by introducing RNA structures that prevent RNA processing (reviewed in ref. [79]). Another important class of SNPs are synonymous mutations, historically overlooked in functional genomics due to a long-standing assumption that they are of neutral effects.[80] However, this assumption is now being challenged by reports implicating "silent" exonic variation in splicing regulation.[73,81] Similarly, recent QTL mapping studies in yeast[59] have shown that synonymous mutations can have large effect sizes and tune effective protein abundances.

© 2019 WILEY Periodicals, Inc.

Altogether, qualitative regulatory variation at the transcript level is likely to influence the overall state of the cell through perturbations of protein interaction networks.

4.2. Post-Translational Modifications and Differential Localization Rewire the Protein Interactome

At the proteomic level, post-translational modifications (PTMs) can alter protein localization, binding efficiency, and induce conformational changes[82,83] (Figure 3B). Engelmann et al.[84] used a method capable of reproducibly profiling the phosphoproteome to capture germline genetically driven differential phosphorylation (phosQTLs). They observed that genetic background drives most of the observed phosphopeptide variation, although the generality of these results with respect to cell type, stimulation conditions, and sample size is currently unknown and requires further study. Conversely, Grossmann and colleagues[85] uncovered hundreds of human phosphotyrosine-dependent PPIs, illustrating how interactome rewiring can be mediated by PTMs. Phosphorylation of globular proteins has further been shown to act as a switch, which can disable or enable PPIs.[86] Beyond PTMs, genetic variation can also affect subcellular localization of transcripts and proteins. Although impaired protein localization is known to be implicated in human diseases,[87] and acknowledging that there has been an accumulation of spatial proteomics datasets,[88] to our knowledge, there has yet to be a systematic investigation of the consequences of genetic variation on protein subcellular localization.

4.3. Perturbations of Protein–Protein Interaction Networks Have Major Phenotypic Consequences

Because several protein-coding polymorphisms have been implicated in various human diseases, there exists a sizeable set of annotated mutations that are amenable to the measurement of protein interaction network perturbations.[89] For example, Sahni and colleagues[90] showed that most disease missense mutations do not dramatically impair protein structure or folding, but rather exert their deleterious effects through other mechanisms such as perturbations of PPIs. They assayed 197 missense mutations in 89 proteins implicated in various Mendelian diseases and found that a quarter of them resulted in complete loss of PPIs and another third led to loss of specific interactions, called "edgetic" interactions.[90,91] Strikingly, distinct mutations in the same gene often caused different interaction perturbation profiles. In contrast, only 8% of the non-pathogenic missense variants assayed were found to perturb interactions. In addition, they found that genes carrying disruptive mutations tended to occupy predominantly central positions in the interactome, an observation that was also reported in the context of cancer.[92,93] In an earlier study, the Roth and Vidal groups[94] mapped a network of 14 000 binary human PPIs and showed that disease-associated alleles were tenfold more likely to perturb interactions than non-disease variants. These works highlight the importance of genetic variation in perturbations of biological networks, and the ensuing manifestation of particular phenotypes. Importantly, these studies focused on perturbations by coding variants, which represent only a fraction of all GWAS signals.

Apart from local perturbations of PPIs, multi-subunit protein complexes can also be affected by genetic variation. A recent meta-analysis of 11 proteomics data sets from humans and mice[6] showed that variations in protein complex abundance and stoichiometry among individuals could be explained to a remarkably large extent by the combined influence of genetic sex and diet (≈13.5% for complex abundance and ≈11.6% of stoichiometry). Proteotypes and variation in the integrative information on cellular biochemical networks at the highest order thus appear to be particularly informative in understanding the effects of both genetic and environmental differences and in identifying mechanisms underlying specific traits. Since multi-subunit protein complexes are responsible for the bulk of the functionality of cells, it follows that perturbation of these complexes is likely to affect the overall cellular state and consequently, mutations of the genes that cause disruption of PPIs in protein complexes are predicted to have the most profound phenotypes. Similarly, the decrease in or loss of expression of members of protein complexes has been shown to percolate to other subunits, resulting in changes in their abundances in breast cancer cell proteomes and indicating potential collateral damage from noncoding variants impacting expression.[95] Strikingly, this propagation effect was reliably detected from proteomic, but not from transcriptomic profiles, suggesting post-transcriptional control.

Local network topology also seems to be important in determining the cellular outcome of particular perturbations. For instance, Celaj and colleagues[96] showed that the direction and magnitude of interactions around a given hub protein, those that make a large number of connections with other proteins in the network, are often concerted in response to environmental perturbations. These results emphasize the importance of local network connectivity in mediating the effects of perturbations and can link back to a recently proposed "omnigenic" model of phenotype.[97] To reconcile the vast, broadly distributed number of loci associated with specific traits, Boyle, Li, and Pritchard proposed that cellular networks are sufficiently interconnected to the extent that most genes that are apparently peripheral in the network are able to exert a non-zero influence on phenotype by their close proximity to core genes that more directly affect phenotype. Because these peripheral genes strongly outnumber core genes, their combined contribution to genetic variation is highly significant and, importantly, determined by network connections.

Finally, in a recent computational meta-analysis of human interactome data, Ghadie and Xia estimated an upper bound of 20% of PPIs that are neutral upon disruption by mutation.[98] This suggests that the global interactome is possibly under strong purifying selection and supports the hypothesis that measurement of the proteotype is useful in understanding the phenotypic consequences of genetic variation. Considering the multiple lines of evidence pointing to a major contribution of proteotypic variation to global cellular and organismal phenotype, we propose that PPI dynamics be investigated in the context of genomic variation, leading to protein-interaction quantitative trait loci, or piQTL.

© 2019 WILEY Periodicals, Inc.

5. Methods for In Vivo Protein-Interaction Dynamics Can Be Linked with Quantitative Genetics to Determine Molecular Origins of Complex Phenotypes

If proteotyping, particularly of PPI networks, is so powerful at bridging genotype-phenotype relationships, why is it not being more broadly applied to sort out the potential origins of specific traits or mechanisms of disease? Proteotype analysis is already being envisioned in a limited scope for specific cases, such as the identification of host–pathogen interactions and mechanisms of cancers and neuropsychiatric diseases.[99,100]

Several methods have been developed to successfully bridge the genetics and protein worlds; the most transformative of these being yeast two-hybrid methods.[101] By coupling PPI to the activity of a transcription factor driving expression of different reporter proteins, one ended up with a highly sensitive detector of PPI that did not require any biochemical analysis.[102] These methods are thus simple and scalable; indeed, the first protein interactome studies were performed with two-hybrid methods.[103,104] In addition, systematic two-hybrid-based mapping of PPIs has removed bias seen in interaction studies centered around proteins that have traditionally received a disproportionate degree of research interest.[105] Protein-fragment complementation assays (PCA) also allow for detection of PPIs in living cells but have the added benefit of allowing for either native or constitutive expression of proteins with normal subcellular localization and post-translational modifications. Furthermore, they can be performed in virtually any living cell or whole organism.[106] PCAs also allow for a variety of reporters, from those that confer survival to cells under selection to fluorescent reporters that allow for detection of spatiotemporal dynamics of protein complexes following cell perturbations.[106–109] A particularly simple way to read out survival-selection PCAs is to integrate sequence barcodes associated with specific genes into cells harboring complementary fragments for a pair of proteins.[110,111] Importantly, there is no longer the need to physically array individual clones for specific PPI protein pairs, saving large amounts of associated materials, manipulations, and waiting times to perform a screen.

Recently, we have demonstrated an even simpler strategy to monitor the consequences of genomic variation and environmental conditions. The idea is based on the observation that homomeric interactions of proteins is widespread, observed among up to 50% of the proteome in several organisms.[108,112,113] We reasoned that the quantity of homomeric complexes in a cell should reflect the integrated effects of complex homeostasis from transcription through degradation. We further reasoned that we could capture in the dynamics of PPIs, any active processes such as post-translational modifications that could affect protein stability, localization, conformation, or allosterically affect the oligomerization of the protein. Thus, assays based on measuring homomeric complexes could reduce the multidimensional problem of measuring the genetic or environmental effects on cell fate to a single variable. In this sense, it might be considered similar to evaluating protein "functionality", that is, the extent to which all molecules of a certain protein are performing their function, whether structural, logistic, or enzymatic, conditional on the integrated effect of multiple variables that may be influenced by genetic variation. Key evidence supports the idea that homomerization may be useful as a proxy marker of functionality: the high proportion of proteins that exist in a homomeric state, the generally positive correlation between homomeric state and activity (see Table S3, Supporting Information, in ref. [8]), the obvious variables that influence both homomerization and protein functionality (e.g., abundance), and the existence of possibly underappreciated parameters that may impact both homomerization and functionality (e.g., conformation dynamics, phase separation).

We implemented a simple survival selection PCA based on the enzyme dihydrofolate reductase (DHFR) previously developed in yeast and created an array of 3500 strains for which we could measure homomeric PPIs and examine the effects of a drug on PCA signals, coining the term homomer dynamics *hd*PCA to describe this approach (Figure 3A).[8,108] We were able to demonstrate that *hd*PCA captured the integrated effects of all levels of protein regulation in response to the immunosuppressant and antiproliferative drug rapamycin. We were also able to recapitulate known and novel mechanistic effects of both rapamycin and the diabetes drug metformin. The *hd*PCA method can capture environmental effects on proteome homeostasis in a simple manner. Although this approach directly reads the interaction among homomers, it may also indirectly acquire information about heteromeric interactions. For instance, among the proteins we studied in the *hd*PCA screen, 55% were hub proteins. Furthermore, as noted above, heteromeric PPIs with hub proteins tend to change in the same direction when cells are treated with an environmental perturbation.[96] Whether *hd*PCA can be used to measure effects of genomic variation on PPI and identify piQTLs remains to be seen. Two things make us optimistic. First, the power of yeast genetics should allow us to quickly create libraries to assay *hd*PCA in different genetic backgrounds and environments.[59,114] Second, in our own study, we observed a remarkable concordance between homomeric complexes that change when treated with metformin and homologous human genes that are identified in GWAS studies of type II diabetes and prostate cancer.[8] The results imply that by, in a sense, treating yeast cells for diabetes, the *hd*PCA screen reveals genes implicated in the disease. It is exciting to anticipate how much more we could get from studies like these if we can simplify the heteromeric screening process in order to practically perform screens with the entire proteome. Some exploratory PCA-based heteromeric screens have already shown their utility in examining the dynamics of the interaction network, adding valuable edge-focused information to the node-focused *hd*PCA approach.[111,115,116] With barcoding of strains and high-throughput sequencing readouts, we anticipate that joint measurements of genomic variation and in vivo PPI dynamics may become possible in the near future.

5.1. Current Challenges to En Masse and Population-Wide Proteotyping

High-throughput proteomics on diverse genetic backgrounds remains, for now, costly and logistically complex. Here, we outline

© 2019 WILEY Periodicals, Inc.

some of the technical obstacles that must be overcome in large-scale proteotyping experiments.

Whole-genome sequencing (WGS) is now the gold standard method to capture coding and non-coding genetic variation in a sample group. For a typical, well-powered QTL study, one would require a number of individuals in the range of 10^2–10^3, depending on the specific goals. In the case of human samples, assuming an optimistic amortized cost of $1000 per whole human genome at 30× coverage (including library preparation, reagents, sequencing, and labor), a group of 100 genetic backgrounds or individuals would still represent a minimum expense of $100 000. For smaller genomes, like *S. cerevisiae* or *E. coli*, the cost is orders of magnitude lower, with up to hundreds of pooled genomes sequenced in the same run.

The expansion of WGS capabilities has enabled the creation of large and diverse genomic panels across multiple kingdoms of life, from yeast[59,117,118] to plants[119] and human cell lines.[120] Nevertheless, leveraging the existing sequence data to perform novel proteotypic analyses on these panels presents important logistic challenges. For instance, systematic molecular measurements could require the physical separation and constant tracking of samples throughout the experiments, to avoid costly single cell capture and re-sequencing. These constraints are difficult to reconcile with high-throughput approaches where cells are often pooled and grown together. Beyond sequencing considerations, pooling and demultiplexing individual backgrounds can also be prohibitive. As an example, multiplexing in targeted MS-based proteomics is possible through the use of isobaric mass tags; however, the maximum number of samples that can be pooled in a single experiment is currently less than a dozen.[121,122] Using chromosomal barcoding techniques,[111,123] it is possible to uniquely tag hundreds or thousands of genetically distinct, whole-genome sequenced individuals to augment these already valuable resources for both the genomics and proteomics communities. These barcoded libraries could then be used for genetic investigations of protein–protein interaction networks and proteotypic features in various genetic backgrounds, obviating the need for sample tracking.

6. Conclusion and Future Perspectives

Interactome-based approaches to genetic and environmental perturbations have the potential to improve our understanding of fundamental biological process, including their dysfunction such as in various diseases. Moreover, in the near future, we envision that these approaches could also facilitate the repurposing of drugs by singling out important network modules.[106,124,125] A case in point is a large-scale interactome mapping of 41 breast cancer cell lines, where Lapek and colleagues have identified PPIs exhibiting deviations from the consensus network, and some of these dysregulated interactions can serve as biomarkers to predict the sensitivity of cell lines across a large panel of drugs.[126] Given that network centrality is prevalent among disease-associated proteins, compounds that target specific interactions while leaving others intact are attractive, especially when global disruption of sub-networks is detrimental to the cell.[127] Altogether, we believe that technological developments that could lower the cost and technical barriers to performing interactome measurements en masse will increase our ability to dissect the molecular and cellular mechanisms underlying various human traits and disease states.

Conflict of Interest

The authors declare no conflict of interest.

Keywords

biochemical networks, expression quantitative trait loci, human diseases, protein fragment complementation assays, protein–protein interaction networks, proteotype, quantitative and complex traits

Received: September 16, 2019
Revised: November 15, 2019
Published online: December 18, 2019

[1] N. J. Timpson, C. M. T. Greenwood, N. Soranzo, D. J. Lawson, J. B Richards, *Nat. Rev. Genet.* **2018**, *19*, 110.
[2] S. L. Edwards, J. Beesley, J. D. French, A. M. Dunning, *Am. J. Hum. Genet.* **2013**, *93*, 779.
[3] A. C. Uzozie, R. Aebersold, *J. Proteomics* **2018**, *189*, 1.
[4] C. Yao, G. Chen, C. Song, J. Keefe, M. Mendelson, T. Huan, B. B. Sun, A. Laser, J. C. Maranville, H. Wu, J. E. Ho, P. Courchesne, A. Lyass, M. G. Larson, C. Gieger, J. Graumann, A. D. Johnson, J. Danesh, H. Runz, S.-J. Hwang, C. Liu, A. S. Butterworth, K. Suhre, D. Levy, *Nat. Commun.* **2018**, *9*, 3268.
[5] J. Carayol, C. Chabert, A. Di Cara, C. Armenise, G. Lefebvre, D. Langin, N. Viguerie, S. Metairon, W. H. M. Saris, A. Astrup, P. Descombes, A. Valsesia, J. Hager, *Nat. Commun.* **2017**, *8*, 2084.
[6] N. Romanov, M. Kuhn, R. Aebersold, A. Ori, M. Beck, P. Bork, *Cell* **2019**, *177*, P1308.e10.
[7] M. L. MacDonald, J. Lamerdin, S. Owens, B. H. Keon, G. K. Bilter, Z. Shang, Z. Huang, H. Yu, J. Dias, T. Minami, S. W. Michnick, J. K. Westwick, *Nat. Chem. Biol.* **2006**, *2*, 329.
[8] B. Stynen, D. Abd-Rabbo, J. Kowarzyk, L. Miller-Fleming, S. K. Aulakh, P. Garneau, M. Ralser, S. W. Michnick, *Cell* **2018**, *175*, 1418.e9.
[9] K. A. Frazer, S. S. Murray, N. J. Schork, E. J. Topol, *Nat. Rev. Genet.* **2009**, *10*, 241.
[10] J. F. Robinson, N. Katsanis, *Oligogenic Disease, in Vogel and Motulsky's Human Genetics* (Eds: M. R. Speicher, A. G. Motulsky, S. E. Antonarakis), **2010**, Springer, Berlin p. 243.
[11] G. R. Cutting, *Nat. Rev. Genet.* **2015**, *16*, 45.
[12] N. Blau, *Hum. Mutat.* **2016**, *37*, 508.
[13] A. Verstraeten, M. Alaerts, L. Van Laer, B. Loeys, *Hum. Mutat.* **2016**, *37*, 524.
[14] K. M. Boycott, M. R. Vanstone, D. E. Bulman, A. E. MacKenzie, *Nat. Rev. Genet.* **2013**, *14*, 681.
[15] C. Fuchsberger, J. Flannick, T. M. Teslovich, A. Mahajan, V. Agarwala, K. J. Gaulton, C. Ma, P. Fontanillas, L. Moutsianas, D. J. McCarthy, M. A. Rivas, J. R. B. Perry, X. Sim, T. W. Blackwell, N. R. Robertson, N. W Rayner, P. Cingolani, A. E. Locke, J. F. Tajes, H. M. Highland, J. Dupuis, P. S. Chines, C. M. Lindgren, C. Hartl, A. U. Jackson, H. Chen, J. R. Huyghe, M. van de Bunt, R. D. Pearson, A. Kumar, et al., *Nature* **2016**, *536*, 41.
[16] K. Kim, S.-Y. Bang, H.-S. Lee, S.-C. Bae, *Nat. Rev. Rheumatol.* **2017**, *13*, 13.

 © 2019 WILEY Periodicals, Inc.

[17] L. Rodriguez-Murillo, J. A. Gogos, M. Karayiorgou, *Annu. Rev. Med.* **2012**, *63*, 63.
[18] D. Sullivan, A. G. Olsson, R. Scott, J. B. Kim, A. Xue, V. Gebski, S. M. Wasserman, E. A. Stein, *JAMA, J. Am. Med. Assoc.* **2012**, *308*, 2497.
[19] D. L. Lacey, W. J. Boyle, W. S Simonet, P. J. Kostenuik, W. C. Dougall, J. K. Sullivan, J. S. Martin, R. Dansey, *Nat. Rev. Drug Discovery* **2012**, *11*, 401.
[20] M. R. Nelson, H. Tipney, J. L. Painter, J. Shen, P. Nicoletti, Y. Shen, A. Floratos, P. C. Sham, M. J. Li, J. Wang, L. R. Cardon, J. C. Whittaker, P. Sanseau, *Nat. Genet.* **2015**, *47*, 856.
[21] B. B. Sun, J. C. Maranville, J. E. Peters, D. Stacey, J. R. Staley, J. Blackshaw, S. Burgess, T. Jiang, E. Paige, P. Surendran, C. Oliver-Williams, M. A. Kamat, B. P. Prins, S. K. Wilcox, E. S. Zimmerman, A. Chi, N. Bansal, S. L. Spain, A. M. Wood, N. W. Morrell, J. R. Bradley, N. Janjic, D. J. Roberts, W. H. Ouwehand, J. A. Todd, N. Soranzo, K. Suhre, D. S. Paul, C. S. Fox, R. M. Plenge, et al., *Nature* **2018**, *558*, 73.
[22] M. D. Gallagher, A. S. Chen-Plotkin, *Am. J. Hum. Genet.* **2018**, *102*, 717.
[23] J. R. Huyghe, S. A. Bien, T. A. Harrison, H. M. Kang, S. Chen, S. L. Schmit, D. V. Conti, C. Qu, J. Jeon, C. K. Edlund, P. Greenside, M. Wainberg, F. R. Schumacher, J. D. Smith, D. M. Levine, S. C. Nelson, N. A. Sinnott-Armstrong, D. Albanes, M. H Alonso, K. Anderson, C. Arnau-Collell, V. Arndt, C. Bamia, B. L. Banbury, J. A. Baron, S. I. Berndt, S. Bézieau, D. T Bishop, J. Boehm, H. Boeing, et al., *Nat. Genet.* **2019**, *51*, 76.
[24] U. M. Marigorta, J. A. Rodríguez, G. Gibson, A. Navarro, *Trends Genet.* **2018**, *34*, 504.
[25] Y. Choi, A. P. Chan, *Bioinformatics* **2015**, *31*, 2745.
[26] I. Adzhubei, D. M. Jordan, S. R. Sunyaev, *Curr. Proto. Human Gen.* **2013**, *76*, 7.20.1.
[27] K. S. Raraigh, S. T. Han, E. Davis, T. A. Evans, M. J. Pellicore, A. F. McCague, A. T. Joynt, Z. Lu, M. Atalar, N. Sharma, M. B. Sheridan, P. R. Sosnay, G. R. Cutting, *Am. J. Hum. Genet.* **2018**, *102*, 1062.
[28] J. Zhou, C. L. Theesfeld, K. Yao, K. M. Chen, A. K. Wong, O. G. Troyanskaya, *Nat. Genet.* **2018**, *50*, 1171.
[29] P. Rentzsch, D. Witten, G. M. Cooper, J. Shendure, M. Kircher, *Nucleic Acids Res.* **2019**, *47*, D886.
[30] G. Yvert, R. B. Brem, J. Whittle, J. M. Akey, E. Foss, E. N. Smith, R. Mackelprang, L. Kruglyak, *Nat. Genet.* **2003**, *35*, 57.
[31] R. B. Brem, G. Yvert, R. Clinton, L. Kruglyak, *Science* **2002**, *296*, 752.
[32] G. T. Consortium, *Nature* **2017**, *550*, 204.
[33] F. W. Albert, L. Kruglyak, *Nat. Rev. Genet.* **2015**, *16*, 197.
[34] A. A. Pai, J. K. Pritchard, Y. Gilad, *PLoS Genet.* **2015**, *11*, e1004857.
[35] S. Chun, A. Casparino, N. A. Patsopoulos, D. C. Croteau-Chonka, B. A. Raby, P. L. De Jager, S. R. Sunyaev, C. Cotsapas, *Nat. Genet.* **2017**, *49*, 600.
[36] B. B. Cummings, J. L. Marshall, T. Tukiainen, M. Lek, S. Donkervoort, A. R Foley, V. Bolduc, L. B. Waddell, S. A. Sandaradura, G. L. O'Grady, E. Estrella, H. M. Reddy, F. Zhao, B. Weisburd, K. J. Karczewski, A. H. O'Donnell-Luria, D. Birnbaum, A. Sarkozy, Y. Hu, H. Gonorazky, K. Claeys, H. Joshi, A. Bournazos, E. C. Oates, R. Ghaoui, M. R. Davis, N. G. Laing, A. Topf, P. B. Kang, A. H. Beggs, et al., *Sci. Transl. Med.* **2017**, *9*, eaal5209.
[37] H. D. Gonorazky, S. Naumenko, A. K. Ramani, V. Nelakuditi, P. Mashouri, P. Wang, D. Kao, K. Ohri, S. Viththiyapaskaran, M. A. Tarnopolsky, K. D. Mathews, S. A. Moore, A. N. Osorio, D. Villanova, D. U. Kemaladewi, R. D. Cohn, M. Brudno, J. J. Dowling, *Am. J. Hum. Genet.* **2019**, *104*, 466.
[38] A. L. Wise, T. A. Manolio, G. A. Mensah, J. F. Peterson, D. M. Roden, C. Tamburro, M. S. Williams, E. D. Green, *Lancet* **2019**, *394*, 533.
[39] Y. Liu, A. Beyer, R. Aebersold, *Cell* **2016**, *165*, 535.
[40] D. Wang, B. Eraslan, T. Wieland, B. Hallström, T. Hopf, D. P. Zolg, J. Zecha, A. Asplund, L. H. Li, C. Meng, M. Frejno, T. Schmidt, K. Schnatbaum, M. Wilhelm, F. Ponten, M. Uhlen, J. Gagneur, H. Hahne, B. Kuster, *Mol. Syst. Biol.* **2019**, *15*, e8503.
[41] S. Schafer, E. Adami, M. Heinig, K. E. C Rodrigues, F. Kreuchwig, J. Silhavy, S. van Heesch, D. Simaite, N. Rajewsky, E. Cuppen, M. Pravenec, M. Vingron, S. A. Cook, N. Hubner, *Nat. Commun.* **2015**, *6*, 7200.
[42] C. Vogel, E. M. Marcotte, *Nat. Rev. Genet.* **2012**, *13*, 227.
[43] A. Battle, Z. Khan, S. H. Wang, A. Mitrano, M. J. Ford, J. K. Pritchard, Y. Gilad, *Science* **2015**, *347*, 664.
[44] C. Cenik, E. S. Cenik, G. W. Byeon, F. Grubert, S. I. Candille, D. Spacek, B. Alsallakh, H. Tilgner, C. L. Araya, H. Tang, E. Ricci, M. P. Snyder, *Genome Res.* **2015**, *25*, 1610.
[45] E. J. Foss, D. Radulovic, S. A. Shaffer, D. M. Ruderfer, A. Bedalov, D. R. Goodlett, L. Kruglyak, *Nat. Genet.* **2007**, *39*, 1369.
[46] R. Aebersold, M. Mann, *Nature* **2016**, *537*, 347.
[47] D. Sasayama, K. Hattori, S. Ogawa, Y. Yokota, R. Matsumura, T. Teraishi, H. Hori, M. Ota, S. Yoshida, H. Kunugi, *Hum. Mol. Genet.* **2016**, *26*, 44.
[48] D. V. Zhernakova, T. H. Le, A. Kurilshikov, B. Atanasovska, M. J. Bonder, S. Sanna, A. Claringbould, U. Vösa, P. Deelen, L. Franke, R. A. de Boer, F. Kuipers, M. G. Netea, M. H. Hofker, C. Wijmenga, A. Zhernakova, J. Fu, *Nat. Genet.* **2018**, *50*, 1524.
[49] R. F. Hillary, D. L. McCartney, S. E. Harris, A. J. Stevenson, A. Seeboth, Q. Zhang, D. C. Liewald, K. L. Evans, C. W. Ritchie, E. M. Tucker-Drob, N. R. Wray, A. F. McRae, P. M. Visscher, I. J. Deary, R. E. Marioni, *Nat. Commun.* **2019**, *10*, 3160.
[50] Y. Liu, A. Buil, B. C. Collins, L. C. Gillet, L. C. Blum, L. Y. Cheng, O. Vitek, J. Mouritsen, G. Lachance, T. D. Spector, E. T. Dermitzakis, R. Aebersold, *Mol. Syst. Biol.* **2015**, *11*, 786.
[51] J. M. Chick, S. C. Munger, P. Simecek, E. L. Huttlin, K. Choi, D. M. Gatti, N. Raghupathy, K. L. Svenson, G. A. Churchill, S. P. Gygi, *Nature* **2016**, *534*, 500.
[52] B. L. Parker, A. C. Calkin, M. M. Seldin, M. F. Keating, E. J. Tarling, P. Yang, S. C. Moody, Y. Liu, E. J. Zerenturk, E. J. Needham, M. L. Miller, B. L. Clifford, P. Morand, M. J. Watt, R. C. R. Meex, K.-Y. Peng, R. Lee, K. Jayawardana, C. Pan, N. A. Mellett, J. M. Weir, R. Lazarus, A. J. Lusis, P. J. Meikle, D. E. James, T. Q. de Aguiar Vallim, B. G. Drew, *Nature* **2019**, *567*, 187.
[53] P. Picotti, M. Clément-Ziza, H. Lam, D. S. Campbell, A. Schmidt, E. W. Deutsch, H. Röst, Z. Sun, O. Rinner, L. Reiter, Q. Shen, J. J. Michaelson, A. Frei, S. Alberti, U. Kusebauch, B. Wollscheid, R. L. Moritz, A. Beyer, R. Aebersold, *Nature* **2013**, *494*, 266.
[54] L. Parts, Y.-C. Liu, M. M. Tekkedil, L. M. Steinmetz, A. A. Caudy, A. G. Fraser, C. Boone, B. J. Andrews, A. P. Rosebrock, *Genome Res.* **2014**, *24*, 1363.
[55] L. M. Holdt, A. von Delft, A. Nicolaou, S. Baumann, M. Kostrzewa, J. Thiery, D. Teupser, *Genetics* **2013**, *193*, 601.
[56] L.-G. Jiang, B. Li, S.-X. Liu, H.-W. Wang, C.-P. Li, S.-H. Song, M. Beatty, G. Zastrow-Hayes, X.-H. Yang, F. Qin, Y. He, *Mol. Cell. Proteomics* **2019**, *18*, 263.
[57] R. J. Hause, A. L. Stark, N. N. Antao, L. K. Gorsic, S. H. Chung, C. D. Brown, S. S. Wong, D. F. Gill, J. L. Myers, L. A. To, K. P. White, M. E. Dolan, R. B. Jones, *Am. J. Hum. Genet.* **2014**, *95*, 194.
[58] B. A. Mirauta, D. D. Seaton, D. Bensaddek, A. Brenes, M. J. Bonder, H. Kilpinen, HipSci Consortium, O. Stegle, A. I. Lamond, *bioRxiv* **2018**, 439216.
[59] R. She, D. F. Jarosz, *Cell* **2018**, *172*, 478.e15.
[60] L. Chen, B. Ge, F. P. Casale, L. Vasquez, T. Kwan, D. Garrido-Martín, S. Watt, Y. Yan, K. Kundu, S. Ecker, A. Datta, D. Richardson, F. Burden, D. Mead, A. L. Mann, J. M. Fernandez, S. Rowlston, S. P. Wilder, S. Farrow, X. Shao, J. J. Lambourne, A. Redensek, C. A. Albers, V. Amstislavskiy, S. Ashford, K. Berentsen, L. Bomba, G. Bourque, D. Bujold, S. Busche, et al., *Cell* **2016**, *167*, 1398.e24.

© 2019 WILEY Periodicals, Inc.

[61] F. Grubert, J. B. Zaugg, M. Kasowski, O. Ursu, D. V. Spacek, A. R. Martin, P. Greenside, R. Srivas, D. H. Phanstiel, A. Pekowska, N. Heidari, G. Euskirchen, W. Huber, J. K. Pritchard, C. D. Bustamante, L. M. Steinmetz, A. Kundaje, M. Snyder, *Cell* **2015**, *162*, 1051.
[62] X. Zhang, R. Joehanes, B. H. Chen, T. Huan, S. Ying, P. J. Munson, A. D. Johnson, D. Levy, C. J. O'Donnell, *Nat. Genet.* **2015**, *47*, 345.
[63] A. Takata, N. Matsumoto, T. Kato, *Nat. Commun.* **2017**, *8*, 14519.
[64] T. Huan, J. Rong, C. Liu, X. Zhang, K. Tanriverdi, R. Joehanes, B. H. Chen, J. M. Murabito, C. Yao, P. Courchesne, P. J. Munson, C. J. O'Donnell, N. Cox, A. D. Johnson, M. G. Larson, D. Levy, J. E. Freedman, *Nat. Commun.* **2015**, *6*, 6601.
[65] Y. Li, M. Oosting, P. Deelen, I. Ricaño-Ponce, S. Smeekens, M. Jaeger, V. Matzaraki, M. A. Swertz, R. J. Xavier, L. Franke, C. Wijmenga, L. A. B. Joosten, V. Kumar, M. G. Netea, *Nat. Med.* **2016**, *22*, 952.
[66] S. Jonsson, G. Sveinbjornsson, A. L. de Lapuente Portilla, B. Swaminathan, R. Plomp, G. Dekkers, R. Ajore, M. Ali, A. E. H. Bentlage, E. Elmér, G. I. Eyjolfsson, S. A. Gudjonsson, U. Gullberg, A. Gylfason, B. V. Halldorsson, M. Hansson, H. Holm, Å. Johansson, E. Johnsson, A. Jonasdottir, B. R. Ludviksson, A. Oddsson, I. Olafsson, S. Olafsson, O. Sigurdardottir, A. Sigurdsson, L. Stefansdottir, G. Masson, P. Sulem, M. Wuhrer, et al., *Nat. Genet.* **2017**, *49*, 1182.
[67] W. E. Kraus, D. M. Muoio, R. Stevens, D. Craig, J. R. Bain, E. Grass, C. Haynes, L. Kwee, X. Qin, D. H. Slentz, D. Krupp, M. Muehlbauer, E. R. Hauser, S. G. Gregory, C. B. Newgard, S. H. Shah, *PLoS Genet.* **2015**, *11*, e1005553.
[68] A. W. Gao, M. G. Sterken, J. uit de Bos, J. van Creij, R. Kamble, B. L. Snoek, J. E. Kammenga, R. H. Houtkooper, *Genome Res.* **2018**, *28*, 1296.
[69] B. L. Pierce, L. Tong, M. Argos, K. Demanelis, F. Jasmine, M. Rakibuz-Zaman, G. Sarwar, M. T Islam, H. Shahriar, T. Islam, M. Rahman, M. Yunus, M. G. Kibriya, L. S. Chen, H. Ahsan, *Nat. Commun.* **2018**, *9*, 804.
[70] B. E. Stranger, L. E. Brigham, R. Hasz, M. Hunter, C. Johns, M. Johnson, G. Kopen, W. F. Leinweber, J. T. Lonsdale, A. McDonald, B. Mestichelli, K. Myer, B. Roe, M. Salvatore, S. Shad, J. A. Thomas, G. Walters, M. Washington, J. Wheeler, J. Bridge, B. A. Foster, B. M. Gillard, E. Karasik, R. Kumar, M. Miklos, M. T. Moser, S. D. Jewell, R. G. Montroy, D. C. Rohrer, D. R. Valley, et al., *Nat. Genet.* **2017**, *49*, 1664.
[71] K. Suhre, M. Arnold, A. M. Bhagwat, R. J. Cotton, R. Engelke, J. Raffler, H. Sarwath, G. Thareja, A. Wahl, R. K. DeLisle, L. Gold, M. Pezer, G. Lauc, M. A. El-Din Selim, D. O. Mook-Kanamori, E. K. Al-Dous, Y. A. Mohamoud, J. Malek, K. Strauch, H. Grallert, A. Peters, G. Kastenmüller, C. Gieger, J. Graumann, *Nat. Commun.* **2017**, *8*, 14357.
[72] E. Park, Z. Pan, Z. Zhang, L. Lin, Y. Xing, *Am. J. Hum. Genet.* **2018**, *102*, 11.
[73] P. Julien, B. Miñana, P. Baeza-Centurion, J. Valcárcel, B. Lehner, *Nat. Commun.* **2016**, *7*, 11558.
[74] J. D. Ellis, M. Barrios-Rodiles, R. Colak, M. Irimia, T. Kim, J. A. Calarco, X. Wang, Q. Pan, D. O'Hanlon, P. M. Kim, J. L. Wrana, B. J. Blencowe, *Mol. Cell* **2012**, *46*, 884.
[75] X. Yang, J. Coulombe-Huntington, S. Kang, G. M. Sheynkman, T. Hao, A. Richardson, S. Sun, F. Yang, Y. A. Shen, R. R. Murray, K. Spirohn, B. E. Begg, M. Duran-Frigola, A. MacWilliams, S. J. Pevzner, Q. Zhong, S. A. Trigg, S. Tam, L. Ghamsari, N. Sahni, S. Yi, M. D. Rodriguez, D. Balcha, G. Tan, M. Costanzo, B. Andrews, C. Boone, X. J. Zhou, K. Salehi-Ashtiani, B. Charloteaux, et al., *Cell* **2016**, *164*, 805.
[76] M. Buljan, G. Chalancon, S. Eustermann, G. P. Wagner, M. Fuxreiter, A. Bateman, M. M. Babu, *Mol. Cell* **2012**, *46*, 871.
[77] Y. Liu, M. Gonzàlez-Porta, S. Santos, A. Brazma, J. C. Marioni, R. Aebersold, A. R. Venkitaraman, V. O. Wickramasinghe, *Cell Rep.* **2017**, *20*, 1229.
[78] M. M. Scotti, M. S. Swanson, *Nat. Rev. Genet.* **2016**, *17*, 19.
[79] K. S. Manning, T. A. Cooper, *Nat. Rev. Mol. Cell Biol.* **2017**, *18*, 102.
[80] R. C. Hunt, V. L. Simhadri, M. Iandoli, Z. E. Sauna, C. Kimchi-Sarfaty, *Trends Genet.* **2014**, *30*, 308.
[81] R. Savisaar, L. D. Hurst, *Hum. Genet.* **2017**, *136*, 1059.
[82] A. Bah, J. D. Forman-Kay, *J. Biol. Chem.* **2016**, *291*, 6696.
[83] N. C. Bauer, P. W. Doetsch, A. H. Corbett, *Traffic* **2015**, *16*, 1039.
[84] B. W. Engelmann, C. J. Hsiao, J. D. Blischak, Y. Fourne, Z. Khan, M. Ford, Y. Gilad, *Sci. Rep.* **2018**, *8*, 12106.
[85] A. Grossmann, N. Benlasfer, P. Birth, A. Hegele, F. Wachsmuth, L. Apelt, U. Stelzl, *Mol. Syst. Biol.* **2015**, *11*, 794.
[86] M. J. Betts, O. Wichmann, M. Utz, T. Andre, E. Petsalaki, P. Minguez, L. Parca, F. P. Roth, A.-C. Gavin, P. Bork, R. B. Russell, *PLoS Comput. Biol.* **2017**, *13*, e1005462.
[87] S. Park, J. S. Yang, Y. E. Shin, J. Park, S. K. Jang, S. Kim, *Mol. Syst. Biol.* **2011**, *7*, 494.
[88] L. Gatto, L. M. Breckels, K. S. Lilley, *Curr. Opin. Chem. Biol.* **2019**, *48*, 123.
[89] P. D. Stenson, M. Mort, E. V. Ball, K. Evans, M. Hayden, S. Heywood, M. Hussain, A. D. Phillips, D. N. Cooper, *Hum. Genet.* **2017**, *136*, 665.
[90] N. Sahni, S. Yi, M. Taipale, J. I. Fuxman Bass, J. Coulombe-Huntington, F. Yang, J. Peng, J. Weile, G. I. Karras, Y. Wang, I. A. Kovács, A. Kamburov, I. Krykbaeva, M. H. Lam, G. Tucker, V. Khurana, A. Sharma, Y.-Y. Liu, N. Yachie, Q. Zhong, Y. Shen, A. Palagi, A. San-Miguel, C. Fan, D. Balcha, A. Dricot, D. M. Jordan, J. M. Walsh, A. A. Shah, X. Yang, et al., *Cell* **2015**, *161*, 647.
[91] N. Sahni, S. Yi, Q. Zhong, N. Jailkhani, B. Charloteaux, M. E. Cusick, M. Vidal, *Curr. Opin. Genet. Dev.* **2013**, *23*, 649.
[92] M. Frenkel-Morgenstern, A. Gorohovski, S. Tagore, V. Sekar, M. Vazquez, A. Valencia, *Nucleic Acids Res.* **2017**, *45*, 7094.
[93] S. Yi, S. Lin, Y. Li, W. Zhao, G. B. Mills, N. Sahni, *Nat. Rev. Genet.* **2017**, *18*, 395.
[94] T. Rolland, M. Taşan, B. Charloteaux, S. J. Pevzner, Q. Zhong, N. Sahni, S. Yi, I. Lemmens, C. Fontanillo, R. Mosca, A. Kamburov, S. D. Ghiassian, X. Yang, L. Ghamsari, D. Balcha, B. E. Begg, P. Braun, M. Brehme, M. P. Broly, A.-R. Carvunis, D. Convery-Zupan, R. Corominas, J. Coulombe-Huntington, E. Dann, M. Dreze, A. Dricot, C. Fan, E. Franzosa, F. Gebreab, B. J. Gutierrez, et al., *Cell* **2014**, *159*, 1212.
[95] C. J. Ryan, S. Kennedy, I. Bajrami, D. Matallanas, C. J. Lord, *Cell Syst.* **2017**, *5*, 399. e5.
[96] A. Celaj, U. Schlecht, J. D. Smith, W. Xu, S. Suresh, M. Miranda, A. M. Aparicio, M. Proctor, R. W. Davis, F. P. Roth, R. P. St. Onge, *Mol. Syst. Biol.* **2017**, *13*, 934.
[97] E. A. Boyle, Y. I. Li, J. K. Pritchard, *Cell* **2017**, *169*, 1177.
[98] M. Ghadie, Y. Xia, *Nat. Commun.* **2019**, *10*, 3205.
[99] N. J. Krogan, S. Lippman, D. A. Agard, A. Ashworth, T. Ideker, *Mol. Cell* **2015**, *58*, 690.
[100] A. J Willsey, M. T. Morris, S. Wang, H. R. Willsey, N. Sun, N. Teerikorpi, T. B. Baum, G. Cagney, K. J. Bender, T. A. Desai, D. Srivastava, G. W. Davis, J. Doudna, E. Chang, V. Sohal, D. H. Lowenstein, H. Li, D. Agard, M. J. Keiser, B. Shoichet, M. von Zastrow, L. Mucke, S. Finkbeiner, L. Gan, N. Sestan, M. E. Ward, R. Huttenhain, T. J. Nowakowski, H. J. Bellen, L. M. Frank, et al., *Cell* **2018**, *174*, 505.
[101] S. Fields, O.-K. Song, *Nature* **1989**, *340*, 245.
[102] S. Grimm, *Nat. Rev. Genet.* **2004**, *5*, 179.
[103] P. Uetz, L. Giot, G. Cagney, T. A. Mansfield, R. S. Judson, J. R. Knight, D. Lockshon, V. Narayan, M. Srinivasan, P. Pochart, A. Qureshi-Emili, Y. Li, B. Godwin, D. Conover, T. Kalbfleisch, G. Vijayadamodar, M. Yang, M. Johnston, S. Fields, J. M. Rothberg, *Nature* **2000**, *403*, 623.
[104] T. Ito, T. Chiba, R. Ozawa, M. Yoshida, M. Hattori, Y. Sakaki, *Proc. Natl. Acad. Sci. USA* **2001**, *98*, 4569.
[105] M. Vidal, *Sci. Signaling* **2016**, *9*, eg7.

© 2019 WILEY Periodicals, Inc.

[106] S. W. Michnick, P. H. Ear, E. N. Manderson, I. Remy, E. Stefan, *Nat. Rev. Drug Discovery* **2007**, *6*, 569.
[107] E. Tchekanda, D. Sivanesan, S. W. Michnick, *Nat. Methods* **2014**, *11*, 641.
[108] K. Tarassov, V. Messier, C. R. Landry, S. Radinovic, M. M. S. Molina, I. Shames, Y. Malitskaya, J. Vogel, H. Bussey, S. W. Michnick, *Science* **2008**, *320*, 1465.
[109] P. H. Ear, S. W. Michnick, *Nat. Methods* **2009**, *6*, 813.
[110] U. Schlecht, M. Miranda, S. Suresh, R. W. Davis, R. P. St. Onge, *Proc. Natl. Acad. Sci. USA* **2012**, *109*, 9213.
[111] U. Schlecht, Z. Liu, J. R. Blundell, R. P. St. Onge, S. F. Levy, *Nat. Commun.* **2017**, *8*, 15586.
[112] E. Krissinel, K. Henrick, *J. Mol. Biol.* **2007**, *372*, 774.
[113] P. L. Kastritis, F. J. O'Reilly, T. Bock, Y. Li, M. Z. Rogon, K. Buczak, N. Romanov, M. J. Betts, K. H. Bui, W. J. Hagen, M. L. Hennrich, M. T. Mackmull, J. Rappsilber, R. B. Russell, P. Bork, M. Beck, A. C. Gavin, *Mol. Syst. Biol.* **2017**, *13*, 936.
[114] U. Weill, I. Yofe, E. Sass, B. Stynen, D. Davidi, J. Natarajan, R. Ben-Menachem, Z. Avihou, O. Goldman, N. Harpaz, S. Chuartzman, K. Kniazev, B. Knoblach, J. Laborenz, F. Boos, J. Kowarzyk, S. Ben-Dor, E. Zalckvar, J. M. Herrmann, R. A. Rachubinski, O. Pines, D. Rapaport, S. W. Michnick, E. D. Levy, M. Schuldiner, *Nat. Methods* **2018**, *15*, 617.
[115] S. Rochette, I. Gagnon-Arsenault, G. Diss, C. R. Landry, *J. Proteomics* **2014**, *100*, 25.
[116] L. Freschi, F. Torres-Quiroz, A. K. Dubé, C. R. Landry, *Mol. BioSyst.* **2013**, *9*, 36.
[117] J. Peter, M. De Chiara, A. Friedrich, J.-X. Yue, D. Pflieger, A. Bergström, A. Sigwalt, B. Barre, K. Freel, A. Llored, C. Cruaud, K. Labadie, J.-M. Aury, B. Istace, K. Lebrigand, P. Barbry, S. Engelen, A. Lemainque, P. Wincker, G. Liti, J. Schacherer, *Nature* **2018**, *556*, 339.
[118] P. K. Strope, D. A. Skelly, S. G. Kozmin, G. Mahadevan, E. A. Stone, P. M. Magwene, F. S. Dietrich, J. H. McCusker, *Genome Res.* **2015**, *25*, 762.
[119] C. Alonso-Blanco, J. Andrade, C. Becker, F. Bemm, J. Bergelson, K. M. Borgwardt, J. Cao, E. Chae, T. M. Dezwaan, W. Ding, J. R. Ecker, M. Exposito-Alonso, A. Farlow, J. Fitz, X. Gan, D. G. Grimm, A. M. Hancock, S. R. Henz, S. Holm, M. Horton, M. Jarsulic, R. A. Kerstetter, A. Korte, P. Korte, C. Lanz, C.-R. Lee, D. Meng, T. P. Michael, R. Mott, N. W. Muliyati, et al., *Cell* **2016**, *166*, 481.
[120] S. Fairley, E. Lowy-Gallego, E. Perry, P. Flicek, *Nucleic Acids Res.* **2019**, https://doi.org/10.1093/nar/gkz836.
[121] N. Pappireddi, L. Martin, M. Wühr, *ChemBioChem* **2019**, *20*, 1210.
[122] B. K. Erickson, C. M. Rose, C. R. Braun, A. R. Erickson, J. Knott, G. C. McAlister, M. Wühr, J. A. Paulo, R. A. Everley, S. P. Gygi, *Mol. Cell* **2017**, *65*, 361.
[123] S. F. Levy, J. R. Blundell, S. Venkataram, D. A. Petrov, D. S. Fisher, G. Sherlock, *Nature* **2015**, *519*, 181.
[124] J. Menche, A. Sharma, M. Kitsak, S. D. Ghiassian, M. Vidal, J. Loscalzo, A.-L. Barabasi, *Science* **2015**, *347*, 1257601.
[125] F. Cheng, R. J. Desai, D. E. Handy, R. Wang, S. Schneeweiss, A.-L. Barabási, J. Loscalzo, *Nat. Commun.* **2018**, *9*, 2691.
[126] J. D. Lapek, P. Greninger, R. Morris, A. Amzallag, I. Pruteanu-Malinici, C. H. Benes, W. Haas, *Nat. Biotechnol.* **2017**, *35*, 983.
[127] N. A. Pabon, Q. Zhang, J. A Cruz, D. L. Schipper, C. J. Camacho, R. E. C. Lee, *Nat. Commun.* **2019**, *10*, 860.

© 2019 WILEY Periodicals, Inc.