Quantitative nature of overexpression experiments

Hisao Moriya
Research Core for Interdisciplinary Sciences, Okayama University, Okayama 700-8530, Japan

ABSTRACT Overexpression experiments are sometimes considered as qualitative experiments designed to identify novel proteins and study their function. However, in order to draw conclusions regarding protein overexpression through association analyses using large-scale biological data sets, we need to recognize the quantitative nature of overexpression experiments. Here I discuss the quantitative features of two different types of overexpression experiment: absolute and relative. I also introduce the four primary mechanisms involved in growth defects caused by protein overexpression: resource overload, stoichiometric imbalance, promiscuous interactions, and pathway modulation associated with the degree of overexpression.

INTRODUCTION
Cellular functions are performed through cooperative actions of thousands of proteins. Intracellular levels of these proteins vary substantially (Kulak et al., 2014; Liebermeister et al., 2014; Figure 1), and the level of each protein has to be highly optimized to maximize cellular functionality (Zaslaver et al., 2004; Dekel and Alon, 2005; Wagner, 2005; Li et al., 2014). The abnormal expression of a protein can have a detrimental effect on cellular functions. Overexpression of some proteins is associated with human conditions, such as Down syndrome and cancer (Tang and Amon, 2013). Researchers have used artificial overexpression of proteins to identify novel proteins involved in biological processes of interest and examine the functions of their target proteins (Rine, 1991; Prelich, 2012). Detailed molecular mechanisms found in individual overexpression experiments are summarized by Prelich (2012). Here I focus on the results of large-scale overexpression experiments using the budding yeast Saccharomyces cerevisiae. I emphasize the often-overlooked quantitative nature of overexpression experiments, as shown by considering normal protein levels. I also discuss some primary mechanistic consequences of protein overexpression as demonstrated in large-scale experiments.

OVEREXPRESSION EXPERIMENTS ARE INHERENTLY QUANTITATIVE
Many large-scale overexpression experiments have been performed using transcriptional induction of strong promoters in yeast to identify the proteins causing growth defects when overexpressed (hereafter called dosage-sensitive proteins; Liu et al., 1992; Espinet et al., 1995; Akada et al., 1997; Stevenson et al., 2001; Boyer et al., 2004; Gelperin et al., 2005; Sopko et al., 2006; Niu et al., 2008; Yoshikawa et al., 2011; Douglas et al., 2012). In most cases, overexpression experiments are considered as qualitative screening methods to identify novel proteins involved in cell morphology and cellular processes, such as the cell cycle (Stevenson et al., 2001; Sopko et al., 2006; Niu et al., 2008). Recently the characteristics of proteins that are strongly associated with dosage-sensitive proteins have been surveyed in an attempt to uncover primary mechanistic consequences of protein overexpression (Gelperin et al., 2005; Sopko et al., 2006; Vavouri et al., 2009; Ma et al., 2010; Yoshikawa et al., 2011; Makanae et al., 2013; Tomala and Korona, 2013).

In gene deletion experiments, the expression of a target protein is stopped by removing the protein-coding regions from the genome. On the other hand, in overexpression experiments, the degree of overexpression of a target protein is quite diverse and varies with the methods and target proteins used. Therefore, to use data on dosage-sensitive proteins obtained by large-scale overexpression experiments for association studies, we need to recognize that the overexpression experiments are inherently quantitative. In the next sections, I discuss the quantitative nature of overexpression experiments, distinguishing them into two different types: absolute and relative.
Figure 1: Expression levels of yeast proteins and their allocations in the yeast proteome. Proteome data (Kulak et al., 2014; Supplemental Table S1) visualized using a proteomap (Liebermeister et al., 2014; www.proteomaps.net/). The area occupied by each process (left) or protein (right) is proportional to its relative protein amount in a cycling yeast population.

Absolute overexpression using promoter swapping

The GAL promoter is commonly used for strong induction of protein expression in the budding yeast. The promoter-swapped gene cassettes are sometimes cloned into multicopy 2-μm plasmids to further increase expression levels (Liu et al., 1992; Espinet et al., 1995; Akada et al., 1997; Stevenson et al., 2001; Boyer et al., 2004; Gelperin et al., 2005; Sopko et al., 2006; Niu et al., 2008; Yoshikawa et al., 2011; Douglas et al., 2012). In these experiments, target proteins are induced at similar high levels. This type of experiment is considered an absolute overexpression experiment. For further discussion, I use the dosage-sensitive protein data set obtained by Sopko et al. (2006) as a representative data set for absolute overexpression experiments. That study is the most extensive, widely cited analysis.

Because expression levels of proteins can vary by more than 10,000-fold (Kulak et al., 2014; Figures 1 and 2), the degree of the overexpression in absolute overexpression experiments should differ by some orders of magnitude. The average expression level of target proteins expressed by genes under the control of the GAL1 promoter on a 2-μm plasmid has been estimated as ~1% of the proteins expressed in the yeast (Tomala and Korona, 2013). According to recent proteome data (Kulak et al., 2014; Figure 1), ~50,000,000 protein molecules are expressed in a budding yeast cell. Only eight glycolytic enzymes are expressed more than 1% of the total proteins (500,000 molecules/cell). Almost all proteins are thus “overexpressed” in the foregoing experiment. However, the degree of overexpression relative to the native level should vary strongly among the analyzed proteins (Figure 2). If the native level of a target protein is 100 molecules/cell, the degree of overexpression will be estimated as 50,000-fold. If the native level of a target protein is 100,000 molecules/cell, the overexpression degree will be 50-fold. Thus the absolute overexpression experiments analyze the consequences of comparably strong production of target proteins independently of their native expression levels. As shown in Figure 3A, proteins with natively low expression are preferentially isolated as the dosage-sensitive proteins in absolute overexpression experiments. If the native expression levels are strongly optimized (see later discussion), the huge degree of overexpression of low-expression proteins can become a bias leading to preferential isolation of those proteins.

Relative overexpression using gene copy number increase

Increasing the gene copy number for the target protein should induce overexpression of the target protein. Multicopy plasmids derived from the 2-μm plasmid are usually used for this purpose in budding yeast (Rine, 1991; Jones et al., 2008). If the expression of the protein is controlled by its native regulatory elements (i.e., promoter and terminator), the expression relative to its native level should increase as the copy number increases. If a multicopy plasmid with a copy number of 10 is used in the experiment, the expression of the proteins coded on the plasmid is expected to increase 10-fold (Figure 2, blue line). This type of experiment is considered a relative overexpression experiment. The genetic tug-of-war (gTOW) experiment (Moriya et al., 2006, 2011, 2012; Kaizu et al., 2010; Makanae et al., 2013) is an example. In the gTOW experiment, the relative overexpression levels (fold increases) causing cellular defects are estimated by measuring the copy number limits of the genes encoding the target proteins. For further discussion, I use the dosage-sensitive protein data set obtained by Makanae et al. (2013) as a representative data set for the relative overexpression experiments; their analysis is the most extensive available.

In relative overexpression experiments, levels of target proteins after overexpression vary, depending on the native expression levels. If a target protein is already highly expressed natively, the after overexpression, a much larger amount of the protein should be present in the cell. As shown in Figure 3A, proteins with high native expression are preferentially isolated as the dosage-sensitive proteins in the relative overexpression experiment. High-level protein expression levels (Table 1) are shown in green circles. The estimated maximal protein expression level (Table 1) is shown in purple.
expression can become a bias in the preferential isolation of such proteins in relative overexpression experiments. This is because the need for massive turnover of proteins causes cellular defects due to process overloads (discussed later).

A series of physiological analyses have been performed using disomic yeast strains with 1 of the 16 chromosomes duplicated in a haploid cell (Torres et al., 2007, 2008, 2010; Sheltzer et al., 2011, 2012; Oromendia et al., 2012; Thorburn et al., 2013; Dephoure et al., 2014; Blank et al., 2015; Bonney et al., 2015). In each disomic cell, almost all of the genes on the duplicated chromosome are overexpressed twofold (Torres et al., 2007). This can be also considered relative overexpression. Analysis of this simultaneous overexpression of hundreds of genes has led to many important insights into the consequences of protein overexpression (Torres et al., 2008; Sheltzer and Amon, 2011; Tang and Amon, 2013; Oromendia and Amon, 2014).

As discussed, absolute and relative overexpression experiments carry different biases due to their quantitative nature. These experiments contribute to the understanding of the consequences of protein overexpression. However, the biases have to be taken into consideration when investigating the mechanistic consequences of protein overexpression using association analyses of large-scale biological data sets. In the following section, I discuss the independent mechanisms of cellular defects upon overexpression of proteins in the context of the quantitative nature of overexpression experiments.

Overexpression experiments have also been performed in other organisms, such as Schizosaccharomyces pombe, Drosophila melanogaster, Arabidopsis thaliana, and Homo sapiens (Prelich, 2012), all using promoter swapping, meaning that they are all absolute overexpression experiments.

CONSEQUENCES OF PROTEIN OVEREXPRESSION

The ways by which protein overexpression can cause cellular defects depend on the protein properties and functions (Prelich, 2012). Several mechanisms of cellular defects upon overexpression of proteins have been suggested from the analyses of dosage-sensitive proteins obtained in genomewide absolute (Vavouri et al., 2009; Ma et al., 2010; Tomala et al., 2014) and relative overexpression experiments (Makanae et al., 2013) and from the analyses of physiological disorders of disomic yeast strains (Sheltzer and Amon, 2011; Tang and Amon, 2013).

Here I focus on four primary mechanisms: resource overload, stoichiometric imbalance, promiscuous interactions, and pathway modulation (Figure 4). The growth defects are largely explained by overload and toxicity (Tomala and Korona, 2013). Overload is an abnormal cellular environment created by enhancement of normal activity or turnover of a protein caused by its overexpression. Toxicity is an effect causing cellular defects due to novel and unrelated properties generated by the overexpression of a protein. Overload is expressed as either resource overload or pathway modulation.

FIGURE 3: Trends in native expression levels of dosage-sensitive proteins and of proteins within various categories. (A) Proteins isolated in absolute and relative overexpression experiments show opposite trends in their native expression levels. Absolute overexpression tends to isolate lowly expressed proteins, whereas relative overexpression tends to isolate highly expressed proteins. Proteins isolated as dosage-sensitive proteins in overexpression experiments are separated by their expression levels, and their proportions are shown. Absolute overexpression, 603 proteins (Sopko et al., 2006); relative overexpression, 682 proteins (Makanae et al., 2013). (B) Proteins within different categories show different trends in their native expression levels, which could be biases to isolate dosage-sensitive proteins in overexpression experiments. Proteins within indicated categories are separated by their expression levels, and their proportions are shown. Transcription factors, 125 proteins (Teixeira et al., 2014; www.yeastact.com/); intrinsic disorder, 439 proteins with GlobPlot score >200 (Linding et al., 2003; http://globplot.embl.de/); protein kinases, 114 proteins (Breitkreutz et al., 2010; www.yeastkinome.com/); membrane proteins; 255 proteins with TMHMM score >100 (Linding et al., 2003; http://globplot.embl.de/); and protein complex members, 1485 proteins (Pu et al., 2009). The original data are given in Supplemental Table S1.
Because the overexpression of glycolytic enzymes might perturb the glycolytic flux, these estimates do not necessarily reflect the maximum expression limits for other proteins. However, if a protein causes a cellular defect at an expression level substantially below this limit, we can assume that mechanisms other than the protein burden effect are involved. Ribosomal proteins are generally highly expressed (Figure 1). Among them, RPL9A and RPS12 have been isolated as dosage-sensitive proteins because they have low copy number limits (Table 1). The estimated protein levels in the overexpression experiments are 1.4–1.9% of the total proteome in terms of amino acids (Table 1), substantially lower than the limit levels calculated using the glycolytic enzymes. We can assume that there is an unknown mechanism other than the protein burden causing the cellular defects upon the overexpression of ribosomal proteins.

Intracellular proteins are folded with the help of molecular chaperones, localized to specific cellular compartments by the translocation and transport machinery, and degraded by the proteasome. Overexpression of proteins with strong demands for those resources might cause resource overloads, leading to cellular defects (Figure 4A). Growth defects in disomic strains have been explained by the overexpression of glycolytic enzymes (Tdh3, Pgk1, Cdc19, and Tpi1; Figure 2, green circles) being determined by the protein burden (Makanae et al., 2013). In gTOW overexpression experiments, these proteins should be expressed up to the levels causing cellular defects due to the protein burden. As shown in Table 1, we can estimate the expression levels by multiplying the native expression levels by the copy number limits of the glycolytic enzymes (the average level is shown as the purple line in Figure 2). The total proteome consists of ~50,000,000 molecules/cell (Kulak et al., 2014). We can estimate the number of amino acids in the total proteome by multiplying the number of protein molecules by their amino acid lengths and adding the results. This estimate gives ~16,000,000,000 amino acids/cell (Supplemental Table S1). Thus we estimate the expression levels of the aforementioned glycolytic enzymes as 14.2–56.5% of the total proteome in terms of amino acids (Table 1); substantially lower than the limit levels calculated using the glycolytic enzymes. We can assume that there is an unknown mechanism other than the protein burden causing the cellular defects upon the overexpression of ribosomal proteins.

Because the overexpression of glycolytic enzymes might perturb the glycolytic flux, these estimates do not necessarily reflect the maximum expression limits for other proteins. However, if a protein causes a cellular defect at an expression level substantially below this limit, we can assume that mechanisms other than the protein burden effect are involved. Ribosomal proteins are generally highly expressed (Figure 1). Among them, RPL9A and RPS12 have been isolated as dosage-sensitive proteins because they have low copy number limits (Table 1). The estimated protein levels in the overexpression experiments are 1.4–1.9% of the total proteome in terms of amino acids (Table 1), substantially lower than the limit levels calculated using the glycolytic enzymes. We can assume that there is an unknown mechanism other than the protein burden causing the cellular defects upon the overexpression of ribosomal proteins.

Intracellular proteins are folded with the help of molecular chaperones, localized to specific cellular compartments by the translocation and transport machinery, and degraded by the proteasome. Overexpression of proteins with strong demands for those resources might cause resource overloads, leading to cellular defects (Figure 4A). Growth defects in disomic strains have been explained...
by the overload of the protein quality control machinery involved in protein folding and degradation (Torres et al., 2010; Oromendia et al., 2012). The protein degradation overload has been also tested by overexpression experiments using the model protein green fluorescent protein (GFP). The addition of a degradation signal to GFP increased its potential to cause cellular defects upon its overexpression (Makanae et al., 2013). As noted earlier, gTOW relative overexpression experiments preferentially isolate natively highly expressed proteins as dosage-sensitive proteins (Figure 3A). Because levels after overexpression of those proteins should be very high (Figure 2), they might include many proteins causing overloads of different cellular resources.

**STOICHIOMETRIC IMBALANCE**

The expression level of each protein (as shown in Figure 1) should be ultimately determined by various types of equilibrium: the balanced numbers of subunits within a protein complex and of protein complexes within a cellular module and balance among cellular modules. Appropriate resource allocation in cellular modules should maximize cellular performance (Liebermeister et al., 2014). For example, the production rate of each subunit of a complex is highly related to the molecular ratio (stoichiometry) of the subunit within the complex (Li et al., 2014). This indicates that cellular systems are optimized to keep protein expression in balance. Because overexpression of a protein complex member perturbs the equilibrium, stoichiometric imbalance is believed to be a primary mechanism of cellular defects upon protein overexpression (balance hypothesis; Papp et al., 2003; Veitia et al., 2008; Figure 4B). Examples of cellular defects caused by overexpression-triggered disruption of the stoichiometric balance in complexes have been reported. Overexpression of a catalytic subunit relative to a regulatory subunit causes constitutive activation/inactivation of the catalytic subunit (de Nadal et al., 1998; Kaizu et al., 2010; Moriya et al., 2011). Overexpression of a subunit in comparison with the expression of other subunits causes an abnormal toxic complex formation (Abruzzi et al., 2002). The disomic yeast can suffer from overloads of the protein quality control resources (most probably caused by unnecessary subunit production) due to stoichiometric imbalances (Torres et al., 2010; Oromendia et al., 2012).

Although stoichiometric imbalance explains the results of individual overexpression experiments, contradictory results have been obtained from large-scale experiments. About 30% of yeast proteins are protein-complex members (Pu et al., 2009). If stoichiometric imbalances of protein complexes are the primary mechanisms of cellular defects after protein overexpression, protein-complex members should be preferentially isolated as the dosage-sensitive proteins. Although initial results supported this hypothesis (Papp et al., 2003), the dosage-sensitive proteins isolated in the absolute overexpression experiment were not preferentially the members of protein complexes (Sopko et al., 2006; Vavouri et al., 2009). However, the dosage-sensitive genes isolated in the relative overexpression experiment preferentially include the complex members (Makanae et al., 2013). In that report, the balance hypothesis was further validated in rescue experiments with simultaneously overexpressed stoichiometric partner subunits. These contradictory results could be explained by differences in the quantitative nature of the experiments. To change the balance, the proteins should be overexpressed in comparison with the native levels. In absolute overexpression experiments, the large variation in the degrees of balance disruptions caused by the overexpressed proteins could hide the cellular defects triggered by stoichiometric imbalance. Protein-complex members tend to be natively highly expressed (Figure 3B). This might also create a bias against the isolation of complex members because the absolute overexpression experiment preferentially isolated proteins with natively low expression, as discussed earlier (Figure 3A).

Note that stoichiometric imbalance in a protein complex does not necessarily cause immediate cellular defects. As described earlier and shown in Figure 4B, cellular defects caused by stoichiometric imbalances in protein complexes depend on the functions and characteristics of the overexpressed proteins. For example, to overload cellular resources, proteins should be highly expressed. It is thus possible that stoichiometric imbalance in a protein with a natively low expression will not cause resource overload.

**PROMISCUOUS INTERACTIONS**

Proteins containing regions without stable structures are called intrinsically disordered proteins (IDPs; Habchi et al., 2014). The ID regions enable flexible protein–protein interactions (Liu and Huang, 2014; Supplemental Table S1). Although the flexible interactions are advantageous to the function of these proteins, there might be some unwanted interactions with nonphysiological partner proteins. The association analysis of absolute overexpression experiments has shown that these promiscuous interactions might be one of the primary mechanisms of cellular defects (Vavouri et al., 2009; Ma et al., 2010; Figure 4C). Despite the tendency of IDPs to have low

---

**TABLE 1:** Estimation of absolute overexpression levels of proteins in relative overexpression experiments.

| Gene name | ORF name | Protein molecules per cell (% of total proteome) (#1) | Gene copy number limit (#2) | Protein length (amino acids) (#3) | Product of #1 × #2 molecules (% of total proteome) (#3) | Product of #1 × #2 × #3 amino acids (% of total proteome) |
|-----------|----------|------------------------------------------------------|-----------------------------|---------------------------------|--------------------------------------------------------|----------------------------------------------------------|
| TDH3      | YGR192C  | 1,575,311 (3.3)                                      | 4.3                         | 332                             | 6,801,949 (14.2)                                        | 2,258,246,952 (14.0)                                    |
| PGK1      | YCR012W  | 561,265 (1.2)                                        | 22.0                        | 416                             | 1,2358,595 (25.8)                                       | 5,141,175,549 (31.9)                                    |
| CDC19     | YAL038W  | 404,162 (0.8)                                        | 31.8                        | 500                             | 12,864,113 (26.9)                                       | 6,432,056,484 (39.9)                                    |
| TPI1      | YDR050C  | 395,237 (0.8)                                        | 68.5                        | 248                             | 27,054,398 (56.5)                                       | 6,709,490,697 (41.6)                                    |
| RPL9A     | YGL147C  | 265,169 (0.6)                                        | 6.2                         | 191                             | 1,642,828 (3.4)                                        | 313,780,172 (1.9)                                       |
| RPS12     | YOR369C  | 258,298 (0.5)                                        | 6.2                         | 143                             | 1,611,136 (3.4)                                        | 230,392,526 (1.4)                                       |
| Total proteome |        | 47,900,214                                           |                             |                                 |                                                        | 16,130,396,439                                         |

The original data are given in Supplemental Table S1.
native expression levels (Figure 3B; Gsponer et al., 2008), they also have been preferentially isolated as dosage-sensitive proteins in gTOW overexpression experiments (Makanae et al., 2013). The hypothetical involvement of promiscuous interactions in cellular defects is thus supported by both absolute and relative overexpression experiments.

Promiscuous interaction–triggered cellular defects are difficult to validate in vivo because of the diversity of the functions and characteristics of IDPs. We do not have good model proteins or model domains for testing such interactions. Regulatory proteins such as transcription factors and signaling molecules often contain ID regions (Supplemental Table S1). Systematic physiological analyses, such as transcriptome analyses of cells overexpressing these IDPs, might distinguish the consequences of promiscuous interactions (e.g., unexpected pathway activation and stress response) from pathway overloads (see later discussion). The degree of false-positive (unexpected) interactions in two-hybrid analyses could be an index for promiscuous interactions because strong promoters are normally used in these experiments (Uetz et al., 2000; Ito et al., 2001).

Strong expression of aggregative proteins, such as the disease-related heterologous proteins misfolded yellow fluorescent protein mutants, causes cellular defects in yeast cells (Geiler-Samerotte et al., 2011; Kaiser et al., 2013; Park et al., 2013). Some protein aggregates are toxic because they capture essential proteins and limiting chaperones into the aggregates (Treusch and Lindquist, 2012; Kaiser et al., 2013; Park et al., 2013). Protein aggregates could cause toxicity, lead to novel and unwanted protein–protein interactions, and overload protein quality control resources. Although overexpressed endogenous proteins sometimes form aggregates in yeast cells, they do not necessarily lead to cellular defects; they might even have a positive effect on environmental adaptation (Alberti et al., 2009). Their toxicities might depend on the proteins with which they interact. Association studies suggest that aggregative proteins tend to be dosage sensitive (Gsponer and Babu, 2012). However, it has not been experimentally demonstrated that protein aggregates are the primary cause of cellular defects after protein overexpression.

PATHWAY MODULATION

The expression level of each protein seems to be optimized to maximize cellular functionality. It is possible that the more the expression level is changed, the more cellular functions are negatively affected. If this were true, proteins with low native expression would be isolated as dosage-sensitive proteins in absolute overexpression experiments. In these experiments, such proteins tend to be overexpressed substantially above their native levels (Figures 2 and 3). Transcription factors, protein kinases, membrane proteins, and IDPs tend to be lowly expressed (Figure 3B). These proteins, except for IDPs, are preferentially isolated as dosage-sensitive proteins in absolute overexpression experiments (Sopko et al., 2006), but not in gTOW relative overexpression experiments (Makanae et al., 2013). One simple explanation is that they are isolated as dosage-sensitive proteins in absolute overexpression experiments because their levels of overexpression are far beyond their physiological concentrations.

Activities of regulatory molecules such as transcription factors and signaling molecules are usually modulated in response to changes in external and internal cellular conditions. Their biochemical characteristics, such as binding constants and enzymatic activities, should differ by some orders of magnitude to avoid unwanted activation/inactivation of target pathways in which they are involved. Due to a mass-action effect, overexpression of inactive forms of these proteins above the threshold might cause untimely activation/inactivation of their target pathways. If this causes cellular defects, these proteins could be dosage sensitive (pathway modulation; Figure 4D; Sheltzer and Amon, 2011). Such pathway modulations have been repeatedly observed in individual overexpression experiments (Prelich, 2012). However, regulatory proteins are not preferentially isolated as dosage-sensitive proteins in gTOW relative overexpression experiments (Makanae et al., 2013). This suggests that the pathway overloads happen only when regulatory proteins are expressed in large excess above their native levels; the responses of biological systems to changes in intracellular parameters are inherently robust (Alon et al., 1999; Little et al., 1999; von Dassow et al., 2000; Kitano, 2004). Further analyses will be required to confirm the pathway modulation mechanisms. Such modulations could be relatively easily detected using transcriptome analyses of the downstream genes specific for each pathway.

CONCLUSION

I have discussed the quantitative nature of overexpression experiments and some primary consequences of protein overexpression as revealed in the large-scale studies. Figure 2 is a conceptual presentation of absolute and relative overexpression experiments. It allows an intuitive understanding of the quantitative nature of overexpression and illustrates the importance of measurements of overexpressed protein levels. Some pioneering studies have revealed the diversity of protein expression levels even in absolute overexpression experiments (Gelperin et al., 2005; Tomala and Korona, 2013; Tomala et al., 2014). It is important to measure precisely the limits of protein expression to obtain a clear picture of the consequences of protein overexpression leading to cellular defects.

Another issue is the confirmation of the hypotheses described. The consequences of protein overexpression are complex because the functions and characteristics of proteins are diverse but interconnected. For example, IDPs include many transcription factors and protein kinases. Overexpression of these regulatory factors could cause both promiscuous interactions and pathway modulations. Overexpression of a protein-complex member could cause pathway activation and overloads. Further systematic analysis of cellular physiology using omics technologies and suppressor mutants (such as in the analyses of disomic strains; Torres et al., 2010) are needed. Systematic genetic interaction analyses (Sopko et al., 2006; Douglas et al., 2012) and analyses using model proteins (Geiler-Samerotte et al., 2011; Makanae et al., 2013; Park et al., 2013; Tomala et al., 2014) should help to dissect the consequences of protein overexpression. Ideally, we should find a biomarker or a reporter associated with each mechanism to assess it properly.

ACKNOWLEDGMENTS

I am sincerely grateful to Takashi Makino, Tohoku University, Sendai, Japan, for various useful suggestions and comments on the manuscript. I am also grateful to Kouji Ishikawa, Okayama University, Okayama, Japan, for his help with data analysis.

REFERENCES

Abruzzi KC, Smith A, Chen W, Solomon F (2002). Protection from free beta-tubulin by the beta-tubulin binding protein Rbl2p. Mol Cell Biol 22, 138–147.

Akada R, Yamamoto J, Yamashita I (1997). Screening and identification of yeast sequences that cause growth inhibition when overexpressed. Mol Gen Genet 254, 267–274.

Alberti S, Halfmann R, King O, Kapila A, Lindquist S (2009). A systematic survey identifies prions and illuminates sequence features of prionogenic proteins. Cell 137, 146–158.
Tomala K, Korona R (2013). Evaluating the fitness cost of protein expression in Saccharomyces cerevisiae. Genome Biol Evol 5, 2051–2060.

Tomala K, Pogoda E, Jakubowska A, Korona R (2014). Fitness costs of minimal sequence alterations causing protein instability and toxicity. Mol Biol Evol 31, 703–707.

Torres EM, Dephoure N, Panneerselvam A, Tucker CM, Whittaker CA, Gygi SP, Dunham MJ, Amon A (2010). Identification of aneuploidy-tolerating mutations. Cell 143, 71–83.

Torres EM, Sokolsky T, Tucker CM, Chan LY, Boselli M, Dunham MJ, Amon A (2007). Effects of aneuploidy on cellular physiology and cell division in haploid yeast. Science 317, 916–924.

Torres EM, Williams BR, Amon A (2008). Aneuploidy: cells losing their balance. Genetics 179, 737–746.

Treusch S, Lindquist S (2012). An intrinsically disordered yeast prion arrests the cell cycle by sequestering a spindle pole body component. J Cell Biol 197, 369–379.

Uetz P, Giot L, Cagney G, Mansfield TA, Judson RS, Knight JR, Lockshon D, Narayan V, Srinivasan M, Pochart P, et al. (2000). A comprehensive analysis of protein-protein interactions in Saccharomyces cerevisiae. Nature 403, 623–627.

Vavouri T, Semple JI, Garcia-Verdugo R, Lehner B (2009). Intrinsic protein disorder and interaction promiscuity are widely associated with dosage sensitivity. Cell 138, 198–208.

Veitia RA, Bottani S, Birchler JA (2008). Cellular reactions to gene dosage imbalance: genomic, transcriptomic and proteomic effects. Trends Genet 24, 390–397.

von Dassow G, Meir E, Munro EM, Odell GM (2000). The segment polarity network is a robust developmental module. Nature 406, 188–192.

Wagner A (2005). Energy constraints on the evolution of gene expression. Mol Biol Evol 22, 1365–1374.

Yoshikawa K, Tanaka T, Ida Y, Furusawa C, Hirasawa T, Shimizu H (2011). Comprehensive phenotypic analysis of single-gene deletion and overexpression strains of Saccharomyces cerevisiae. Yeast 28, 349–361.

Zaslaver A, Mayo AE, Rosenberg R, Bashkin P, Sberro H, Tsalyuk M, Surette MG, Alon U (2004). Just-in-time transcription program in metabolic pathways. Nat Genet 36, 486–491.