

makes CpG islands ideal sequences for easy access by transcription activators and RNA Pol II complexes, hence their prevalent use in the promoters of house-keeping genes in mammals. Previously, poly (dA:dT) tracts in the yeast genome were shown to have similar molecular properties and transcriptional activating functions (Anderson and Widom, 2001; Iyer and Struhl, 1995). It appears that some CpG promoters have evolved to regulate rapid signal-induced transcription in response to diverse stimuli. This requires a signal-induced activator, which converts pre-engaged RNA Pol II into a form that can efficiently undergo transcription elongation coupled to RNA splicing. Thus, to end with

the golfing analogy, CpG islands are the “transcriptional tee off” areas of the mammalian genome that provide a nucleosome-depleted surface for the efficient assembly and partial activation of RNA Pol II complexes prior to their productive launch down genomic “fairways” in a signal-dependent manner.

REFERENCES

Anderson, J.D., and Widom, J. (2001). *Mol. Cell Biol.* *21*, 3830–3839.

Bentley, D.L. (2005). *Curr. Opin. Cell Biol.* *17*, 251–256.

Core, L.J., and Lis, J.T. (2008). *Science* *319*, 1791–1792.

Hargreaves, D.C., Horng, T., and Medzhitov, R.

(2009). *Cell*, this issue.

Iyer, V., and Struhl, K. (1995). *EMBO J.* *14*, 2570–2579.

Jang, M.K., Mochizuki, K., Zhou, M., Jeong, H.S., Brady, J.N., and Ozato, K. (2005). *Mol. Cell* *19*, 523–534.

Ramirez-Carrozzi, V.R., Braas, D., Bhatt, D.M., Cheng, C.S., Hong, C., Doty, K.R., Black, J.C., Hoffmann, A., Carey, M., and Smale, S.T. (2009). *Cell*, this issue.

Ramirez-Carrozzi, V.R., Nazarian, A.A., Li, C.C., Gore, S.L., Sridharan, R., Imbalzano, A.N., and Smale, S.T. (2006). *Genes Dev.* *20*, 282–296.

Sims, R.J., 3rd, Belotserkovskaya, R., and Reinberg, D. (2004). *Genes Dev.* *18*, 2437–2468.

Suzuki, M.M., and Bird, A. (2008). *Nat. Rev. Genet.* *9*, 465–476.

Disorder, Promiscuity, and Toxic Partnerships

Edward M. Marcotte^{1,2,*} and Mark Tsechansky²

¹Department of Chemistry and Biochemistry

²Center for Systems and Synthetic Biology, Institute for Cellular and Molecular Biology
University of Texas at Austin, 2500 Speedway, Austin, TX 78712-1064, USA

*Correspondence: marcotte@icmb.utexas.edu

DOI 10.1016/j.cell.2009.06.024

Many genes are toxic when overexpressed, but general mechanisms for this toxicity have proven elusive. Vavouri et al. (2009) find that intrinsic protein disorder and promiscuous molecular interactions are strong determinants of dosage sensitivity, explaining in part the toxicity of dosage-sensitive oncogenes in mice and humans.

In yeast, approximately 15%–20% of genes are toxic when overexpressed (Gelperin et al., 2005; Sopko et al., 2006), a phenomenon known as dosage sensitivity. A similar fraction of genes are known to be essential for growth in standard laboratory conditions and can lead to cell death upon deletion or downregulation (Giaever et al., 2002). Too much expression of a gene is not the same as too little, and it is not surprising that the mechanisms of toxicity appear quite different. What has been surprising is that a general understanding is emerging for the causes of cell death stemming from loss of function but not from overexpression. In the case of loss of function, essential genes tend to pref-

erentially interact with each other, both as hubs in gene networks and as components of essential protein complexes (Figure 1). If any one component of an essential complex is deleted, the whole system fails. This trend explains a large fraction of the cases in which reducing or abolishing expression of a particular gene results in lethality (Hart et al., 2007; Wang et al., 2009). In contrast, dosage sensitivity has thus far eluded any such general explanation. One is now provided, at least in part, by Vavouri et al. (2009) in this issue.

This notable lack of a general mechanism for dosage sensitivity has been a source of considerable frustration given its contribution to many diseases, includ-

ing a wide variety of amyloid, prion, and other diseases of protein misfolding and aggregation. For example, one of the first mouse models to develop the predominant features of Alzheimer’s disease entails overexpressing a variant of the human amyloid precursor protein (Games et al., 1995). Overexpression of dosage-sensitive oncogenes also leads to a variety of human cancers. For example, it has recently been shown that overexpression (up to 100-fold) of the angiotensin II receptor type I enables a highly invasive phenotype in mammalian epithelial cells, thereby defining a subpopulation of breast cancers (10%–20%) that might respond to treatment with angiotensin receptor antagonists (Rhodes et al., 2009). A better

notion of why some genes are toxic when overexpressed and others are not might therefore help in understanding mechanisms of disease progression and even in designing new therapeutics.

In order to search for general mechanisms of dosage sensitivity, Vavouri and colleagues tested a variety of likely factors, including protein abundance, aggregation potential, half-life, and the degree to which codon usage is optimized for protein translation, all of which failed to provide a reliable prediction (Vavouri et al., 2009). Instead the strongest predictors of dosage sensitivity are intrinsic protein disorder and the tendency to participate directly in large numbers of pairwise protein-protein interactions. The latter is especially predictive when either partner contains functional motifs completely contained in a single linear stretch of amino acids. These factors do not explain all (or even most) of the trend—this is clearly only part of the story. Nonetheless these factors provide successful predictions of dosage-sensitive genes in yeast and metazoans, to the extent that a computer algorithm trained to identify such proteins correctly predicts the dosage toxicity of 6 of 8 genes tested in worms.

One interpretation, offered by Vavouri et al., is that intrinsically disordered proteins are more likely to participate in promiscuous protein interactions when overexpressed, simply as a consequence of mass action. This explanation draws support from two intriguing sets of contrasting observations: First, dosage-sensitive genes show higher numbers of direct protein interactions (measured by the yeast two-hybrid assay) than dosage-insensitive genes but interestingly do not share more interaction partners in protein complexes (as measured by affinity purification or coimmunoprecipitation). This is in strong contrast to genes essential upon knockout or knockdown (Figure 1) and suggests that there is something special about the excess participation specifically in pairwise interactions. Conceivably, the yeast two-

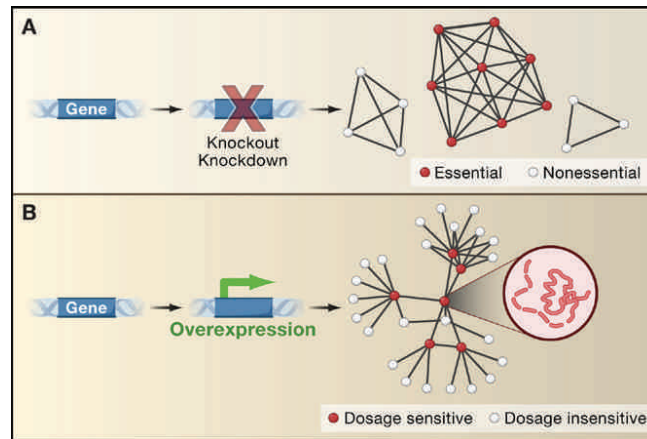


Figure 1. Too Much Gene Expression Is Not the Same as Too Little

(A) Genes that lead to cell death or growth defects upon knockout or knockdown tend to preferentially associate with each other in the same protein complexes, shown here by representing proteins as circles and protein-protein associations by lines forming a network of protein interactions. Complexes of essential proteins tend to contain more protein constituents than nonessential complexes. These trends explain a large fraction (>30%) of the essential genes in yeast cells (Hart et al., 2007; Wang et al., 2009).

(B) In contrast to genes that are essential upon knockout or knockdown, a different mechanism appears to underlie the toxicity observed upon gene overexpression (Vavouri et al., 2009). Such dosage-sensitive genes tend to participate in larger numbers of direct pairwise protein interactions, shown here with dosage-sensitive genes forming hubs in a network of yeast two-hybrid protein interactions. Dosage-sensitive genes also tend to be intrinsically disordered, as illustrated by the dashed line for the protein structure pictured in the inset. Overexpression of an intrinsically disordered protein may increase its already high degree of interaction promiscuity, leading to toxic partnerships and detrimental effects on the cell.

hybrid assay might preferentially detect interactions that result from non-native expression of highly interactive (or even just sticky) proteins; the assay requires constitutive nuclear expression of proteins whose native expression is neither constitutive nor nuclear, and in fact, hub proteins are known to exhibit increased disorder (Haynes et al., 2006). Nonetheless, the signal for dosage sensitivity appears stronger than this trend alone. Second, dosage-sensitive proteins show increased intrinsic disorder, but not increased tendencies to self-aggregate, at least as predicted computationally. Thus, dosage sensitivity does not appear to be predominantly driven by self-aggregation, although it is possible that this trend might be overturned with direct experimental data. Nonetheless, such aggregation would still be consistent with the notion of mass-action-driven interaction promiscuity being a major determinant of dosage sensitivity.

This model in which increased dosage drives toxic interactions by mass action—tested in yeast, flies, and

worms—might also explain a fraction of dosage-sensitive oncogenes in humans and mice. These cases abound in the literature, especially for oncogenic transcription factors, such as the gene *Oct-4*. *Oct-4* is a homeobox gene responsible for maintaining pluripotency in embryonic stem cells. Strikingly, as the expression level of *Oct-4* is varied from 0% to 150% relative abundance, the incidence of tumors formed following injection into syngeneic mice varies from 4% to 64%–83%, an effect possibly stemming from promiscuous activation of growth factors (Gidekel et al., 2003). Vavouri et al. systematically analyze a set of dosage-sensitive oncogenes. They show that genes known to be amplified and causally linked to cancer tend to encode proteins that are enriched with intrinsically disordered segments, contain or bind linear functional motifs, and make many direct pairwise physical interactions with other proteins. Thus, at least a substantial portion of these genes are consistent with the model.

Intrinsic disorder and protein-protein interaction promiscuity account only for a portion of dosage-sensitivity cases. This begs the question of what accounts for the rest of the cases? A test of the dosage-sensitive yeast genes assembled by Vavouri et al. shows that the 839 dosage-sensitive genes are strongly statistically enriched for proteins with transcription factor and DNA-binding activities ($p \leq 10^{-10}$; hypergeometric probability), with 110 of the proteins involved in some aspect of transcriptional regulation. The set is enriched for many other regulatory functions as well, including 50 genes known to participate in phosphorylation, dephosphorylation, or autophosphorylation, and 30 in RNA transport. Thus, the notion of mass-action-driven protein interaction promiscuity can probably be generalized to mass-action-driven promiscuity of

many types of macromolecular interactions, including transcription factor-target interactions, kinase/phosphatase-target interactions, and so on. Vavouri et al. point out a prime candidate for overexpression-induced promiscuity leading to toxicity, overexpression of microRNAs, which might lead to many off-target gene regulatory interactions. Therefore, although the current data primarily implicate promiscuous protein interactions and disordered proteins, a broader exploration of regulatory promiscuity may yet reveal mechanisms of dosage sensitivity with even greater predictive value.

REFERENCES

- Games, D., Adams, D., Alessandrini, R., Barbour, R., Berthelette, P., Blackwell, C., Carr, T., Clemens, J., Donaldson, T., Gillespie, F., et al. (1995). *Nature* 373, 523–527.
- Gelperin, D.M., White, M.A., Wilkinson, M.L., Kon, Y., Kung, L.A., Wise, K.J., Lopez-Hoyo, N., Jiang, L., Piccirillo, S., Yu, H., et al. (2005). *Genes Dev.* 19, 2816–2826.
- Giaever, G., Chu, A.M., Ni, L., Connelly, C., Riles, L., Veronneau, S., Dow, S., Lucau-Danila, A., Anderson, K., Andre, B., et al. (2002). *Nature* 418, 387–391.
- Gidekel, S., Pizov, G., Bergman, Y., and Pikarsky, E. (2003). *Cancer Cell* 4, 361–370.
- Hart, G.T., Lee, I., and Marcotte, E.R. (2007). *BMC Bioinformatics* 8, 236.
- Haynes, C., Oldfield, C.J., Ji, F., Klitgord, N., Cusick, M.E., Radivojac, P., Uversky, V.N., Vidal, M., and Iakoucheva, L.M. (2006). *PLoS Comput. Biol.* 2, e100. 10.1371/journal.pcbi.0020100.
- Rhodes, D., Ateeq, B., Cao, Q., Tomlins, S., Mehra, R., Laxman, B., Kalyana-Sundaram, S., Lonigro, R., Helgeson, B., Bhojani, M., et al. (2009). *Proc. Natl. Acad. Sci. USA*. Published online June 1, 2009. 10.1073/pnas.0900351106.
- Sopko, R., Huang, D., Preston, N., Chua, G., Papp, B., Kafadar, K., Snyder, M., Oliver, S.G., Cyert, M., Hughes, T.R., et al. (2006). *Mol. Cell* 21, 319–330.
- Vavouri, T., Semple, J., Garci-Verdugo, R., and Lehner, B. (2009). *Cell*, this issue.
- Wang, H., Kakaradov, B., Collins, S.R., Karotki, L., Fiedler, D., Shales, M., Shokat, K.M., Walther, T.C., Krogan, N.J., and Koller, D. (2009). *Mol. Cell. Proteomics* 8, 1361–1381.

Chromatin Places Metabolism Center Stage

Andreas G. Ladurner^{1,2,*}

¹Genome Biology Unit

²Structural and Computational Biology Unit

European Molecular Biology Laboratory, Meyerhofstrasse 1, 69117 Heidelberg, Germany

*Correspondence: ladurner@embl.de

DOI 10.1016/j.cell.2009.06.025

Dynamic changes in histone and transcription factor acetylation modulate gene expression. A study in *Science* (Wellen et al., 2009) reports that changes in glucose metabolism alter the availability of acetyl-CoA, the essential cofactor for protein acetylation. These findings reveal a direct connection between central metabolism and mammalian gene expression.

Animals constantly fine-tune their profiles of gene expression in order to adapt to the physical demands of the environment and to the availability of nutrients, be it feast or famine. Moreover, different cell types vary in their metabolic and gene expression requirements. In a recent issue of *Science*, Wellen et al. (2009) provide evidence for a mechanism in mammalian cells that links the production of acetyl-CoA from citrate to the regulation of metabolic genes through alterations in histone acetylation. These findings illustrate how central metabolism, too often viewed as a subject for textbooks, still conceals plenty of surprises.

Under normal circumstances, most adult cells utilize available nutrients efficiently, primarily for energy production.

This situation is markedly different in cells that grow rapidly or that metabolize nutrients quickly, such as in the fetus, during cellular insults (such as extensive DNA damage) or in tumors. Cancer cells, in particular, become very adept at balancing the need for energy production with anabolic processes, such as the increased synthesis of lipids, nucleic acids, and proteins. Otto Warburg's hypothesis (Warburg, 1956; Hsu and Sabatini, 2008) that cancer cells tweak their metabolic pathways in favor of aerobic glycolysis has largely been validated experimentally, including the recent identification of pyruvate kinase isoforms that help to switch glucose metabolism from the production of CO₂ to anabolic growth (Christofk et al., 2008).

Cells maintain this metabolic balancing act by integrating the control of gene expression with metabolism and by compartmentalization between mitochondria (for energy generation), the cytoplasm (for biomolecule production), and the nucleus (for gene expression). Although this division of labor promotes efficiency in dealing with available nutrients, the system needs to be able to reshuffle its global activities to meet the increased energy demands of growing cells. Growth factors stimulate cells to take up more nutrients by changing gene expression patterns, a process which is promoted by shuttling transcription factors into the nucleus or by activating signaling kinases, for example. In addition, it is now clear that in order to run the whole system of metabolic control