Molecular Systems Biology (2005) doi:10.1038/msb4100006

www.molecularsystemsbiology.com
Article number: 2005.0003

molecular systems biology

NEWS AND VIEWS

Multifunctional genes

Jeroen van de Peppel and Frank CP Holstege

Department for Physiological Chemistry, University Medical Center Utrecht, Utrecht, The Netherlands

Molecular Systems Biology 29 March 2005; doi:10.1038/msb4100006

How many genes encode proteins involved in multiple cellular processes? Genetic pleiotropy becomes apparent when mutations in a single gene give rise to diverse phenotypes. Such pleiotropy is well documented for several human disease genes but has not yet been systematically investigated. In their article, Dudley *et al* (2005) describe the first genome-wide study of genetic pleiotropy. Their findings have implications for several fields, including genetics, drug discovery, bioinformatics and evolution.

To discover and study pleiotropic genes, the authors have harnessed the workhorse of functional genomics, the budding yeast *Saccharomyces cerevisiae*. Single gene deletion strains, representing all 4700 nonessential genes, were analyzed for growth under 21 different conditions, ranging from induction of DNA damage to iron limitation. Of these mutant strains, 551 exhibited a phenotype in only one or two conditions. Strikingly, 216 strains show growth defects in 3–14 conditions, a much higher degree of pleiotropy than may be expected based on a random distribution of phenotypes.

It is impressive that this study encompasses newly generated data, aimed at studying pleiotropy in particular. Whereas systematic gene silencing using RNAi is still at the teething stage for other model organisms (Huppi *et al*, 2005), the yeast research community has been able to establish a comprehensive collection of strains bearing single gene knockouts of all known genes (Giaever *et al*, 2002). Systematic screening of phenotypes in yeast has therefore been described before. Here, the study design was specifically aimed at countering pitfalls that may confound analyses of pleiotropy. The study made use of diploid strains, which are probably less prone to spurious mutations, and included a very diverse set of conditions to determine unrelated phenotypes. The amount of newly generated data backs up the claim that the developed approaches are cost-effective. It also holds the promise of more systematic analyses of gene function in the future.

Does finding diverse growth defects after deleting a single gene implicate a role for that gene in diverse processes? This is extremely difficult to demonstrate unequivocally. It also depends on the definition of what constitutes diverse function. Should this be completely diverse molecular function or identical molecular function, but required for different processes or under different conditions? The question also depends on how phenotypes should be interpreted and linked to cellular function. Some seemingly diverse conditional growth defects may reflect a common deficiency in a single underlying cellular process. Nevertheless, the authors' arguments that pleiotropy often does imply roles in diverse cellular processes are quite strong. These arguments are formed by analyzing the functional categories of genes represented by groups of pleiotropic genes that share subsets of phenotypes. The analysis is an excellent demonstration of the power of biclustering algorithms (Tanay *et al*, 2004), which are becoming increasingly important as functional genomic studies grow to be more data-rich.

Several classes of genes are found to be over-represented in the pleiotropic genes discovered in yeast. Vacuolar organization and biogenesis genes seem highly pleiotropic, as do genes involved in transcription by RNA polymerase II. The latter may be expected given that transcription is an integral part of most cellular responses. A similar argument may be made for genes involved in intracellular transport, in particular of proteins. It would have been interesting to learn whether particular subtypes of transcription factors, such as global coregulators, are more pleiotropic than others. However, such analyses are also limited by the degree and accuracy to which this type of subclassification has been systematically carried out for database annotation. Another limitation is the necessary restriction to nonessential genes, that is, genes not required under standard conditions of growth. It is to be expected that genes with roles in diverse cellular processes are more likely to be essential and may perhaps also encompass other classes of pleiotropic genes.

This study is the first of its kind. Despite representing a preliminary, likely conservative estimate of pleiotropy, it is clear that there is a relative overabundance of genes with roles in diverse processes. This finding underscores the requirement for taking pleiotropy into account when considering drug treatment, drug design and side effects. Similar analyses in mammalian cells will be required to assess whether prevalently targeted proteins have pleiotropic characteristics. Another question that arises from this study is how the findings should be placed in evolutionary theory. Does pleiotropy dampen adaptability or do organisms benefit from multifunctional genes? Extending such analyses to more organisms is required to investigate how widespread pleiotropy is, and it will also determine whether pleiotropic genes are more highly conserved.

The finding of pleiotropic genes may also be able to shed light on the validity of figures depicted in the current epidemic of publications containing gene interaction networks (e.g. Kemmeren *et al*, 2004). These networks arise from analyses of diverse types of functional genomic data and reveal relationships between genes. Such networks contain many hubs. These are genes that are connected to a significantly large number of other genes in the network. One prediction is that such hubs should be over-represented in the pleiotropic genes

described here. If this is not always the case, then such a comparison may at least reveal which types of networks best capture genes with related phenotype profiles and are therefore more likely to represent related function.

Validation of gene networks is only one potential additional application of the results of this study. Other applications of the phenotype data, such as discovering novel categories of functionally related genes, are described by the authors. As with other systematically derived data sets, more applications will follow. Besides providing thought-provoking insights into genetic pleiotropy, the work itself is also multifunctional.

References

Dudley AM, Janse DM, Tanay A, Shamir R, Church GM (2005) A global view of pleiotropy and phenotypically derived gene function in yeast. *Mol Sytems Biol* 2005 doi:10.1038/msb4100004

Giaever G, Chu AM, Ni L, Connelly C, Riles L, Veronneau S, Dow S, Lucau-Danila A, Anderson K, Andre B, Arkin AP, Astromoff A, El-Bakkoury M, Bangham R, Benito R, Brachat S, Campanaro S, Curtiss M, Davis K, Deutschbauer A, Entian KD, Flaherty P, Foury F, Garfinkel DJ, Gerstein M, Gotte D, Guldener U, Hegemann JH, Hempel S, Herman Z, Jaramillo DF, Kelly DE, Kelly SL, Kotter P, LaBonte D, Lamb DC, Lan N, Liang H, Liao H, Liu L, Luo C, Lussier M, Mao R, Menard P, Ooi SL, Revuelta JL, Roberts CJ, Rose M, Ross-Macdonald P, Scherens B, Schimmack G, Shafer B, Shoemaker DD, Sookhai-Mahadeo S, Storms RK, Strathern JN, Valle G, Voet M, Volckaert G, Wang CY, Ward TR, Wilhelmy J, Winzeler EA, Yang Y, Yen G, Youngman E, Yu K, Bussey H, Boeke JD, Snyder M, Philippsen P, Davis RW, Johnston M (2002) Functional profiling of the *Saccharomyces cerevisiae* genome. *Nature* **418:** 387–391

Huppi K, Martin SE, Caplen NJ (2005) Defining and assaying RNAi in mammalian cells. *Mol Cell* **17:** 1–10

Kemmeren P, Kockelkorn TT, Bijma T, Donders R, Holstege FC (2004) Predicting gene function through systematic analysis and quality assessment of high-throughput data. *Bioinformatics* [E-pub ahead of print 16 November 2004]

Tanay A, Sharan R, Kupiec M, Shamir R (2004) Revealing modularity and organization in the yeast molecular network by integrated analysis of highly heterogeneous genomewide data. *Proc Natl Acad Sci USA* **101:** 2981–2986