View metadata, citation and similar papers at core.ac.uk
brought to you by CORE
provided by Elsevier - Publisher Connector

Cell

Genetic Redundancy: New Tricks for Old Genes

Ran Kafri,[1,*] Michael Springer,[1] and Yitzhak Pilpel[2,*]
[1]Department of Systems Biology, Harvard Medical School, Boston, MA 02115, USA
[2]Department of Molecular Genetics, Weizmann Institute of Science, Rehovot 76100, Israel
*Correspondence: ran_kafri@hms.harvard.edu (R.K.), pilpel@weizmann.ac.il (Y.P.)
DOI 10.1016/j.cell.2009.01.027

Many crucial components of signal transduction, developmental, and metabolic pathways have functionally redundant copies. Further, these redundancies show surprising evolutionary stability over prolonged time scales. We propose that redundancies are not just archeological leftovers of ancient gene duplications, but rather that synergy arising from feedback between redundant copies may serve as an information processing element that facilitates signal transduction and the control of gene expression.

Functional redundancy due to gene duplications is a characteristic feature of many biological systems. Examples of redundant gene duplicates are replete within almost every signaling, developmental, or metabolic context. Such redundancies include those, for example, in the Hox gene cluster and among myogenic regulators (MyoD, Myf5, myogenin, and Mrf4). In signaling cascades, crucial signaling components are frequently associated with redundant isoforms exemplified by the MAP kinases, NF-κB inhibitors, Wnt proteins, and more. Complementing these numerous individual examples, four recent systematic studies in the budding yeast *Saccharomyces cerevisiae* now independently provide comprehensive lists of functionally redundant gene pairs (Dean et al., 2008; DeLuna et al., 2008; Kafri et al., 2008; Musso et al., 2008; Figure 1B). Interestingly, despite this prevalence of documented examples covering a diversity of organisms and systems, a question rarely asked is, "why have two when one seems to be enough?"

Functions of Redundant Duplicates

The central question we explore here is whether there are recognizable functional facets through which functional redundancy may be exploited by biological systems. Given that no two genes are absolutely identical, we use the term "redundancy" to refer to pairs of homologous genes that, by means of partial overlap in functions, compensate for each other's loss. The unprecedented availability of a systematic list of biologically redundant gene duplicates from one species allows an examination of the biological roles represented among redundant duplicates. Inspection of the 239 redundant duplicates of *S. cerevisiae* (Dean et al., 2008; DeLuna et al., 2008; Kafri et al., 2008; Musso et al., 2008) reveals that ~28% of these are associated with signal transduction (such as, the Ser/Thr receptor tyrosine kinases and G-coupled protein receptors), ~25% are metabolic enzymes, ~16% are ribosomal proteins, ~8% are membrane transporters, and ~7% are

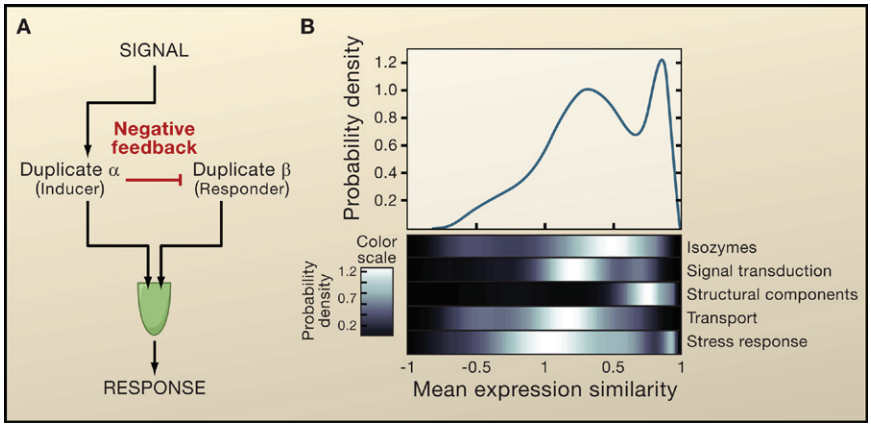

Figure 1. New Roles for Genetic Redundancy

(A) Redundancy of two duplicate genes (α and β) used by a responsive circuit. Although duplicates α and β perform the same molecular function, they differ in how they are regulated. Specifically, duplicate β is repressed directly or indirectly by duplicate α such that its protein level responds reciprocally to changes in the levels of its partner. In such cases, an inductive signal upregulating α will result in downregulation of its partner, β. Given that both duplicates perform the same molecular function, output is defined as the sum of the levels of the two proteins. In this design, fluctuations of α may be balanced by a reciprocal response from β generating a nonfluctuating output.

(B) Similar expression patterns for redundant gene duplicates. 239 pairs of redundant gene duplicates were collated and an expression similarity score (Pearson coefficient) was calculated separately for each duplicate pair (DeLuna et al., 2008; Kafri et al., 2008; Musso et al., 2008). Shown is the probability density (calculated using MATLAB's kernel density function) for the expression similarity scores (http://longitude.weizmann.ac.il). The lower panel shows the relative abundance (depicted by a continuous color scale) of different functional groups across the expression similarity scale. For a given functional classification, lighter colors on the color scale correspond to a higher proportion of duplicate gene pairs associated with the expression similarity score specified by the x axis. Redundant components of signal transduction pathways or isozymes are differentially regulated. This contrasts with functionally redundant structural components such as ribosomal proteins that are tightly coregulated.

stress response genes. Interestingly, in vertebrates, duplicates that have been conserved since the split between cartilaginous and bony fish were found to be specifically enriched for signal transduction and developmental functions (Putnam et al., 2008). Similar enrichment for signal transduction functions among gene duplicates in plants has also been reported.

The association of redundancy with signal transduction should not be surprising. In fact, redundant isoforms of kinases, phosphatases, and other posttranscriptional regulators occur in most signal transduction pathways. Yet, from the systems biology perspective, this tendency of redundant isoforms to be associated with signaling components raises intriguing possibilities. In particular, it may be interpreted to suggest the existence of mechanisms by which functional overlap contributes to or facilitates signal transduction.

If cellular networks do, as we would like to suggest, exploit the redundancy of signaling components then we would expect regulatory designs that allow redundant isoforms to work synergistically (Figure 1A). The 59 redundant pairs of signal transduction genes collected from *S. cerevisiae* (Dean et al., 2008; DeLuna et al., 2008; Kafri et al., 2008; Musso et al., 2008) reveal that redundant isoforms are not coregulated in their expression patterns (Figure 1B). Such differential regulation of redundant duplicates has been presumed to be a general attribute of evolutionarily conserved redundancies. Although we currently lack systematic evidence to substantiate these conclusions for organisms other than yeast, it is interesting to speculate further. For example, from a series of well-studied vertebrate developmental pathways it has been noted that redundant regulators are typically temporally or spatially distinct in their expression patterns (Kafri et al., 2006). Furthermore, redundant duplicates are typically crossregulated by negative feedback that allows one of the redundant isoforms (the responder) to respond to an alteration in expression or function of its partner (the inducer) (Figure 1A). Knocking out one isoform results in an upregulation of its redundant partner. Good examples of this are provided by the four master regulators of vertebrate skeletal muscle development: MyoD, Myf5, myogenin, and Mrf4, encoded by the MRF gene family. These four transcription factors specify differentiation of mesoderm to skeletal muscle and originated (with the appearance of fish) from early gene duplication events. Interestingly, despite their long evolutionary separation, these transcriptional regulators have largely conserved their functional redundancies. Organisms may exploit mutual repression among such redundant regulators, for example, to overcome stochastic fluctuations in protein expression. In such cases, expression of one redundant copy may be induced when expression of the repressing partner is temporarily reduced, thus negating the disruption.

A key aspect of these regulators of skeletal muscle development is the negative feedback observed between members of a redundant pair. For example, MyoD and Myf5 are expressed in separate cell lineages (Haldar et al., 2008), but mutations in MyoD induce increased proliferation of the Myf5-positive cell lineage thus boosting expression of the Myf5 redundant isoform. In this case, extracellular signals regulate a "responsive circuitry" comprising MyoD and Myf5 that effectively buffers against MyoD mutations. In the case of Mrf4 and myogenin, the responsive circuitry induces expression of myogenin to circumvent the effects of mutations in Mrf4. Importantly, such responsive backup circuits are not unique to the myogenic pathway and have been reported for numerous developmental and signaling pathways (for example, redundancy of the developmental regulators Pax1 and Pax9 in mouse or dlx3 and dlx7 in zebrafish).

A Tighter Grip on Flux Control

A key requirement of cellular regulation is the maintenance of various biochemical or metabolic fluxes (that is, the rate of biochemical interconversions and flow of metabolites in the cell) despite large changes in the external milieu and nutrient availability. On evolutionary time scales, adaptation to extreme environmental change is sometimes achieved by preservation of gene duplicates. This is illustrated by observations of adaptive gene amplifications in response to antibiotics, anticancer drug treatments, and nutrient limitations. Furthermore, a genome-wide analysis showed that enzymes acting in reactions with higher metabolic fluxes (that is, pathways with increased metabolic rates) are more likely to have duplicate partners (Conant and Wolfe, 2007; Papp et al., 2004).

For a single cell, the ability to quickly and efficiently respond to fluctuating environments is crucial and offers an obvious evolutionary advantage. One avenue through which functional redundancy could be used to facilitate this ability is by exploiting the differential efficiencies generated by divergence. For example, in yeast, the HXT gene family encodes a redundant set of membrane hexose transporters with varying affinities for glucose and consequently different transport efficiencies (Ozcan and Johnston, 1999). This variation together with glucose-tuned regulation enables the control of glucose fluxes: high-affinity transporters are expressed when glucose is limited, and low-affinity transporters are expressed when glucose is abundant (Ozcan and Johnston, 1999). This design allows the cell to adapt to changes in the availability of external glucose. Other examples of this phenomenon include transport of iron, copper, manganese, zinc, and other metals. In some cases, high-affinity transporters may also have the advantage of higher specificity, that is, their affinity for competing ligands may be reduced. As a consequence, they may transport more of their designated target ligand compared with other competing molecules. Low-affinity transporters for some ligands may have a unique additional role. As external nutrients decrease, the influx of ligand transported by a low-affinity transporter will decrease at a concentration where the high-affinity transporter is still saturated. The cell can respond to this decrease in flux by upregulating the high-affinity transporter. Through this mechanism, the low-affinity transporter acts both as a transporter in replete environments and as a sensor when there is a decrease in nutrients. In all of these cases, functional redundancy of differentially expressed partners is exploited to increase the cell's ability to sense and respond to unpredictable changes in the environment. We suggest that although in many of these cases specialization of at least one of the duplicates has occurred, the protein's function is optimally performed through synergy between the protein duplicates.

Processing Xenobiotic Information

Some of the largest gene families manifest partial redundancy that is collectively used for sophisticated processing of external information. Many of these gene families evolved to interact with a huge repertoire of previously unspecified xenobiotic chemicals. Two prominent examples are the olfactory receptors that detect volatile odorants and the cytochrome P450 enzymes that detoxify foreign substances in the liver. How can a limited set of 500–1000 mammalian odorant receptors bind to and uniquely "sense" an unlimited set of volatile compounds? The solution appears to rely heavily on functional overlap and partial redundancy among the olfactory receptors. The prevailing notion is that the binding spectra of each of the receptors may overlap widely so that each odorant may be bound by a unique combination of receptors, each with potentially different affinities. Partial redundancy between the detectors appears to underlie that ability to uniquely represent each of the inputs prior to further neuronal processing. Redundancy among olfactory receptors did not disappear during evolution, and the repertoire grew to become one of the largest gene families in the mammalian genome. As in other cases mentioned above, partial compensation may arise when an olfactory receptor becomes mutated. For instance, upon mutation of the highest-affinity receptor for a particular odorant, the second highest-affinity receptor may now bind to the odorant, preserving the ability to detect that odorant, albeit with a potentially different olfactory sensation (Lancet et al., 1993). Here, too, exploiting redundancy for accurate information processing requires paralogs that are dissimilar in expression, that is, each olfactory neuron "chooses" randomly to express only one olfactory gene from the repertoire and to exclude the rest.

The Selective Advantage of Redundancy

From an evolutionary perspective, redundancy is thought to buffer phenotypes from genomic variations by reducing the phenotypic cost of mutations, consequently increasing an organism's ability to evolve (evolvability) (Kirschner and Gerhart, 2005). Yet this very fact also renders redundancy evolutionarily unstable and functional overlap is, typically, rapidly lost due to divergence. Yet, recent evidence shows that this instability is not the inevitable fate of all redundant pairs. In fact, for a significant proportion of duplicates, redundancy is stable on evolutionary time scales, with examples ranging from 80 (Tischler et al., 2006) to 100 (DeLuna et al., 2008; Musso et al., 2008) million years of evolutionary conservation. The prolonged evolutionary conservation of these overlapping functions attests to their importance to the fitness of the organism.

The first theoretical basis offering an explanation for the conservation of genetic redundancy was provided by Nowak (Nowak et al., 1997). In that work, the authors considered a population of organisms in which some essential function is redundantly performed by genes at either of two loci, *A* and *B*. Using elegant mathematical arguments, Nowak described three scenarios under which both redundant loci could stably coexist in the population. The first of these scenarios described situations where one of the loci, say *A*, works at a somewhat higher efficiency than its partner but is also exposed to higher mutation rates leading to nonfunctional forms, *a* and *b*. In these situations, Nowak and colleagues calculated that the alleles for both redundant duplicates could stably coexist through selection of the *Ab* genotype. The second scenario describes duplicates that are redundant only with respect to a certain function, with genes being maintained by selection because of another independent function. In the last scenario, a certain gene fails, from time to time, to correctly perform its function (although defects are not heritable). In cases where such failures occur at high rates, it is evolutionarily advantageous if this function is compensated for by a redundant duplicate. Note that the first two models consider conservation of redundancy without assuming any evolutionary advantage for the redundant state. Moreover, in none of the above models is it assumed that organisms have specifically evolved ways to use existing functional overlap.

An alternative to Nowak's approach is that redundancy itself constitutes a selective advantage. Specifically, we propose that in certain systems, regulatory connections have evolved to exploit functional overlaps resulting in evolutionarily advantageous functions. Perhaps one such function could be the use of crossregulated redundant duplicates to downplay stochastic noise at the protein level (Figure 1A). In such cases, compensation for gene loss may be merely a side effect of mechanisms that use functional redundancy in the wild-type organism. It should be noted that we do not consider noise regulation as an exclusive possibility but rather as an example of a functional overlap that confers a selective advantage on the wild-type organism.

Weak Linkages, Robustness, and Evolvability

Living cells consist of a highly stochastic internal environment with randomly fluctuating concentrations of proteins and regulators. In addition to this variation, allelic variation generates unpredictable and unique genomic backgrounds. Within this framework, cells need to evolve means to produce precisely fine-tuned responses to a wide variety of signals. In other words, molecular evolution has been challenged with the requirement to invent highly accurate signaling systems that are built with randomly fluctuating components.

Inevitably, meeting this challenge requires weak regulatory associations, that is, weak linkages (Kirschner and Gerhart, 2005), between signaling components to minimize noise propagation together with tight associations between signals and responses. It is plausible that functional redundancy may be exploited as a tool to reconcile these conflicting requirements. It is, for example, easy to imagine how a randomly fluctuating crucial regulator could be reciprocally coupled with a redundant partner (Figure 1A). In such cases, linkages between the regulating protein and its target are weak as the target responds equally to two different (redundant) protein regulators, both carrying the same message through noisy channels. In contrast, linkages between the signal and response, that is, the message delivered by the redundant pair, are tight as internal fluctuations are buffered by the redundancy within the system.

Genetic redundancy has long been viewed as being the fuel of evolutionary change by decoupling genotypic information from its phenotypic expression and consequently from selection. Yet,

redundancy is expected to "burn out" in the process of evolutionary change. Here, we have argued that redundancy, in certain contexts, may be evolutionarily stable. In light of this, we now ask whether such "conserved redundancies" may also accelerate and facilitate evolutionary change. From the preceding discussion, it is not difficult to envision how conserved redundancies may facilitate weak regulatory linkages that buffer against the cells' internal stochasticity. We would like to suggest that in the context of cellular regulation, genetic variability is fundamentally no different than the stochastic fluctuations of proteins. In fact, the two processes may differ primarily in their time scales, the former representing fluctuations with a lifetime equal to or longer than the lifetime of the individual. In this sense, evolvability and robustness simply represent two sides of the same coin and may both be facilitated by functional redundancy.

Conclusion

Validation and identification of a selective advantage for the redundancy of gene duplicates may be difficult as it requires untangling the evolutionary benefits of the pair's redundant attributes from their nonredundant (divergent) attributes. Nevertheless, two lines of evidence could indicate the direct benefit of existing redundancy on a function: first, the evolutionary conservation of the functional overlap and, second, the nontrivial regulatory design that uses it. The focus of recent systematic studies of single and double gene deletions on gene duplicates originating from an ancient whole-genome duplication, together with several other studies (Kafri et al., 2008; Tischler et al., 2006), provides evidence for strong conservation of redundancy among certain gene duplicates. Notably, it has been previously recognized that duplicates originating from ancient whole-genome duplications tend to maintain overlapping roles (Conant and Wolfe, 2008). The dissimilar expression patterns of these redundant gene pairs together with their measured capacity to compensate for each other's loss suggest that they may be involved in responsive circuitry. Together these data strongly implicate redundancy as a conserved design of genetic networks. Furthermore, from examination of the functions of redundant genes it is becoming clear that redundant partners are often associated with proteins involved in signaling and posttranslational protein modifications such as phosphorylation and ubiquitination. These findings illuminate redundancy as a central yet unexplored component of cellular signaling and call for new studies to test how functional overlap facilitates these functions.

ACKNOWLEDGMENTS

Y.P. thanks the European Research Council for support and the Department of Systems Biology, Harvard Medical School for their hospitality. R.K. thanks the Human Frontiers Science Program for funding.

REFERENCES

Conant, G.C., and Wolfe, K.H. (2007). Mol. Syst. Biol. *3*, 129.

Conant, G.C., and Wolfe, K.H. (2008). Nat. Rev. Genet. *9*, 938–950.

Dean, E.J., Davis, J.C., Davis, R.W., and Petrov, D.A. (2008). PLoS Genet. *4*, e1000113 10.1371/journal.pgen.1000113.

DeLuna, A., Vetsigian, K., Shoresh, N., Hegreness, M., Colon-Gonzalez, M., Chao, S., and Kishony, R. (2008). Nat. Genet. *40*, 676–681.

Haldar, M., Karan, G., Tvrdik, P., and Capecchi, M.R. (2008). Dev. Cell *14*, 437–445.

Kafri, R., Levy, M., and Pilpel, Y. (2006). Proc. Natl. Acad. Sci. USA *103*, 11653–11658.

Kafri, R., Dahan, O., Levy, J., and Pilpel, Y. (2008). Proc. Natl. Acad. Sci. USA *105*, 1243–1248.

Kirschner, M., and Gerhart, J. (2005). The Plausibility of Life: Resolving Darwin's Dilemma (New Haven, CT: Yale University Press).

Lancet, D., Sadovsky, E., and Seidemann, E. (1993). Proc. Natl. Acad. Sci. USA *90*, 3715–3719.

Musso, G., Costanzo, M., Huangfu, M., Smith, A.M., Paw, J., San Luis, B.J., Boone, C., Giaever, G., Nislow, C., Emili, A., and Zhang, Z. (2008). Genome Res. *18*, 1092–1099.

Nowak, M.A., Boerlijst, M.C., Cooke, J., and Smith, J.M. (1997). Nature *388*, 167–171.

Ozcan, S., and Johnston, M. (1999). Microbiol. Mol. Biol. Rev. *63*, 554–569.

Papp, B., Pal, C., and Hurst, L.D. (2004). Nature *429*, 661–664.

Putnam, N.H., Butts, T., Ferrier, D.E., Furlong, R.F., Hellsten, U., Kawashima, T., Robinson-Rechavi, M., Shoguchi, E., Terry, A., Yu, J.K., et al. (2008). Nature *453*, 1064–1071.

Tischler, J., Lehner, B., Chen, N., and Fraser, A.G. (2006). Mol. Biol. Evol. *20*, 772–774.