Genetic Redundancy

Darwin's theory of evolution and subsequently population genetics have stressed the role that natural selection plays in the fixation of beneficial mutations and the elimination of deleterious ones. By contrast, the neutral theory shows what level of genetic diversity we might expect when a mutant gene is selectively equivalent to a wild type. These mathematical theories provide the logical foundation on which our understanding of nature rests; they give us insight into how adaptations could have developed through the fixation of innumerable mutations, each of small effect. Nonetheless, in genetic systems there is a great deal of diversity that is neither immediately beneficial nor always neutral yet has been preserved over long periods of time. Such genetic information is placed in a category of its own: genetically redundant information.

The term "redundancy" is somewhat problematic because it might lead one to assume that this information has no adaptive value in any circumstance. To appreciate the value of redundancy, however, one need only consider devices engineered by humans. Modern aircraft, for example, are controlled by sophisticated avionic circuits. Only a single system is required to fly the aircraft, and because each system is costly to the manufacturer, one might suppose that each airplane is equipped with only a single copy. In fact, most aircraft have four replicas of this circuitry, so that if one or more fails, there will be a backup to take its place. Redundancy ensures the robustness of the aircraft in the face of damage to the system. We can tell that a system is redundant precisely because damage leaves it functioning exactly as it did before the damage occurred. Therefore, the redundancy is not neutral, because it has a value when its duplicate is removed; nor is it immediately advantageous, because, in the absence of damage, duplicates can be removed without adverse consequences.

What is true of engineered devices often is also true of evolved devices. Thus, it comes as no surprise to find extensive redundancy among the genomes of living organisms. The best-characterized example of redundancy in genetic systems (apart from the redundancy of the genetic code itself) is polyploidy. In diploid genomes, for example, there are at least two copies of each autosomal allele. Diploidy has been explained as an adaptation to minimize the impact of deleterious recessive mutations: the second allele can mask the negative effects of mutation. A further characteristic of diploidy is haplo-insufficiency on the loss of an allele, or the reduction in fitness that results from that loss.
This article focuses on redundancy among independent genes rather than on polyploidy, since the former type of redundancy is harder to understand. Following the removal, or “knockout,” of both alleles of a gene, there often is no detectable, or “scoreable,” change in the phenotype or fitness of the organism. For example, knockout of the *Drosophila* genes *gooseberry* and *sloppy paired*, the mouse genes *tanascin* and *Hox*, the *Arabidopsis* gene *far*, or the yeast gene *myosin* do not result in a quantifiable change in phenotype under laboratory conditions. Therefore, these genes are thought to possess a degree of redundancy similar to that of avionics systems.

From the perspective of evolution, redundancy presents a series of fascinating questions. How does redundancy evolve in the first place? Once it has evolved, how is it maintained? Is redundancy simply a matter of preserving duplicates of a given gene, or does it involve the restructuring of whole genetic networks? What fraction of the genome is redundant, and how does this fraction relate to factors that influence the rate of mutation? What are the implications of large-scale redundancy for the rate of evolution of a species? Each of these questions is addressed in the sections that follow.

**Evolution and Maintenance of Redundancy**

The two principal theories describing the origin of redundancy are functional shift and gene duplication. Functional shift is the phenomenon in which two independent genes evolve toward some degree of overlap in their function, or toward a third, novel function. In contrast, gene duplication occurs when a single gene duplicates, resulting in the presence of two identical genes at different sites in the genome. For a population of individuals to possess redundant genes, the duplicate must spread to the point of fixation. This can happen even when there is a probability of both copies of the original gene being lost through mutation, as long as the rate of recurrent duplication is on the same order as the rate of mutation.

Once a degree of redundancy has arisen, it must be maintained. This presents a problem because random inactivation of one gene from a redundant pair will have minimal phenotypic effects, or even none. Several mechanisms for preserving redundancy have been proposed. These theories can be characterized by their focus on cumulative benefit, mutational error buffering, developmental error buffering, pleiotropy, or regulatory elements.

Cumulative benefit theories assume that having an extra active copy of a gene increases the quantity of the product (e.g., a particular protein) that the gene makes. If this increases fitness, genetic redundancy is easy to understand. Thus, each eukaryotic cell harbors multiple copies of the mitochondrial genome, which enables cells to metabolize efficiently, and multiple copies of tRNA and mRNA genes for efficient translation. Eliminating copies will reduce fitness; hence, redundancy is maintained by stabilizing selection acting on the ensemble of identical genes.

Mutational error buffering theories define genetic error buffering as any mechanism that reduces mutational load. Consider two genes that are otherwise identical but have slightly different mutation rates and slightly different efficiencies. The gene with the higher mutation rate will, on average, be eliminated from the population; that is, it becomes a pseudogene. If it is the less efficient gene that also possesses the lower rate of mutation, redundancy can be preserved, because loss of the more efficient gene owing to mutation will bring a relatively large drop in fitness. If the more efficient gene had the lower mutation rate, there would be no selective reason to maintain redundancy, because loss of the less efficient gene would not bring a substantial loss of fitness. Thus, only by allowing a slight asymmetry in each gene’s abilities to perform its respective functions might redundancy be preserved.
Under theories of developmental error buffering, it is observed that during the early growth and development of an organism, many different and strongly interlinked processes must occur. Sometimes random events or environmental variability disturb one or more of these processes. Developmental error buffering denotes mechanisms that can reduce the effects of such perturbations—mechanisms that, in effect, improve the stability of the phenotype. Duplicated or redundant pairs of genes can be maintained if one can take over the lost function when the other becomes defective owing to an event during development. This can occur only if the redundant member of the pair becomes defective at a lower rate than the other member of the pair does. In other words, a function that acts as a developmental buffer must have a low error rate and must be in support of a developmentally unstable function with a high mutation rate.

Some form of pleiotropy can also ensure the conservation of redundant function. The term “pleiotropy” refers to cases in which a single gene affects more than one function and therefore experiences selection in more than one context. Consider two genes with two independent functions. Assume that one gene (the pleiotropic gene) can, in addition to its primary function, also perform the function of the other gene, though less efficiently. Mutations to the pleiotropic gene can eliminate either its unique function, or both functions. Redundancy will be preserved whenever the rate of elimination of the pleiotropic function is lower than elimination of the unique function. In other words, when one gene's pleiotropic function is more robust than the other gene's unique function, the correlated function can be preserved.

Regulatory element theories note that there is redundancy not only among coding regions but also among regulatory elements associated with structural genes. For example, if each of two duplicated genes is accompanied by duplicates of subsets of regulatory elements (where each subset overlaps to some degree through shared elements), the redundant genes can be maintained by selection that acts through their unique patterns of expression and the shared regulatory element. The shared regulatory element controls the correlated function. If one assumes that the shared regulatory element is a smaller mutational target than the coding region, this will prolong the half-life of the redundant function. It is not sufficient, however, to prevent one or more shared elements from becoming silenced in the long term. To preserve redundancy indefinitely requires, as in the pleiotropic model, some asymmetry in mutation and/or in efficacy of the regulators.

Is Redundancy Simple Preservation or Large-scale Restructuring?

Genetic redundancy can be the direct consequence of carrying duplicate copies of a gene, or the outcome of nonlinear interactions among many genes that are coordinately regulated. As in an ecosystem where predators are able to switch to other prey when their favored resources becomes limiting, so genes can modify their pattern of regulation when some target genes are removed or damaged. In this case, redundancy arises because each gene interacts with numerous others according to the abundance of all concerned. Redundancy is not a property of individual genes, but of a network.

In an effort to determine which type of redundancy is more prevalent in yeast, A. Wagner searched for correlations between fitness effects and genetic similarity in yeast knockout mutants. If gene duplication accounts for redundancy, then we expect that after knockout of a gene, the degree of fitness reduction will correlate positively with the genetic divergence between the target of the knockout and its duplicate. We expect genes that are more similar to compensate more effectively for the loss of their corresponding paired genes. This, however, is not what we find. There is no significant correlation between genetic divergence and the magnitude of the protective effect on a knockout. There are, however, many knockouts that have a minimal effect on fitness, even though the removed genes have a very weak similarity to other genes in the genome. This lends credibility to the hypothesis that redundancy emerges as a network effect.
Proportion of Redundancy and Relation to Factors Influencing Mutation Rates

What fraction of the genome is redundant, and how does this fraction relate to factors that influence the rate of mutation? These questions are largely unanswerable at present. In most organisms that have been sequenced so far, at least one-third of the genes in the genome belong to paralogous gene families—that is, they share homology through duplication. Assuming that redundancy exists wherever there is strong genetic similarity, this might lead us to expect very high levels of redundancy; however, we know from yeast studies that this is not always the case. Furthermore, selection for redundancy may have more to do with shoring up the genome against errors than with selection for redundancy in itself. This is likely because robustness in resistance to errors does not directly increase the reproductive fitness of the redundant genome, but it enables populations with high redundancy to outlast those with little redundancy.

Redundancy has been documented best among eukaryotes—organisms with large genomes and low mutation rates. Eukaryotes are thought to have undergone one or more whole-genome duplications during their evolutionary history. Mutation rates vary from a low of around $8.4 \times 10^{-11}$ per base pair per replication in *Drosophila* to highs of $4.0 \times 10^{-10}$ in *E. coli* and $2.4 \times 10^{-8}$ in bacteriaphage lambda.

Interestingly, higher mutation rates (like those observed in viruses and bacteria) are not associated with higher levels of redundancy; if anything, the reverse appears to be true. Why this should be so is at present controversial. Many viruses and bacteria have “overlapping genes,” stretches of DNA (or RNA) that encode more than one protein. This is the opposite of redundancy, because a single mutation is likely to affect two or more proteins. Through mechanisms that are not entirely understood, overlapping genes may actually make populations of viruses in the host more stable. Alternatively, viruses may have been selected for very small genomes and high replication rates, making redundancy disadvantageous.

Implications of Redundancy for Rates of Evolution

An important consequence of redundancy is the reduction of genetically encoded variance among phenotypes within a population. This occurs because redundancy tends to obscure the effects of mutations on phenotypes. We know from Fisher's fundamental theorem that selection is more effective when there is greater heritable variance among phenotypes, so we expect that redundancy will reduce the rate of adaptive evolution.

To understand this relationship, consider the consequences of complete redundancy. Adaptive evolution would come to a halt because no mutation would ever produce any change in phenotype, and the rate of fixation of mutations would then be determined exclusively by the rate of mutation, as is predicted by the neutral theory. There are circumstances, however, when redundancy may promote adaptive evolution—particularly in cases where local deviations from the wild type phenotype are deleterious. The effect of partial redundancy is to allow a sufficient quantity of genetic polymorphism to arise by neutral drift, so that new, nonredundant mutations can fall within the vicinity of the second fitness peak. Thus, the phenotype is able to “hop” from one fitness peak directly to another without having to go through intermediate phenotypes of lesser viability. Redundancy could provide for more effective exploration of genetic space. Neither of these two potential consequences of redundancy has yet been tested empirically.

Bibliography

Find this resource:


Find this resource:

DAVID C. KRAKAUER

Copyright © 2019. All rights reserved.