

Thinking about genetic redundancy

JAMES H. THOMAS

Partial functional redundancy among genes is frequently observed in a wide range of organisms and processes, but the selective value of such redundancy is not immediately apparent. Any fully redundant function should be evolutionarily unstable: unless selection acts to maintain the redundancy it will tend to be lost by mutational drift. I discuss four possible mechanisms by which selection might act to maintain genetic redundancy.

I follow common usage in defining two genes as being redundant if each can partially or fully substitute for the function of the other. Genetic criteria that establish genetic redundancy experimentally are discussed in Box 1. Even if only clearly established cases are considered, genetic redundancy appears to be a common phenomenon and thus warrants careful thought. It is perhaps surprising that redundancy is so prevalent, since it is not immediately obvious what selective advantage it might confer. Possession of two fully redundant genes should, on evolutionary time scales, be an unstable condition. Each gene would inevitably suffer mutation and become nonfunctional. In the absence of significant selection for both genes, a nonfunctional allele would probably become fixed in the population by chance and the redundancy would be lost. Although this scenario applies only where the two genes are perfectly redundant and there is no selective advantage associated with having both genes, a similar argument might suggest that even partially redundant genes would tend to lose their redundancy. In this case, and in the absence of specific selection, genetic drift would tend to reduce the degree of functional overlap between the two genes. These thoughts prompt a careful consideration of mechanisms by which natural selection might act to maintain genetic redundancy.

I will consider four classes of explanation for the existence of redundant genes or genetic processes, shown schematically in Fig. 1. First, in redundancy selected for cumulative function, two genes perform the same function but are independently selected on the basis of enhanced efficiency or speed. Second, in redundancy selected for fidelity, two genes perform similar or distinct functions that enhance the fidelity of some process. Selection acts on the enhanced fidelity achieved by the combined activity of the two genes. Third, in redundancy selected for divergent functions,

two genes have overlapping functions, but are selected on the basis of distinct functions unique to each; selection acts only on functions peculiar to each gene, but serves indirectly to maintain the functional overlap. Fourth, in redundancy selected for emergent function, two genes perform closely related functions, but presence of both functions results in the appearance of an emergent property that neither gene alone can perform; selection for this emergent property accounts for the maintenance of redundancy. Finally, in addition to these four models, I consider chance redundancy, which arises from an evolutionarily recent gene duplication where neither copy of the gene has yet drifted to become nonfunctional. I also discuss how these explanations of genetic redundancy can be applied to redundancy at the levels of the cell and of tissues.

Redundancy selected for cumulative function

In redundancy selected for cumulative function, each redundant gene has the same function and each is selected for its cumulative contribution to the quantity of some gene product. An example of this type of

Box 1. A genetic test for redundancy

A fairly rigorous genetic test for functional redundancy between two genes is to determine whether the double mutant has additional mutant defects when compared with each single mutant⁴⁹ (see Figure). In a number of cases, loss-of-function mutations in either of two genes have little or no effect on some specific process, while the double mutant is strongly deficient in that process. There are many clearly demonstrated examples of such phenotypic synergy (see, for example, Refs 21, 22, 25, 28, 35, 40). Some caution is warranted when interpreting such interactions. First, if the mutations analysed cause a gain of function, the synergy may not reflect the function of the normal gene product. Second, partial loss-of-function mutations in two genes may show phenotypic synergy that has nothing to do with redundancy. Third, the process under study should be well defined and it should be clear that the genes have an unusual and specific synergy for that process. The often-observed phenomenon that two sick mutants get sicker when combined is difficult to interpret conclusively. In the absence of specific phenotypic synergy, genetic redundancy has often been invoked as an explanation when null mutations in a gene have no discernible phenotype (see, for example, Refs 41–44) or one that appears less severe than expected (see, for example, Refs 45–48). In such cases, functional redundancy is a sound explanation of the results, but an alternative possibility is that the gene performs a non-redundant function that has not been discerned. Such a function might be subtle and yet confer a selective advantage on animals carrying the normal gene. Alternatively, the function could be critical, but evident only under some other environmental condition, in some unassessed (or even unknown) part of the life cycle or in some specific assay not usually performed.

Genotype	Phenotype
a^-	defect A
b^-	defect B
$a^- b^-$	additional new defect

Genetic test for functional redundancy. The phenotypic consequences of eliminating the function of gene a or b, or both. The $a^- b^-$ double mutant has a novel defect in addition to the phenotypes of each single mutant.

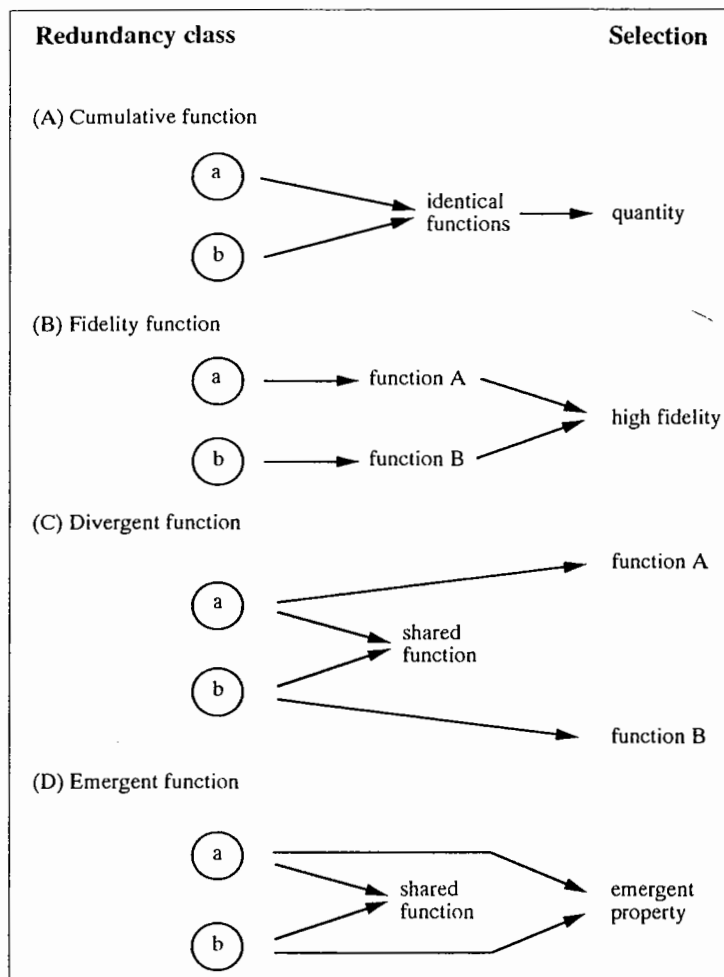


FIG 1

Mechanisms of selection for genetic redundancy. Redundant genes are symbolized by encircled lower-case letters. Arrows indicate both the redundant and individual functions of each gene; individual functions are to the left, and the function on which selection acts is to the right.

redundancy is found in the ribosomal RNA genes. These genes are present in multiple identical copies in all eukaryotes; this appears to be because maximal transcription from one copy of the gene cannot produce ribosomal RNA at a sufficiently high rate to support optimal growth¹. Another such example that reflects a requirement for high levels of gene product is seen in the chorion genes of insects. In *Drosophila*, two clusters of chorion genes are specifically amplified in ovarian follicle cells when large amounts of their gene products are required². It is presumed that this gene amplification, which transiently produces many redundant sets of chorion genes, is selected for production of the large amounts of chorion protein required in the eggshell. Other examples that probably represent this sort of redundancy include the histone genes of many organisms, the cuticle collagen genes and muscle actin genes of *Caenorhabditis elegans* and a variety of other proteins that are rapidly synthesized at high levels. This sort of redundancy abounds at the cellular level. In most large organisms, a single cell of a given type is rarely large enough to perform a required function. Instead, between hundreds and millions of essentially

identical copies of a single cell type are used to achieve the functional volume required.

Redundancy selected for fidelity

In redundancy selected for enhanced fidelity, genes have an individual function that may be either similar or distinct, but together they promote high fidelity for some process. An example of redundancy that is probably selected for fidelity is seen in the G1 cyclin genes of yeast. *CLN1*, *CLN2* and *CLN3* form a functionally redundant family of genes that are required for progression of the cell cycle from G1 to S phase³. Loss of function of *CLN1* or *CLN2* has little or no phenotypic effect⁴, while loss of *CLN3* causes a slight delay in entry into S phase⁵. Cells that are doubly mutant for certain combinations of these genes have a more severe cell-cycle delay than the single mutants^{4,6,7}, and triply mutant cells are permanently arrested in G1 (Ref. 4). The close functional similarity of these genes and the graded defects in cell-cycle timing seen in the single, double and triple mutants have led to the suggestion that together these genes promote robust, high-fidelity control of the cell cycle⁸.

This type of redundancy may also be found in the yeast genes *CIN8* and *KIP1*, which encode members of the kinesin family of microtubule-based motor proteins^{9,10}. Inactivation of both genes causes a fatal defect of the mitotic spindle. Inactivation of either gene alone has no obvious effect on cells grown at 26°C or below, indicating that the two genes are redundant for mitotic spindle function at low temperatures. Deletion of *CIN8* reduces the highest viable growth temperature; however, on closer analysis, deletion of *CIN8* can be seen to substantially reduce the fidelity of chromosome segregation, even during growth at low temperatures. *CIN8* and *KIP1* may perform closely related functions that together promote the desired high fidelity of chromosome segregation¹¹.

Redundancy that promotes fidelity is probably pervasive in a variety of high-fidelity processes, such as DNA replication and transcription, RNA splicing and translation, chromosome segregation, control of the cell cycle and development. Genes that act to prevent cancer are a particularly interesting case in which intense selection for high fidelity must exist. In large metazoan animals, the very large number of cells present in the mature animal demands correspondingly low frequencies of somatic mutation and of the other events that underlie the formation of malignant tumors¹²⁻¹⁴. This may select for extensive redundancy among genes that function either to prevent the accumulation of such cancer-causing events or to forestall their deleterious consequences. For example, many cancers arise by mutation of more than one gene¹⁵⁻¹⁷. Such mutations may affect functionally redundant pathways, each of which can restrain cell proliferation. At the level of the cell or tissue, redundancy selected for fidelity may also be common. For example, regulative cell interactions may

promote the required high fidelity of development. In the central nervous system, extensive neuronal redundancy may be required to produce high-fidelity command of language, arithmetic and other complex tasks. A simpler example of redundancy in the nervous system is chemotaxis in *Caenorhabditis elegans*, in which response to specific chemicals is mediated by a group of partially redundant sensory neurons¹⁸.

The concept of redundancy selected for fidelity is similar to that of 'epigenetic load' proposed by Nasmyth *et al.*¹⁹ As they have pointed out, redundancy that promotes fidelity may apply not only to normal growth conditions but also to variant conditions, thus permitting a process to continue as normal under a wider range of environmental conditions than would otherwise be possible.

Redundancy selected for divergent functions

In redundancy selected for divergent functions, two genes are individually important for their capacity to perform distinct functions. However, the functions of the genes overlap, or intersect, in some respect. The genes are redundant for this intersecting function, but they are selected for their related but divergent functions. A number of examples that probably represent this sort of redundancy have been described, although the strength of the evidence varies²⁰⁻²⁵. I will describe two particularly clear instances. In budding yeast, two functionally redundant genes, *KSS1* and *FUS3*, which encode serine/threonine kinases, have been identified. Disruption of *KSS1* causes slow growth and defective adaptation to mating factor²⁶. Disruption of *FUS3* blocks cell-cycle arrest in response to mating pheromone but leaves other responses intact, causing a partial defect in mating²⁷. Disruption of both *KSS1* and *FUS3* causes a strong mating defect, blocking all known responses to mating pheromone²⁸. The *FUS3* and *KSS1* proteins have considerable sequence similarity to one another and to the vertebrate family of MAP kinases²⁹. Their redundancy suggests that both gene products can activate the transcription factor *STE12*, an event required for mating responses. In contrast, only *FUS3* can phosphorylate *FAR1*, a protein responsible for linking mating pheromone response to cell-cycle arrest³⁰. This probably accounts for the defective cell-cycle arrest seen in cells that are mutant for *FUS3* alone. The exact cause of the defect seen in cells that are mutant for *KSS1* is less clear, but appears to involve the product of the *WHI1* gene, a protein homologous to cyclins²⁶. The relationship between *KSS1* and *FUS3* is a good example of redundancy selected for divergent functions, because both their distinct and shared functions have been identified. Presumably, each gene is individually selected on the basis of its non-redundant function in the cell cycle (*KSS1*) or in response to mating pheromone (*FUS3*).

This type of redundancy is also seen with the two cell-surface receptors encoded by *lin-12* and *glp-1*, which control developmental cell interactions in *C. elegans*³⁰. While loss of either *lin-12* or *glp-1* function causes failure of two separate specific sets of cell interactions³¹⁻³³, loss of function of both genes results in failure of additional cell interactions that rarely or never fail in either single mutant^{34,35}. *lin-12* and *glp-1* must be functionally redundant for these additional cell

interactions, but functionally distinct for the other cell interactions. This is particularly interesting from an evolutionary perspective: *lin-12* and *glp-1* lie within a few kilobases of one another and their amino acid sequences are over 50% identical³¹. It seems likely that the two genes arose by duplication and subsequent divergence of a single precursor gene, and that during their divergence each acquired distinct functions while retaining some functional overlap with the other gene.

It is interesting to speculate why, in the course of evolution, genes encoding overlapping redundant functions have not drifted to the point where they are no longer redundant. Obviously, in some instances, insufficient time will have elapsed since the appearance of the redundancy to allow such drift to become complete. However, there is an alternative explanation for why such redundancies may be quite stable even over very long periods. I propose that, where the divergent functions of the redundant genes are closely related, selection for those divergent functions will often suffice to maintain functional overlap between the two genes. For example, with the yeast kinases *KSS1* and *FUS3*, each is individually selected to respond to related signals, to catalyse the kinase reaction and to maintain an appropriate kinase target specificity. Since these individual properties of the proteins are closely related to their redundant functions, an unavoidable consequence of their common selected features may be that the genes tended to retain some functional overlap.

Redundancy selected for an emergent property

In genetic redundancy selected for an emergent property, two genes perform similar functions and each alone can confer a wild-type phenotype for that function. However, the presence of two such functions results in an emergent or synthetic property. The emergent property is logically distinct from the individual functions of the redundant processes, but is dependent on the existence of two or more such processes. This idea may best be illustrated by binocular vision. In primates, each of the two eyes performs redundant functions in detecting color, shape, movement, etc. However, an individual eye is largely incapable of direct depth perception. Binocular depth perception is thus an emergent property of our visual system that is dependent on having two fully functional and largely redundant eyes. Selection for two eyes may also act at other levels, for example, on the enhanced visual acuity afforded by having two images to compare in the brain (redundancy selected for fidelity). However, even if no other selection were involved, the emergent property of depth perception is itself an important selectable trait. A genetic example of such an emergent property is color vision³⁶. Color is perceived by comparing the activation of three distinct cone-pigment proteins. The function and molecular structure of each cone pigment is very similar, but each has a slightly different absorbance spectrum. Numerous fine color gradations can be detected by comparing the relative activation of the three pigments. Under conditions of bright light, vision is mediated only by cones, yet genetic loss of any one cone pigment results in vision that is virtually normal, except for impaired color discrimination. Thus, the three cone-pigment genes are functionally redundant

for all aspects of vision except color perception; it seems that the existence of three distinct pigments is selected on the basis of the emergent property of color vision.

Emergent properties that arise from genetic redundancy may be more common than we realize, since their detection may often require specific consideration of the emergent property. Other genetic examples might include two genes encoding similar products that are spatially segregated within a cell, thus allowing the collection of spatial information. Alternatively, similar gene products might have overlapping ligand-binding specificities or affinities, permitting emergent inference of ligand type or amount (odorant receptors may have this property³⁷). Emergent properties may also be common at the levels of the cell and the organ. Tactile spatial information is gathered from large numbers of virtually identical mechanosensory neurons scattered across the skin of vertebrates. A simple form of tactile information is seen in response to touch stimulation in *C. elegans*, in which genetically identical sensory cells produce forward or reverse movement, dependent only on the position of the stimulated touch-responsive neuron^{38,39}.

Concluding remarks

Obviously, selection for genetic redundancy may not always fall neatly into one of these four proposed classes. To take a simple example, some high-fidelity processes may require quantities of gene products that cannot be produced by a single gene; in such cases, selection for cumulative function will also promote fidelity. More generally, any given case of genetic redundancy (including some of those discussed above) may involve more than one of the explanations suggested; often, it will be difficult to determine which of these types of selection best explains a given case. However, careful consideration of the possibilities is a good first step toward such a goal.

Finally, two genes may be perfectly redundant and the redundancy may not yet have been lost by genetic drift. This situation may have arisen by a gene duplication that, by chance, became fixed in a population. If the duplication arose recently enough, both copies of the gene might remain fully functional. The frequency of such perfect redundancy will depend on the rate of unselected gene duplication relative to that of unselected mutation to non-functionality. Since these rates, especially that of gene duplication, are not easily determined, it is difficult to assess how frequently this might occur. Nevertheless, the rarity of large near-perfect duplications of protein-coding DNA in most organisms argues that such perfect redundancy is uncommon. Another way that perfectly redundant genes might have arisen is by the action of some historical selection that is no longer relevant¹¹; however, the prevalence of this mechanism may be even more difficult to assess.

Acknowledgements

I thank L. Avery, L. Hartwell and I. Herskowitz for contributing important ideas to this review. I also thank L. Avery, C. Berg, D. Birnby, L. Hartwell, I. Herskowitz, E. Malone, A. Paulovitch, D. Reiner, W. Schackwitz, F. Tax and J. Vowels for critical reading of the text.

References

- 1 Long, E.O. and Dawid, I.B. (1980) *Annu. Rev. Biochem.* 49, 727-764
- 2 Spradling, A.C. (1981) *Cell* 27, 193-201
- 3 Richardson, H.E., Wittenberg, C., Cross, F. and Reed, S.I. (1989) *Cell* 59, 1127-1133
- 4 Hadwiger, J. et al. (1989) *Proc. Natl Acad. Sci. USA* 86, 6255-6259
- 5 Cross, F. (1988) *Mol. Cell. Biol.* 8, 4675-4684
- 6 Lew, D.J., Marini, N.J. and Reed, S.I. (1992) *Cell* 69, 317-327
- 7 Linskens, M., Tyers, M. and Futcher, B. (1993) *Cell* 72, 487-489
- 8 Nasmyth, K. and Dirick, L. (1991) *Cell* 66, 995-1013
- 9 Hoyt, M.A., He, L., Loo, K.K. and Saunders, W.S. (1992) *J. Cell Biol.* 118, 109-120
- 10 Roof, D.M., Meluh, P.B. and Rose, M.D. (1992) *J. Cell Biol.* 118, 95-108
- 11 Goldstein, L.S.B. (1993) *J. Cell Biol.* 120, 1-3
- 12 Armitage, P. and Doll, R. (1954) *Br. J. Cancer* 8, 1-12
- 13 Cairns, J. (1975) *Nature* 200, 197-200
- 14 Nowell, P.C. (1975) *Science* 194, 23-28
- 15 Land, H., Parada, L.F. and Weinberg, R.A. (1983) *Nature* 304, 596-602
- 16 Sinn, E. et al. (1987) *Cell* 49, 465-475
- 17 Hansen, M.F. and Cavenee, W.K. (1988) *Trends Genet.* 4, 125-128
- 18 Bargmann, C.I. and Horvitz, H.R. (1991) *Neuron* 7, 729-742
- 19 Nasmyth, K. et al. (1991) *Cold Spring Harbor Symp. Quant. Biol.* 61, 9-20
- 20 Breeden, L. and Nasmyth, K. (1987) *Cell* 48, 389-397
- 21 Lamphier, M.S. and Ptashne, M. (1992) *Proc. Natl Acad. Sci. USA* 89, 5922-5926
- 22 Higashijima, S., Michiue, T., Emori, Y. and Saigo, K. (1992) *Genes Dev.* 6, 1005-1018
- 23 Culotti, J.G., von Ehrenstein, G., Culotti, M.R. and Russell, R. (1981) *Genetics* 97, 281-305
- 24 Burbulys, D., Trach, K.A. and Hoch, J.A. (1991) *Cell* 64, 545-552
- 25 Thomas, J.H., Birnby, D.A. and Vowels, J.J. (1993) *Genetics* 134, 1105-1117
- 26 Courchesne, W.E., Kunsawa, R. and Thomer, J. (1989) *Cell* 58, 1107-1119
- 27 Elion, E.A., Grisafi, P.L. and Fink, G.R. (1990) *Cell* 60, 649-664
- 28 Elion, E.A., Brill, J.A. and Fink, G.R. (1990) *Cold Spring Harbor Symp. Quant. Biol.* 61, 41-49
- 29 Boulton, T.G. et al. (1990) *Science* 249, 64-67
- 30 Yochem, J. and Greenwald, I. (1989) *Cell* 58, 553-563
- 31 Greenwald, I., Sternberg, P. and Horvitz, H.R. (1983) *Cell* 34, 435-444
- 32 Priess, J., Schnabel, H. and Schnabel, R. (1987) *Cell* 51, 601-611
- 33 Austin, J. and Kimble, J. (1987) *Cell* 51, 589-599
- 34 Lambie, E.J. and Kimble, J. (1991) *Development* 112, 231-240
- 35 Bowerman, B., Tax, F.E., Thomas, J.H. and Priess, J.R. (1992) *Development* 116, 1113-1122
- 36 Nathans, J. et al. (1992) *Annu. Rev. Genet.* 26, 403-424
- 37 Buck, L. and Axel, R. (1991) *Cell* 65, 175-187
- 38 Chalfie, M. and Au, M. (1989) *Science* 243, 1027-1033
- 39 Chalfie, M., Thomson, J.N. and Sulston, J.E. (1983) *Science* 221, 61-63
- 40 Rabin, R.S. and Stewart, V. (1992) *Proc. Natl Acad. Sci. USA* 89, 8419-8423
- 41 Greenwald, I. and Horvitz, H.R. (1980) *Genetics* 96, 147-164
- 42 Pereira, A., Doshen, J., Tanaka, E. and Goldstein, L.S.B. (1992) *J. Cell Biol.* 116, 377-383

REVIEWS

- 43 Smith, S.A., Kumar, P., Johnston, I. and Rosamond, J. (1992) *Mol. Gen. Genet.* 235, 285-291
- 44 Zhuang, Y. *et al.* (1992) *Proc. Natl Acad. Sci. USA* 89, 12132-12136
- 45 Kim, S.K. and Horvitz, H.R. (1990) *Genes Dev.* 4, 357-371
- 46 Joyner, A.L. *et al.* (1991) *Sciences* 251, 1239-1243
- 47 Soriano, P., Montgomery, C., Geske, R. and Bradley, A. (1991) *Cell* 64, 693-702
- 48 Roulier, E.M., Fyrberg, C. and Fyrberg, E. (1992) *J. Cell Biol.* 116, 911-922
- 49 Guarente, L. (1993) *Trends Genet.* 9, 362-366

J.H. THOMAS IS IN THE DEPARTMENT OF GENETICS, UNIVERSITY OF WASHINGTON, SEATTLE, WA 98195, USA.