Science & Society

EMBO reports

Understanding redundancy and resilience

Redundancy in life is provided by distributing functions across networks rather than back-up systems

Philip Hunter*

Redundancy is fundamental for all organisms to cope with environmental stress and harmful mutations. It plays a vital role in all key processes from genetics to development, immunity, nervous systems, musculoskeletal systems and visual processing. Understanding of how redundancy works is also highly relevant for developing novel therapeutics to treat genetic disorders by bypassing dysfunctional alleles or for cancer therapy to overcome tumour cells' resistance mechanisms. Redundancy is also a critical feature for synthetic biology's applications in medicine, manufacturing or agriculture to ensure that the systems are resilient against external or internal disruptions.

As a field of study, redundancy dates back around a century to first observations of gene duplication in *Drosophila*. The genomic revolution opened a new chapter with the ability to knockout genes and directly observe the impact on an organism's phenotype and fitness. This led to the identification of gene pairs that appear to be mutually redundant to some degree, and the idea of distributed robustness whereby protection and adaptive ability are conferred by networks rather than individual genes.

First clues from gene pairs

Redundancy rooted in genes can manifest itself in different ways. In the case of redundant gene pairs, either two genes encode the same protein, or two different genes play an equivalent role and can substitute for each other. Redundancy in biology, however, does not usually operate in the way that, say, standby systems do in computation, with one ready to take over in the event of failure. There is a contradiction though: how would a truly redundant gene resist accumulation of deleterious mutations given that these would have no immediate impact on the resulting phenotype? There are indeed examples of gene pairs where mutations only cause damage when both are affected. Such redundant genetic pairs have arisen repeatedly via gene duplication either within a chromosome or across to another chromosome.

> *"Redundancy in biology however does not usually operate in the way that, say, standby systems do in computation, with one ready to take over in the event of failure."*

The first clue came almost a century ago, following analysis of the bar mutation in *Drosophila melanogaster:* some normal parents in *Drosophila* can give birth to male mutants with reduced bar-like eyes, hence the name "bar mutation". A landmark study by Sturtevant and Morgan (1923) found that it results from an unequal crossing over during gametogenesis after which two genes from separate parental chromosomes appeared together in the same chromosome of the offspring. This was finally characterized as a gene duplication effect after another 13 years (Bridges, 1936; Muller, 1936). It had been thought that gene duplication had no immediate consequences under the assumption that the new duplicates remain passive until a subsequent mutation in the original copy disrupts the associated function. Yet, this idea should have been scotched in the light of those successive bar studies in *Drosophila*.

The genomic era

The genomic era then enabled biologists to evaluate the phenotypic impact of knocking out genes known or suspected to have vital functions. When such a knockout had no apparent impact, it was concluded that there must be a back-up, although the situation is usually more complex and nuanced than that. One of the first genes that was identified as seemingly being redundant is the gene for the PrPc prion protein, a glycoprotein located on the surface of cells of the CNS especially. Misfolding of PrPc is associated with neurodegenerative diseases such as transmissible spongiform encephalopathies or Alzheimer's.

The original finding that mice lacking the PrPc protein on their cell surfaces appeared to develop and behave with no outward signs of any problems came out of work to establish the protein's role in these diseases (Büeler *et al*, 1992). At that time, there was little idea about the normal function of PrPc, but the conclusion was that whatever it was there must be some back-up to explain why deletion had no discernible impact on phenotype. After all, PrPc is highly conserved in mammals so it was assumed that it must play a critical role.

Further studies have gradually shed light on the roles that PrPc may have, although there is still considerable uncertainty. A 2014 review noted growing evidence for PrPc involvement in the regulation of cell proliferation and differentiation in various types of mammalian stem cells (Halliez *et al*, 2014). It also reported several studies, suggesting a role for PrPc in cell adhesion, extra-cellular matrix interactions and the cytoskeleton. The conclusion was that

Freelance Journalist, London, UK
*Corresponding author. E-mail: ph@philiphunter.com
Philip Hunter is a freelance journalist in London, UK
DOI 10.15252/embr.202254742 | EMBO Reports (2022) 23: e54742 | Published online 14 February 2022

animals lacking the PrPc gene did overtime suffer adverse consequences, even if these would not be immediately apparent.

Further evidence for the role of prion-related proteins came in 2020 by a French team led by Jean-Luc Vilotte from the University Paris Saclay (Passet *et al*, 2020), who studied the Shadoo protein, which belongs to the prion family and is also evolutionarily conserved. Knockout experiments in mice revealed a significant increase in embryonic lethality during post-implantation stages, as well as growth retardation of young pups and a defect in lactation among females. Vilotte does think though that proteins of the prion family provide some degree of mutual back-up. "We believe that there is some redundancy between PrPc and other related proteins, such as Doppel or Shadoo, and potentially a genetic adaptation of some knockout genotypes", he said. Whereas previous experiments had shown that the presence of the Shadoo protein appeared necessary for normal embryogenesis in the absence of PrPc, a recent preview, on which Vilotte was also an author, indicated that when both were knocked out, the impact on the overall transcriptome was relatively small. The authors suggest that some form of genetic reconfiguration takes place to buffer against the impact of both genes being hobbled by mutation (https://www.biorxiv.org/content/10.1101/2021.10.22.465458v1). Vilotte also suggested that the absence of one or more PrPc-related proteins might only induce a drastic phenotype when the animal was challenged in some way, or when a particular threat arose. He gave an example from another study by a Japanese group reporting that the PrPc protein conferred some protection against influenza A infection in mice (Chida *et al*, 2018).

Redundancy of virulence

Broader insights into redundancy have been obtained through research on bacterial pathogenesis, which led to the identification of five specific types of redundancy that are also widely applicable across life: molecular, target, pathway, cellular process and system redundancy.

Molecular redundancy is broadly understood as covering for the loss-of-function if genes are damaged or lost. Target redundancy is more specific to pathogens, as it involves effector molecules that bind to target proteins in the host cell. Redundancy at this level arises when different effector molecules can elicit the same reaction in the host. This differs from molecular diversity and is more elusive to study because there may be no sequence or structural similarity between the effector molecules, while the functional overlap may only materialize under specific conditions.

Pathway redundancy is common across all life and evolved to protect against perturbations that could disrupt vital processes, such as mutations and environmental changes. In bacterial pathogenesis, it works through effector molecules that are capable of manipulating two or more proteins in the host pathway to provide some resilience against loss of any one effector.

> *"Pathway redundancy is common across all life and evolved to protect against perturbations that could disrupt vital processes, such as mutations and environmental changes."*

Cellular process redundancy is closely related to pathway redundancy but occurs just within the cell. It works through effector molecules that compensate for each other by targeting redundant or complementary host pathways that collectively control a single cellular process. One of the best-known examples is the unfolded protein response (UPR), which keeps the load of unfolded proteins in the cell from accumulating to potentially lethal levels. As it is critical for cell function, the UPR is activated through three distinct sensory pathways to ensure that the cell can respond to multiple signs of stress around the endoplasmic reticulum where protein folding takes place. The relevance for bacterial pathogenesis is that a pathogen can hijack one signalling process while leaving the others intact to ensure that the cell does not break down through accumulation of unfolded protein products.

Finally, system redundancy elevates the process to a higher level, such as the whole cell. One prominent example is cell death, which can be accomplished through apoptosis, necrosis, pyroptosis and autophagy. Apoptosis, or programmed cell death, is the most highly regulated and responsible for killing off around 60 billion cells a day in the average human adult, in response to triggers, internal stress, malfunction or degeneration of some kind. Necrosis is a unpremeditated response to acute cell injury, infection or sudden deterioration. Pyroptosis lies somewhere between these two. It is a short-term response to some highly inflammatory process and often part of the antimicrobial response. Autophagy maintains healthy function of eukaryotic cells by clearing or recycling ageing or damaged components, including mitochondria, but autophagy has also become an adaptive response to stress or disease, either promoting survival of the cell or cell death if that appears a more practical option. These four distinct processes overlap and together maximize the chance that damaged, diseased or compromised cells are destroyed and removed.

How redundancy evolves

A major question is how different mechanisms and levels of redundancy evolved and are maintained in the face of mutations that tend to erode genes that no longer provide an essential function. Plants are a natural source of inspiration given their tendency to gene duplication. Flowering plants or angiosperms have undergone multiple whole-genome duplications during their 200 million years of evolution to the extent that most known genes have at least one copy. Relaxed selection on gene duplicates allows for mutations to accumulate and lead to pseudogenization (function loss). Other times, redundant copies are retained for a long time due to the slow pace of genetic drift in large populations. Another option is that duplicates are retained through selection on novel, adaptive functions. In some cases, genes that originally had duplicate functions evolved separate roles that are each maintained independently through natural selection.

> *"Plants are a natural source of inspiration given their tendency to gene duplication."*

Plants have therefore been widely studied for identifying gene pairs with redundant functions, which have been difficult to find for various reasons, including the challenge of detecting phenotypes associated with random knockouts. Recently, machine

learning has been applied to identify gene pairs with redundancy in plants (Chen *et al*, 2010).

Such work has inspired the development of predictive models of genetic redundancy incorporating omics and mutant phenotype data sets, but the ultimate objective is to equate genetic redundancy with fitness of the organism, for unless it contributes to that, it will not be maintained through natural selection. "Genetic redundancy ultimately needs to be defined on the basis of fitness effects if we are interested in understanding its evolutionary implications", explained Shin-Han Shiu, a plant biologist at the University of Wisconsin-Madison, USA. "The use of multiple definitions is mainly an attempt to approximate fitness. As we gather more fitness data in plants and initiate a similar modelling study in yeast, it becomes clear that what we learned from the multiple definitions capture overlapping but distinct properties compared to defining genetic redundancy based on fitness".

> *"... the ultimate objective is to equate genetic redundancy with fitness of the organism, for unless it contributes to that, it will not be maintained through natural selection."*

Having established a successful model in plants and yeast (Cusack *et al*, 2021), Shiu's aim is now to extend it to animals where gene duplication events are considerably rarer. "A key component determining the degree of redundancy is history of recent duplication", he added. "I would expect animal lineages in general to have a lower degree, with the potential exception of some teleost and amphibian lineages with recent history of polyploidy".

The next steps for Shiu's team, and the field generally, are to expand the scope beyond just gene pairs. First, this means assessing the extent and scope of redundancy and fitness loss for particular gene pairs, taking account of varying environmental conditions. "Then we can ask the question whether some contributing factors are context-specific or general and, perhaps more importantly, to assess globally the impact of the environment on genetic interactions, with genetic redundancy being one of them", Shiu explained. The second aim is to extend the scope to higher orders of genetic redundancy beyond gene pairs, to get closer to the real complexity of genetic interactions in organisms. "After all, genes are in families and there are some degrees of redundancy between any two members", Shiu said. "There are already serious complications when you consider even just three genes in this context, let alone an entire family. If you add environment into the picture, it is quite a challenging problem".

The wider point is that redundancy in biology often occurs at the network level in a flexible way without the need to maintain critical functions. According to Charlie Boone, a specialist in yeast genetics at the University of Toronto, Canada, duplicated genes only retain their redundancy if they acquire some additional role. "If a duplicated gene is retained it is not just sitting there as a back-up in case something fails otherwise it would not be selected for, at least unless something fails quite regularly in a particular environment or developmental program", he explained. However, where duplicated genes have diverged in function, they may still contribute to redundancy at a higher level. This can occur through bypass suppression, a form of genetic rewiring where the phenotypic impact of a mutation is avoided by switching to other genes. Studies by Boone and colleagues in yeast found that 124, or 17%, of 728 essential yeast genes were dispensable in the sense that their loss or mutation could be overcome by bypass suppression (van Leeuwen *et al*, 2020).

As Boone noted, genetic mechanisms driving such suppression are relevant to our understanding of genome architecture and evolution. It may also have relevance for future treatment of human genetic disorders by exploiting the resilience of some patients to disease to identify novel therapeutic interventions.

Efficient use of spare capacity

Such resilience can also be explained in terms of distributed robustness. In nature, it evolved as an inherent part of the system and is maintained on the basis of fitness or selective advantage. One study to predict the robustness of multilayer biological networks confirmed that these likely evolved by avoiding strong degree–degree correlations, which means they avoid the tendency for nodes of the network to form preferential connections with similar ones. This increases overall flexibility and makes the network more resistant to perturbation or disturbance (Mondragón, 2020).

Such research has drawn on findings from technological or data networks, where tolerance against deliberate attack can be improved by incorporating spare capacity large enough to cater for loss of the biggest node in the system. In a network of multiple nodes, such redundancy can be provided more economically than having to duplicate every significant component. Karthik Raman from the Department of Biotechnology, Bhupat & Jyoti Mehta School of Biosciences in India, has compared such "spare capacity" in data networks with concepts of distributed robustness in biology (https://arxiv.org/abs/2109.12358).

> *"... redundancy has sometimes been recruited through evolution to fine tune processes to yield further selective advantage."*

His findings have potential for commercial and medical applications of redundancy insights, with Raman identifying two key categories. "Firstly, there is genetic engineering of plants. To obtain desired traits through engineering, one would need to be able to predict the impact when a gene is manipulated. And an important consideration of that impact is the degree of genetic redundancy with other genes", he explained. "Secondly are medical applications. Genetic disease due to loss-of-function can result from genes with a relatively low degree of genetic redundancy. Thus, there is potential for gene therapy using sequences that can complement the loss-of-function".

Raman added that life uses its spare capacity highly efficiently, such that the system does not need much over-provisioning to provide adequate resilience against mutations and disturbance. At the same time, redundancy has sometimes been recruited through evolution to fine-tune processes to yield further selective advantage. This is the case for musculoskeletal systems, which have a lot of redundancy that does not appear to confer any benefits

in terms of back-up. Instead, redundant structures have been honed for improved coordination at least in humans (Stanev & Moustakas, 2019). At the simplest level, humans have more muscles than are strictly necessary for the required range of movements, and yet fail to provide much back-up, arguing in favour of their role in coordination.

Spatial redundancy in the retina and the visual cortex has also been recruited to improve visual processing. The point is that images in general contain information at varying resolutions. The visual systems of humans and many animals have evolved to exploit this not for back-up, but to focus the high-resolution part of the visual system on relevant areas to obtain more detail, as it provides selective advantage for escaping predators or hunting prey. A recent study demonstrated that a quick initial analysis of low-spatial frequency information guides the slower processing of high spatial frequency detail (Petras *et al*, 2021).

There are hopes of exploiting a better understanding of redundancy in cancer therapeutics now; many drugs that target some apparently critical pathway for tumour cells failed because of evolving redundancy mechanisms. This is especially true for traditional chemotherapy, which increases the rate of genetic and epigenetic alterations that eventually lead to resistance mechanisms in surviving cells (Lavi, 2015). It could also affect the efficiency of immunotherapy to direct the host's T cells against the tumour, but it is not quite clear yet what role redundancy will play there. As a result, a more holistic system perspective has now been widely adopted in cancer research to develop computational approaches that model the dynamic behaviour and evolution of cancer cells. This could assist clinicians in better predicting the success chance and risk of relapse of treatment schemes. After a century of erratic progress, research of biological redundancy is getting ready to progress beyond fundamental research.

References

Bridges CB (1936) The bar "gene" a duplication. *Science* 83: 210 – 211

Büeler H, Fischer M, Lang Z, Bluethmann H, Lipp HP, DeArmond SJ, Prusiner SB, Aguet M, Weissmann C (1992) Normal development and behaviour of mice lacking the neuronal cell-surface PrP protein. *Nature* 356: 577 – 582

Chen HW, Bandyopadhyay S, Shasha DE, Birnbaum KD (2010) Predicting genome-wide redundancy using machine learning. *BMC Evol Biol* 10: 357

Chida J, Hara H, Yano M, Uchiyama K, Das NR, Takahashi E, Miyata H, Tomioka Y, Ito T, Kido H *et al* (2018) Prion protein protects mice from lethal infection with influenza A virus. *PLoS Pathog* 14: e1007049

Cusack SA, Wang P, Lotreck SG, Moore BM, Meng F, Conner JK, Krysan PJ, Lehti-Shiu MD, Shiu SH (2021) Predictive models of genetic redundancy in *Arabidopsis thaliana*. *Mol Biol Evol* 38: 3397 – 3414

Halliez S, Passet B, Martin-Lannerée S, Hernandez-Rapp J, Laude H, Mouillet-Richard S, Vilotte J-L, Bérringue V (2014) To develop with or without the prion protein. *Front Cell Dev Biol*. https://doi.org/10.3389/fcell.2014.00058

Lavi O (2015) Redundancy: a critical obstacle to improving cancer therapy. *Can Res* 75: 808 – 812

Mondragón RJ (2020) Estimating degree–degree correlation and network cores from the connectivity of high–degree nodes in complex networks. *Sci Rep* 10: 5668

Muller HJ (1936) Bar duplication. *Science* 83: 528 – 530

Passet B, Castille J, Makhzami S, Truchet S, Vaiman A, Floriot S, Moazami-Goudarzi K, Vilotte M, Gaillard A-L, Helary L *et al* (2020) The prion-like protein Shadoo is involved in mouse embryonic and mammary development and differentiation. *Sci Rep* 10: 6765

Petras K, Ten Oever S, Dalal SS, Goffaux V (2021) Information redundancy across spatial scales modulates early visual cortical processing. *NeuroImage* 244: 118613

Stanev D, Moustakas K (2019) Modelling musculoskeletal kinematic and dynamic redundancy using null space projection. *PLoS One* 14: e0209171

Sturtevant AH, Morgan TH (1923) Reverse mutations of the bar genes correlated with crossing over. *Science* 57: 746 – 747

van Leeuwen J, Pons C, Tan G, Wang JZ, Hou J, Weile J, Gebbia M, Liang W, Shuteriqi E, Li Z *et al* (2020) Systematic analysis of bypass suppression of essential genes. *Mol Sys Biol* 16: e9828

