

# Appearances can be deceiving: phenotypes of knockout mice

Ivana Barbaric, Gaynor Miller and T. Neil Dear

Advance Access publication date 20 June 2007

## Abstract

In the field of mammalian functional genomics, one of the main aims in the post-genomic era is to elucidate the function of all genes in the genome. The powerful technology of gene targeting in embryonic stem cells has enabled the simple generation of mice lacking a specific gene. However, it is evident that in a proportion of such knockout mice no deviation in phenotype could be detected. Advancements in the field of mouse phenotyping and use of extensive phenotyping tests on each knockout showed that abnormal phenotypes were sometimes detected in physiological areas where they were not initially anticipated, or only manifested under certain conditions, emphasizing the need for careful phenotypic investigation. Nevertheless, the effect of some genes became evident only upon inactivation of another gene, pointing to the phenomenon of biological robustness. Unlike in yeast, this phenomenon has not yet been analysed systematically in the mouse. In this review, we present examples of mouse knockouts that lend support to the concept of robustness, discuss the mechanisms by which it may have evolved, as well as speculate on the reasons for its evolution.

**Keywords:** *genetic redundancy; robustness; phenotype*

## BACKGROUND

Readers may be aware of the amusing allegories written by William Sullivan and Douglas Kellogg on the relative merits of investigating processes using genetic versus biochemical approaches [1]. The analogy, not meant to be taken too seriously, concerns Bill, a retired geneticist, and Doug, a retired biochemist, and their attempts to ascertain how cars work while observing a car production plant. The geneticist resorts to tying the hands of individual workers in the plant and looking at what the cars were like that came off the production line. To his surprise, when he ties the hands of a vice-president, the cars rolled out with no obvious defect. Depending on which story you read (Bill's or Doug's), the conclusion is that either (i) the vice-president had no important role in the production of cars or (ii) there was more than one vice-president, so that another could take over his

duties. This reflects the subject of this review—the generation of mouse mutations with no discernible phenotype and what conclusions we can draw from such experiments.

The first mouse knockout line was made in 1989 [2], marking a major breakthrough in mouse genetics. Gene targeting in mouse ES cells enabled the ablation of any gene in the mouse genome, which would typically be followed by an investigation of the consequences on development, morphology and physiology. Almost 20 years on this technique is a standard tool in mouse genetics and is now used in a large-scale format to resolve one of the major aims of mouse genetics in the post-genomic era: assigning a function to each gene in the mouse genome [3, 4]. Yet, despite the progress in technology, some of the challenges in analysing mouse knockouts remain the same and this is when there is no discernable phenotype, not an

Corresponding author: T. Neil Dear, Mammalian Genetics of Disease Unit, School of Medicine and Biomedical Sciences, University of Sheffield, Beech Hill Road, Sheffield S10 2RX, United Kingdom. Tel: +44 0114 271 3781; Fax: +44 0114 226 8898; E-mail: T.Dear@Sheffield.ac.uk.

**Ivana Barbaric** is a post-doctoral research scientist in the Department of Biomedical Science at the University of Sheffield.

**Gaynor Miller** is a Lecturer in the Mammalian Genetics of Disease Unit in the School of Medicine at the University of Sheffield. Her main area of research is the molecular genetics of muscular dystrophies.

**Neil Dear** is the Head of the Mammalian Genetics of Disease Unit in the School of Medicine at the University of Sheffield. His main area of research is deciphering disease pathways using genetic techniques.

uncommon occurrence upon generating null mutations in the mouse.

It is difficult to estimate what proportion of mouse null mutants generated to date have no observable phenotype. This is because publishing 'negative' data is often difficult, accompanied by the fact that frequently a researcher will hold off publishing in the hope of eventually finding a phenotype. It is considered a common enough occurrence that one journal (*Molecular and Cellular Biology*) devotes a section to reporting such minimal or absent phenotypes. Some idea of the scale of the problem can be ascertained from data obtained for other organisms where systematic approaches have been used to generate loss-of-function mutations. In *Arabidopsis*, only a small proportion of several hundred null mutants had an identifiable phenotype [5]. Data from yeast is particularly useful because loss-of-function mutations have been systematically generated for every gene in the genome. A report of the targeting of genes in yeast showed that ~40–60% of the mutants had no detectable phenotype in the assays used (growth defects, shape and size abnormalities) [6–8]. One more complex model that simulates cell behaviour found that 26% of yeast genes involved in metabolic pathways did not contribute to metabolic phenotype [9].

### Searching for a phenotype

An estimated 10–15% of the genes in the mouse genome have been inactivated using gene targeting technology [3], but it is not clear, or easy to assess, what proportion of these mice have no detectable phenotype. This is partly due to the fact that the degree of phenotypic investigation of knockout mice varies greatly between laboratories, as well as the possibility that many knockout mice with no phenotype have not been published. However, we can get a rough idea from looking at some of the larger gene families where all of the gene members have been knocked out and there is at least one member that has no detectable phenotype e.g. *Casp12* of the 10 caspase gene family knockouts [10], *Adcy4* of the seven adenylate cyclase gene family knockouts [10], *Capn5* of the six calpain gene family knockouts [11], *Pmm1* of the two phosphomannomutase gene family knockouts [12]. Again, this is a biased estimate as, perhaps, redundancy in gene families makes absence of a phenotype more likely than in orphan genes. Nevertheless, it is not unreasonable to suggest that amongst knockout mice

generated, ~10–15% did not have an overt phenotype.

If inactivation of a gene does not lead to an observed abnormal phenotype, there are three possibilities: (i) the abnormal phenotype is present under the conditions currently being used but is yet to be discovered, (ii) the abnormal phenotype will only become evident under environmental conditions that have not yet been tested or (iii) there is no abnormal phenotype.

Using inductive reasoning, can we conclude that a mouse does not have a phenotype after comprehensive efforts to find one? As mentioned earlier, one significant problem in the body of phenotyping data available for knockout mice is the enormous variability in the phenotyping that has been undertaken. This includes the number of phenotyping assays, the physiological systems and processes investigated, and the protocols used. Typically knockout mice for a specific gene are generated in a laboratory with restricted expertise and scientific interests and in many cases mice will, understandably, only be examined in certain phenotypic areas. When nothing abnormal is detected, it is difficult to draw any conclusions. As Lewis Wolpert commented on the supposed absence of a phenotype: «*I say, have you taken your mice to the opera? Can they still tell Wagner from Mozart?*» [13].

The scientific community is well aware of this issue and several large-scale programmes have commenced in order to assist in identifying phenotypes. Efforts have been made to standardize protocols as much as possible, for example, EMPReSS (European Mouse Phenotyping Resource for Standardized Screens) offers a broad range of standardized phenotyping protocols [14]. Of course, if a lab is interested in gene function *per se*, they are likely to pursue broader phenotyping assays themselves. However, most labs are interested in genes involved in specific processes or diseases. They are unlikely to invest effort to find a phenotype for a mutant that is outside of their area of expertise or interest, with neither the time, money nor scientific interest to justify the effort. One way to circumvent this is to create large-scale phenotyping platforms for mice that are accessible to the scientific community but run at specialized centres. For example, the German Mouse Clinic has offered a standardized broad phenotyping platform for mutant mice for the last few years [15], and they have found new phenotypes in 40% of the lines they have

studied [16]. Systematic inactivation of mouse genes has now begun under the auspices of the EUCOMM (European Conditional Mouse Mutagenesis) programme [4] and the Knockout Mouse Project (KOMP, [3]). This is being linked to large-scale phenotyping platforms to analyse these mice in detail e.g. EUMODIC (European Mouse Disease Clinic, [17]).

There are some good examples of unexpected phenotypes that have come about through expanded phenotyping assays. The melanocortin 5 receptor gene (*Mcr5*) knockout mouse initially demonstrated no abnormal response to melanocortin peptides in a variety of phenotyping assays. However, fortuitously, the mice were put through a swim test to examine stress-induced analgesia (the swimming inducing a mild stress response in the mouse) and it was noticed that knockout mice took longer to dry than their wild type counterparts. This reflected a lack of water repulsion in the coats of these mice due to reduced sebum components and implicated the *Mcr5* gene in the regulation of exocrine gland function [18]. Phenotypic abnormalities in previously generated knockout mice have been expanded by investigating new phenotypic areas. For example, *Lif* knockout mice were originally observed to have defects in embryo implantation, yet subsequent studies have revealed immune and neurological abnormalities [19, 20].

## GENES WITH SMALL EFFECTS ON PHENOTYPE

One prediction of the analysis of gene function is that for certain genes the phenotype will be too subtle to detect using small numbers of mice. The theory of neutral selection proposes that genes with only very small effects on fitness can become fixed in the population. There is some evidence for this, e.g. non-essential genes in yeast appear to make marginal fitness contributions [21]. Such genes might have very little discernable effect at the individual level but a significant effect at the population level. Tautz [22] refers to this as the *Biological Uncertainty Principle*. Selection becomes a predictable evolutionary pressure when a population is sufficiently large or selection is sufficiently strong. So consider a population where a new gene has a selective advantage ( $s$ ) in a diploid organism with an effective population size ( $N_e$ ) (i.e. the size of an ideal population that would act the same as the actual population after

taking into account differences in reproductive success and non-random mating). If  $s > 1/2N_e$ , then natural selection will determine the gene frequency. So, for example, even in a relatively small effective population size of, say, 50, a gene with only a 1% improvement on fitness will have a selective advantage. If this weak selection does exist, then it would seem unlikely that the functions of some genes can be uncovered through phenotypic analysis with the numbers of animals that are typically analysed. For example, consider a gene that reduces the red blood cell count by 1% in female mice. Using the haematocrit value for C57BL/6J females of 42.6% [23], the analysis of 1568 mice (784 each of knockout mice and wild type controls) would be needed to reliably detect that difference [24]. This is then compounded by the inherent variability even in inbred mice, making finding subtle phenotypes more difficult [25].

Although this highlights difficulties in assigning clear phenotypes to all genes, the current knockout strategies are still the way to proceed to assign gene function as clearly many genes have a phenotype when a loss-of-function mutant is generated. However, we will find it virtually impossible to differentiate between those genes that have a small effect on phenotype, and those that are masked by redundancy as discussed subsequently.

## MICE WITH NO PHENOTYPE

Despite the difficulties detailed above in detecting a phenotype, there are numerous cases of gene inactivation where mice appear healthy and normal, although one might have expected a significant effect on phenotype based on the known function of the gene and its expression. Such a reduction in the variability of the phenotype to genetic or environmental changes is referred to as *phenotypic robustness*. When genetic changes do not lead to variability of the phenotype, this is referred to as *genetic robustness*. Robustness can be divided into two types on the basis of evolutionary origin [26]:

1. **Adaptive.** Where robustness is the result of natural selection as it increases the fitness of the genotype.
2. **Intrinsic.** Where robustness is an inherent feature of a certain molecular pathway but it is not the direct result of selection for buffering *per se*. For example, where isozymes function under varying regulatory conditions.

It is notable that in yeast, central metabolic pathways appear to have more alternatives than other pathways [27]. This might reflect intrinsic robustness where central metabolic pathways must function under variable physiological and environmental conditions. However, it may also be adaptive—central metabolic pathways are critical for the organism's survival and a back-up mechanism may be advantageous. It has been found in an analysis of transcriptional and signal transduction networks that parallel pathways connecting a regulator to a regulated molecule are not, as is commonly perceived, rare but are actually quite common [28].

A subclass of intrinsic robustness is congruent robustness [29]. This refers to genetic robustness that is a side effect of environmental robustness. For example, consider an enzyme that is expressed at high levels to function under different environmental conditions (perhaps some of which are rarely, or ever, encountered). Any reduction in enzyme activity may still enable the organism to function under the commonly encountered conditions.

## MECHANISMS OF GENETIC ROBUSTNESS

Whereas the importance of robustness is immediately apparent—it enables the organism to endure changes in its environment or genetic make up, the mechanisms by which it arises are not well understood. The mechanisms are considered to be dichotomous: (i) genetic buffering—where alternative pathways for a process exist in the organism, or (ii) functional complementation—where genes are to some extent redundant in function [30]. Two genes are considered to be redundant if they can fully or partially substitute each others functions.

### Paralogous gene redundancy

The mouse genome contains multiple gene families, each consisting of multiple genes that are closely related in structure and physiological function. Approximately 50% of mouse genes have a related member (a paralogue or a pseudogene) in the genome [31]. These genes have almost certainly arisen by gene duplication. It would seem reasonable to assume that directly after duplication, the two genes were fully redundant and carried out identical roles. There is some evolutionary evidence for this initial redundancy. For example, in the *Gdnf* gene family, which consists of four gene members in mice,

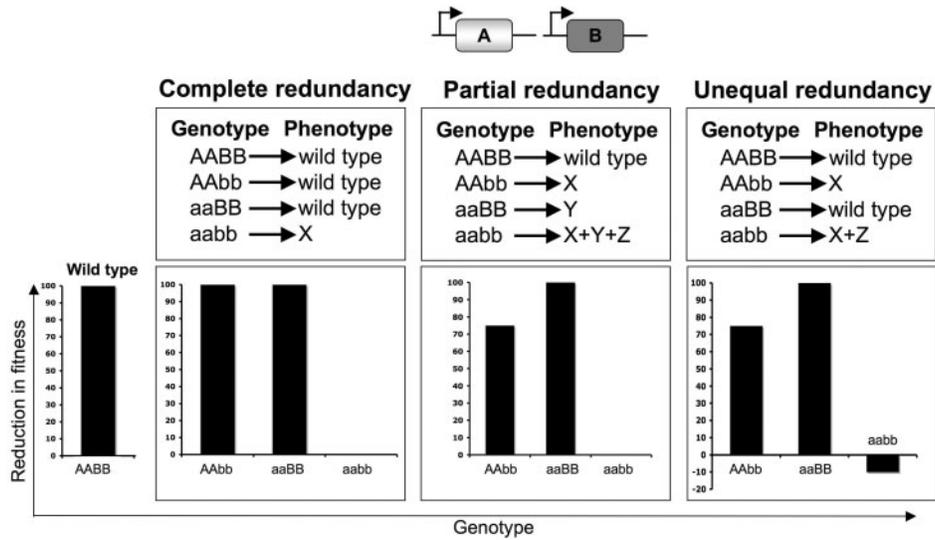
the gene encoding neurturin (*Nrtm*) was lost in the frog but persephin (*Pspn*) was lost in the chicken suggesting true functional redundancy in the early tetrapods [32]. Subsequent divergence would then have proceeded.

Paralogous genetic redundancy is often cited as a mechanism to account for lack of a knockout phenotype. A classic example is the genes *MyoD* and *Myf5* involved in skeletal muscle development. Mice lacking one or other of these genes are similar to wild type mice [33, 34], while the double knockout is completely lacking in skeletal muscle [35]. Redundancy of paralogues can be seen at its most extreme in some of the *Hox* genes. For example, inactivation of five of the six *Hox10a,c,d* alleles, or five of the six *Hox11a,c,d* alleles results in a much less severe phenotype than when all six are inactivated in both cases [36].

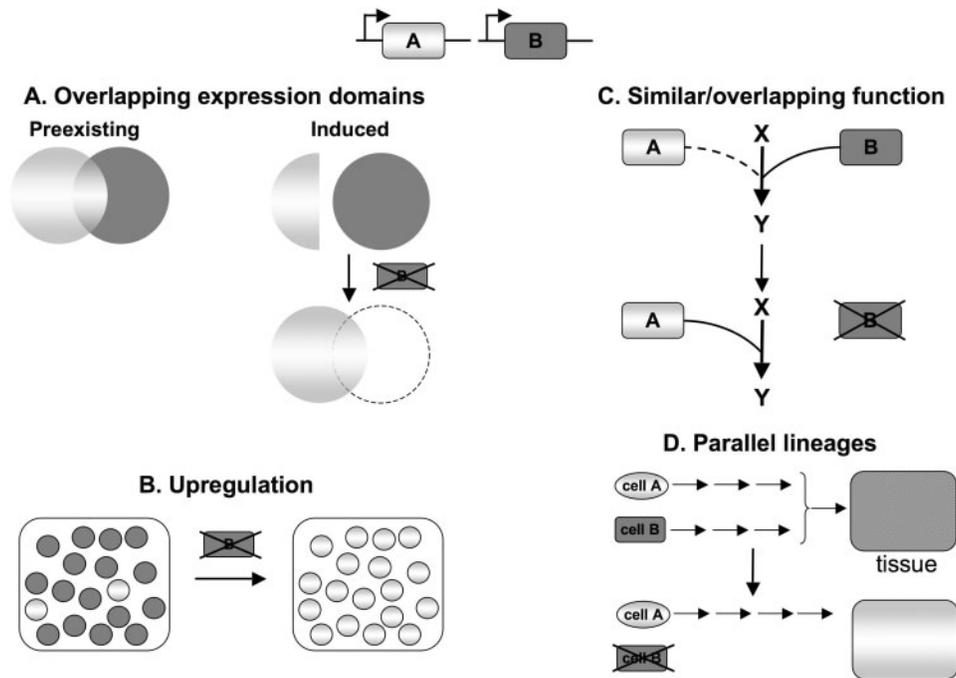
Genetic redundancy can be complete or partial (Figure 1; [37]). The major difference between these two is that in complete redundancy the genes have very similar, if not identical, roles. It is not clear, however, if any such examples exist although they are theoretically possible. Even in the case of *MyoD* and *Myf5*, the individual knockouts have some differences—*Myf5*-null mice have abnormal rib development [33] and delayed trunk muscle development while *MyoD*-null mice have delayed limb and brachial arch muscle development [38].

There are various criteria that must be met for such compensation to occur. First, overlapping expression of the two genes is required, either already present or induced as a result of the loss of function of a gene (Figure 2A). For example, in the case of the partial redundancy between *MyoD* and *Myf5*, they are both expressed in myogenic precursors and developing mouse skeletal muscle [39]. Upregulation of the redundant gene can also occur (Figure 2B). For example, in the skeletal muscles of mice in which *MyoD* has been inactivated, *Myf5* mRNA expression is increased [34], presumably reflecting that MyoD assists in regulating levels of *Myf5* mRNA transcription under normal conditions.

In such cases, a second requirement is similarity in the function of the expressed proteins (Figure 2C). This similarity can be of a varying degree. For example, the two coding sequences may have, at least potentially, identical functions. There is some evidence for this for two members of the MyoD family as a knockin of the *MyoG* (myogenin) coding sequence into the *Myf5* locus shows that myogenin



**Figure 1:** A genetic test for redundancy. Shown are the expected phenotypes that would result in specific scenarios of redundancy. The bar charts shown below exemplify each scenario semi-quantitatively. For the y-axis, 100% indicates a wild type phenotype, while anything less than that indicates an abnormal phenotype resulting in some reduction in fitness. Different phenotypes resulting from different combination of inactive genes would be reflected in different reductions in the overall fitness of the phenotype.



**Figure 2:** Some of the mechanisms of, and requirements for, genetic robustness. (A), overlapping expression of genes with similar functions would enable gene A to cover for the function of gene B in some cells. (B), in the absence of gene B, gene A expression may be increased to compensate. (C), for both of the previous scenarios, it is essential that gene A and B have similar functions. (D), gene A and B may be active in different cell types but activate genetic pathways such that both cell types can differentiate along the same pathway.

protein can rescue *Myf5*-null rib defects [40], suggesting that the MyoG protein fulfils a very similar, if not identical, function to *Myf5* but in a different cell type. In some cases, it may be that two distinct cell

types can commit themselves to the same tissue and when one is lost due to mutation in a critical gene, the other retains (or newly acquires) the ability to differentiate into the required cell type (Figure 2D).

The immune system is a good example of this sort of redundancy. Invading microbes can be destroyed by a variety of different mechanisms and cell types, such as T-dependent and T-independent antibodies, T cell-mediated killing and macrophages.

Differences between human and mouse gene families may explain why some mouse knockouts do not recapitulate the human condition. Consider, for example, the *OCRL1* gene, which encodes a phosphatidylinositol 4,5-bisphosphate 5-phosphatase. This gene is mutated in Lowe syndrome, a rare genetic disorder in humans that results in serious physical and mental problems. Yet, the mouse *Ocr1* knockout appears unaffected. However, mice have a related gene, *Inpp5b*, which is not present in humans. Inactivation of this gene results in only a mild phenotype, while the *Ocr1-Inpp5b* double knockout is embryonic lethal [41]. Thus *Inpp5b* may be able to protect mice from any deleterious effects that would normally result from the absence of *Ocr1*.

### Unequal genetic redundancy in paralogues

After gene duplication, it is assumed that diversification proceeds such that one or both genes acquire new functions or characteristics. Interestingly, in a study using the valuable genetic and biochemical data from yeast, Wagner [42] produced evidence to suggest that this divergence is not equal amongst paralogues—one member of a gene pair appears to acquire more complicated protein interactions and regulatory mechanisms than the other. Immediately after gene duplication, it is possible to envisage that the typical scenario is that one copy becomes mutated such that it still has some activity, but less than the other ‘wild type’ version. A null mutation in the ‘wild type’ version would result in an obvious phenotype while a mutation in the less active version would have no phenotype. As both genes continue to diverge, the scenario evolves with the ‘wild type’ version gaining new function relative to the less active version. One potential explanation for the unequal divergence of the two copies is that if the gene pair diverged symmetrically then the probability of a deleterious mutation occurring is higher than if they had diverged asymmetrically. This is because in asymmetric divergence one gene copy has lost more functions than the other and is less affected by mutations, whereas in symmetrically diverged genes, both gene copies remain highly susceptible to deleterious mutations [42].

In discussing this phenomenon in *Arabidopsis*, Briggs *et al.* [43] highlight that in a typical unequally redundant gene pair one gene is essential, whereas the other is not. Therefore, the knockout for the essential gene would have an obvious phenotype while the second knockout would have a minimal or absent phenotype. There are many examples of this unequal redundancy in mice. As an example, consider a prominent gene family such as the caspases where knockout alleles have been generated for most members. The gene family consist of 14 members. Some knockouts have severe phenotypes e.g. loss of function of *Casp8* results in prenatal lethality [44] and *Casp9* in perinatal lethality [45]. In contrast, *Casp1* knockout mice are relatively healthy with some defects in cytokine processing and resistance to LPS-induced shock [46, 47] while *Casp12* knockouts have no detectable phenotype [48]. Another example of unequal redundancy is the *Prrx1* and *Prrx2* genes. They are closely related and even exhibit similar expression during development. Yet, the knockout phenotypes of these genes are strikingly different. *Prrx1* knockout mice exhibit severe skeletal abnormalities [49], whereas *Prrx2* knockout mice show no discernable phenotype. The double knockout mice, however, suffer from many anomalies in addition to the ones observed in *Prrx1* knockouts [50].

### Genetic redundancy between non-paralogous genes

Homologous genes performing similar functions are often cited as a mechanism for genetic redundancy. In some cases, as exemplified earlier, double and triple knockouts of members of the same gene family support this. Yet, analysis of far more comprehensive data sets from other organisms do not support the notion that genetic redundancy due to overlapping similar gene function is the only mechanism for genetic redundancy. For example, one study, focussing on transposon-tagged set of null mutants in yeast, found that the genes with no apparent null phenotype were no more similar to their closest paralogues in the genome (either in terms of sequence or expression pattern) than those with a clear phenotype [42]. An independent study on a larger gene deletion set showed a slightly higher proportion of genes with paralogues in the ‘no phenotype’ class compared to the essential class, but even so the differences were not startling [7, 8]. Thus, genetic redundancy due to overlapping gene

function is unlikely to explain the lack of phenotype in some null mutants and further suggests that gene duplications contribute only a proportion of the robustness against mutation. This evokes the role of non-paralogous gene pairs in redundancy. Non-paralogous genetic redundancy can be separated into non-paralogues (i) with domain homology and (ii) with no genetic similarity at all.

The extent of redundancy between seemingly unrelated genes may be considerable. Consider the example of the *daf-18* gene in *Caenorhabditis elegans*. A loss-of-function mutation in the mouse homologue of this gene (*Pten*) results in embryonic lethality [51], whereas most *daf-18*<sup>-/-</sup> worms develop normally. Yet, an RNAi screen uncovered 27 different unrelated genes that when inactivated caused a strong sterility or embryonic lethality phenotype in *daf-18*<sup>-/-</sup> mutants [52].

### Domain redundancy

The studies on the role of gene duplication in robustness against mutation have focussed primarily on whole gene duplications where the genes still retain a high degree of similarity. Pasek *et al.* [53] have considered the case of ‘partial duplicate genes’—genes originally generated through gene duplication, but subsequent events (e.g. exon deletion or duplication, gene fusion, exon shuffling) have resulted in the genes only sharing certain domains. In yeast 10% (241/2407) of the genes they analysed were partial duplicates i.e. sharing one or more protein domains while not being true paralogues. They came up with an estimate of between 10% and 37% for the contribution of domain redundancy to robustness against mutation in yeast.

### Genetic redundancy between genes with no genetic similarity

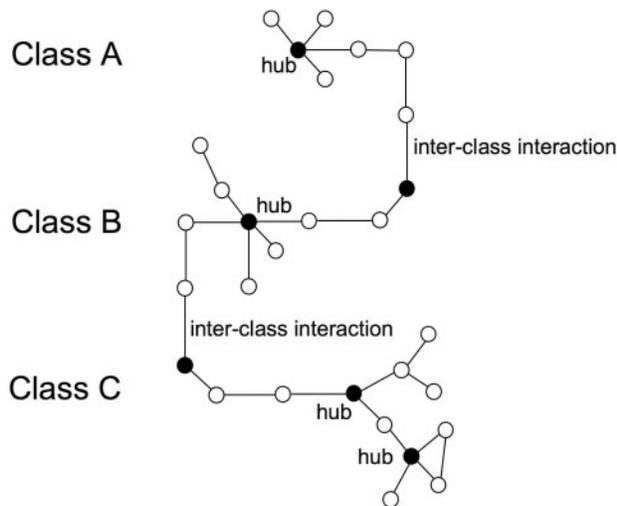
A genetic difference in a single gene can result in either a significant reduction, or even what seems to be a complete absence, of an abnormal phenotype and this redundancy may be complete or partial. One example of the buffering of phenotype by structurally unrelated genes is the potential role of heat shock proteins in buffering mutations. Consider the specific example of Hsp90 in *Drosophila*. When this gene is mutated, widespread phenotypic variation results from other mutations, previously silent in the presence of the wild type Hsp90 [54]. Thus, Hsp90 is able to buffer against genetic mutations that would normally have a phenotypic effect.

### Genetic redundancy at the network level

Networks form an important level of physiological control and further reflect the complexity of robustness against mutation. For example, consider the network of potential protein–protein interactions (PPIs) in the cell. Some proteins are connected to many other proteins in the network (these are termed ‘hubs’) while others have relatively few connections (and are termed ‘non-hubs’) [55]. So those proteins with many connections mediate interactions between many of the less connected proteins. It was shown in yeast, worms and the fly that highly connected proteins—network hubs—are more likely to be essential for survival than unexceptionally connected proteins [56–59]. This is referred to as the central-lethality rule [56]. However, there is a report disputing the above findings and not finding any evidence that the number of links emanating from a protein in a PPI network is correlated with its likelihood of being essential [60]. Nevertheless, apart from PPI networks, transcriptional networks also demonstrate a similar trend, with transcription factors with many gene targets more likely to be essential [58, 61]. Indeed, in transcriptional networks there is evidence that essential transcriptional regulators are more likely to have connections between different classes of regulators (e.g. chromatin condensation versus transcription initiation) rather than within the same functional class [62]. Thus, overall evidence suggests that genes situated at network hubs or that connect regulators of different classes are less likely to exhibit robustness against mutation (Figure 3). Why are hubs more likely to be essential? Analysis by He and Zhang suggests that the answer is simply due to numbers—those proteins with many connections are more likely to have an essential interaction than those with few connections [55].

### THE EFFECT OF GENETIC BACKGROUND

In mice there are numerous examples where the phenotypic effects of a mutation are influenced by the genetic background [63]. The various inbred strains of laboratory mice show considerable genotypic and, as a result, phenotypic variation. This can be clearly seen in the enormous amount of variation for phenotypes already analysed in inbred strains. For example, C3H/HeJ and C57BL/6J—the two most common strains used in mouse genetic



**Figure 3:** Example of a hypothetical network of proteins based on Reference [62]. Lines indicate proteins that are directly connected in some way e.g. a protein–protein interaction or transcription effector and target. Filled circles indicate essential proteins the loss of which results in lethality. The classes represent different classes of regulators e.g. in the case of transcriptional regulators, those that affect chromatin condensation, transcription initiation, RNA processing or nuclear export [62]. Essential proteins are more likely to be localized at hubs or interact with regulators of more than one class.

research—differ by more than 1 SD for 43% (342/748) of the phenotypes compared between them in the Mouse Phenome Database [64]. Observation of genetic robustness may depend on the mouse strain used as specific alleles of genes present in certain inbred mouse strains have been shown to either partially or completely mask the effects of genetic mutations. For example, a variant of the *Mtap1a* gene in inbred strains such as AKR/J and 129P2 can protect mice carrying the *tubby* mutation from hearing loss, but not from other effects of the *tubby* mutation, such as obesity and retinal degeneration [65]. A variant allele of the *Pla2g2a* gene carried by AKR/J but not C57BL/6J mice is able to alter the phenotype severity resulting from the *Apc*<sup>Min</sup> mutation, significantly reducing the number of intestinal tumours that arise [66, 67].

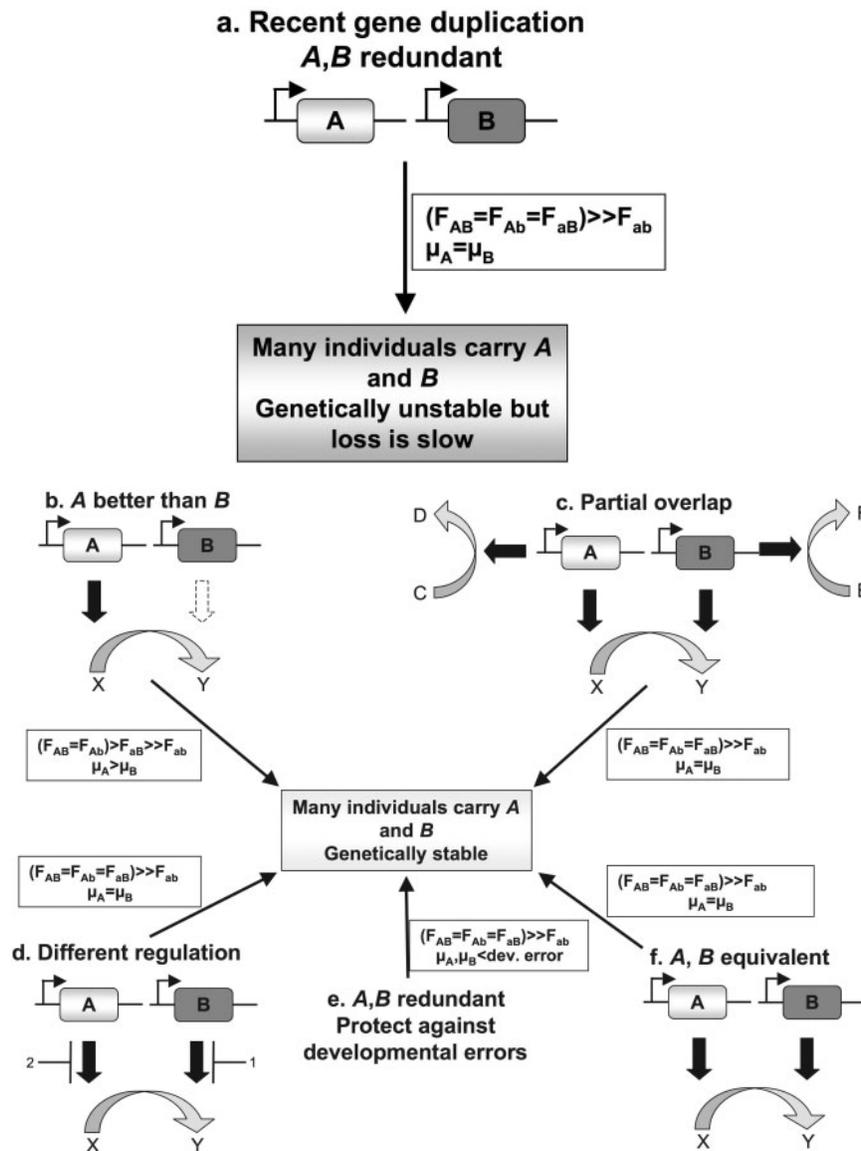
New mutations can also generate genetic robustness. For example, an ENU mutagenesis screen was performed on mice containing an inactivating mutation in the *Mpl* gene that results in thrombocytopaenia. Mutant mice were identified in which the low platelet counts were ameliorated, and subsequent positional cloning identified mutations in the

*Myb* gene. *Mpl*-null mice homozygous for these mutant *Myb* alleles actually showed platelet levels higher than wild type mice [68].

## WHY DOES GENETIC REDUNDANCY EXIST AND HOW DID IT EVOLVE?

Theodosius Dobzhansky once stated «*Nothing in biology makes sense except in the light of evolution*» [69]. Genetic redundancy appears somewhat contradictory from an evolutionary point of view. The redundancy is logical as it creates a ‘fail safe back-up’ for important processes and protects the organism against mutation, thereby increasing the robustness of the organism. Yet the lack of a strong selection pressure acting on an individual gene would predict that the gene would be evolutionarily unstable. Some scenarios have been proposed in which redundancy might develop. Many of these models were proposed by Nowak *et al.* [70].

1. **Recent gene duplication:** Classic evolutionary theory predicts that most often a duplicate gene will be lost shortly after its creation due to accumulating deleterious mutations [71]. It has been argued, however, that under certain conditions genetic redundancy will be selected. For example, consider two genes *A* and *B* that provide complete redundancy and are the result of recent duplication of a common ancestral gene. As long as their mutation rates are similar it will take a long time (more than  $10^7$  generations) for one of the genes to become extinct. This is because only *ab* individuals will be selected against, and both *a* and *b* alleles will be removed at equal rates. Because mutation rates of *A* and *B* are the same, the rates of formation of the *a* and *b* alleles are the same (Figure 4A; [70]).
2. **Asymmetric redundancy:** Gene *A* performs the function slightly better than *B*, but *AB* is no better than *A* alone. Provided that the mutation rate of *A* is higher than *B*, then both genes will be maintained in the population because although gene *B* is less fit, it has a lower mutation rate (Figure 4B; [70]).
3. **Partial overlapping functions:** Both genes are retained because they have other independent functions for which selection pressure is applied, yet they are fully redundant for one particular function (Figure 4C; [70]).



**Figure 4:** Potential evolutionary mechanisms by which genetic redundancy could evolve. See text for details. Boxes indicate relative fitness values of the phenotypes from the alleles A, a, B and b, and the relative mutation rates of the 2 loci A and B required in each scenario. F, fitness of phenotype;  $\mu$ , mutation rate gene.

4. **Side effect of biological robustness:** A robust system can tolerate a certain degree of change, both internal and external, yet still maintain normal function [72]. This is a fundamental feature of organisms—many internal processes remain constant despite changes in the environment e.g. mammals maintain body temperature and basal metabolic processes despite significant changes in food intake and external temperature. Genetic redundancy may be a side effect of the evolution of such robustness. For example, consider genetic redundancy as a side effect of the advantage of having two genes perform

similar roles [73]. Two isozymes may perform a similar role but their regulation may be different such that the levels of enzyme can be finely tuned with different mechanisms being used to control the level of one versus the other, and as result, the overall enzyme activity. This would increase the robustness of the organism by allowing it to operate in more than one situation or condition (Figure 4D; [70]). If only a single gene were present, the regulatory mechanism would have to be simpler and, concomitantly, less responsive to other changes within or outside the cell. In addition to this, when

one gene is lost, the redundant partner could be re-programmed to cover [73]. Again, *MyoD* and *Myf5* are an example of this, as *Myf5* is up regulated when *MyoD* is absent [34].

5. **To protect against errors during development:** One model considers errors that may occur during development of the organism from a one-cell embryo to the adult e.g. somatic mutation, stochastic variation in cell-cell signalling and patterning. Such errors act as a selection pressure for redundancy (Figure 4E; [70]). The authors point out that this model predicts that redundancy should be more common in developmental genes, that have specific patterns of expression, as opposed to those housekeeping genes absolutely required in all cells at all times (as mutation of the latter would result in death of that cell). This would be interesting to test.
6. **Fail-safe mechanisms:** In the simplest scenario, having a back up mechanism for an important pathway may enhance the fitness of an organism by increasing its ability to tolerate what would normally be deleterious mutations (Figure 4F).
7. **Enhancing genetic variation to be acted on by evolution:** For example, the discovery that heat shock proteins are able to buffer against the effects of other genetic mutations (see above) has led to the suggestion that perhaps, such buffering ability enables genetic variation to accumulate and thereby promote evolutionary change in an organism through physiological and developmental processes that would otherwise tolerate very little genetic change [54] (see also review by Rutherford [74]).

## SUMMARY

Functional annotation of the genes from the mouse genome is under way and one of the major tools utilized to fulfil this goal is inactivation of the gene of interest followed by analysing the phenotype of the resulting mutant mice. A caveat of this approach is that the gene function cannot be discerned if a mouse knockout has no detectable abnormal phenotype. Hence, there is tendency to be disappointed when a mouse knockout has no overt phenotype. Similarly, some knockouts do not adequately reflect a clinical phenotype expected from similar mutations in a human orthologue. Yet, although they may not provide us with the information on specific gene function or a model for

a human disease, there is a lot to be learned from ‘no phenotype’ knockouts. Consider the mouse *Hprt* gene, one of the first ever mouse mutations generated by gene targeting in embryonic stem cells. In humans, inactivating mutations in the orthologous gene result in the highly debilitating illness Lesch–Nyhan disease. However, in mice, inactivation of *Hprt* results in mice with none of the behavioural symptoms of this disease [75], possibly because striatal dopamine levels, believed to be responsible for the manifestation of the disease symptoms, are not reduced sufficiently to result in overt disease [76]. There was a suggestion that this might be due to functional redundancy between the *Aprt* and *Hprt* genes in the purine salvage pathway [77], but the subsequent double knockout for these two genes also has none of the abnormal behaviour typical of human patients [78]. Regardless, this difference between mouse and human opens up an opportunity to investigate potential therapies for this disease. There may be some simple compensatory mechanism for the lack of *Hprt* in mice that averts a massive reduction in dopamine levels in the brain. If so, identification of this mechanism may reveal a potential therapeutic avenue to reduce the loss of striatal dopamine in human patients. For example, *Hprt*-null mice do not exhibit hyperuricaemia seen in Lesch–Nyhan patients due to the presence of urate oxidase (uricase), an enzyme humans do not have. However, uricase alone cannot account for the difference between *Hprt*-deficient mice and humans as *Uox* and *Hprt* double-knockout mice also appear behaviourally normal [79].

Although at first often disappointing, mouse knockouts that do not exhibit a predicted phenotype may be of importance in revealing redundant gene networks, alternative pathways or modifiers. Furthermore, discrepancies between the symptoms of human diseases and mouse phenotypes are invaluable in analysing the differences between the species, and these differences may even provide clues for possible therapies.

Mario Capecchi, a pioneer in the field of gene targeting, stated when referring to knockout mice: «*I don't believe there is a single mouse that doesn't have a phenotype . . . . . We just aren't asking the right questions.*» [63]. The presence of genetic redundancy does not have to contradict this view. More than one way of achieving a specific function will increase the robustness of an organism and potentially provide it with a selective advantage in the long

term [72]. Yet, a mouse lacking a functional copy of a gene that is backed up by a second gene still has a phenotype resulting from the lack of the first gene, it may just not be evident at the level of the whole organism. For example, consider two metabolic pathways A and B that result in the synthesis of a substance C. If pathway A is inactivated by mutation in a gene, pathway B will still enable C to be produced. So if phenotype means 'presence of C', then there is no obvious phenotype, yet if phenotype means 'is pathway A functioning' then a phenotype is present. It is conceivable that a back up may result in an alternative way of doing something, rather than directly replacing it. So, as Capecchi notes, the questions need to be the right ones.

The phenomenon of 'no phenotype' knockouts does not diminish the value of large-scale knockout projects aimed at elucidating function of each gene in the mouse genome, but it does signify the need of achieving this aim in conjunction with other forward and reverse genetics approaches. Final conclusions on the extent of genetic redundancy in mammals are awaiting the results of large-scale analyses, but the observations made so far give us a glimpse of the rising challenge in functional genomics: elucidating the way genes interact within genetic networks and pathways to define the complexity of the living organism.

### Key Points

- Phenotypic and genetic robustness is a common feature of organisms.
- The mechanisms for genetic robustness are varied.
- Redundancy can evolve for genes with fully overlapping functions.

### References

1. The Salvation of Doug and the Demise of Bill. <http://review.ucsc.edu/spring04/bio-debate.html> (23 May 2007, date last accessed).
2. Thompson S, Clarke AR, Pow AM, *et al.* Germ line transmission and expression of a corrected HPR1 gene produced by gene targeting in embryonic stem cells. *Cell* 1989;**56**:313–21.
3. Austin CP, Battey JF, Bradley A, *et al.* The knockout mouse project. *Nat Genet* 2004;**36**:921–4.
4. Auwerx J, Avner P, Baldock R, *et al.* The European dimension for the mouse genome mutagenesis program. *Nat Genet* 2004;**36**:925–7.
5. Bouche N, Bouchez D. Arabidopsis gene knockout: phenotypes wanted. *Curr Opin Plant Biol* 2001;**4**:111–7.
6. Smith V, Chou KN, Lashkari D, *et al.* Functional analysis of the genes of yeast chromosome V by genetic footprinting. *Science* 1996;**274**:2069–74.
7. Winzeler EA, Shoemaker DD, Astromoff A, *et al.* Functional characterization of the *S. cerevisiae* genome by gene deletion and parallel analysis. *Science* 1999;**285**:901–6.
8. Giaever G, Chu AM, Ni L, *et al.* Functional profiling of the *Saccharomyces cerevisiae* genome. *Nature* 2002;**418**:387–91.
9. Deutscher D, Meilijson I, Kupiec M, *et al.* Multiple knockout analysis of genetic robustness in the yeast metabolic network. *Nat Genet* 2006;**38**:993–8.
10. Deltagen and Lexicon Knockout Mice. <http://www.informatics.jax.org/external/ko/index.html> (23 May 2007, date last accessed).
11. Franz T, Winckler L, Boehm T, *et al.* Capn5 is expressed in a subset of T cells and is dispensable for development. *Mol Cell Biol* 2004;**24**:1649–54.
12. Cromphout K, Vleugels W, Heykants L, *et al.* The normal phenotype of Pmm1-deficient mice suggests that Pmm1 is not essential for normal mouse development. *Mol Cell Biol* 2006;**26**:5621–35.
13. Wolpert L, Ghysen A, Garcia-Bellido A. Debatable issues. *Int J Dev Biol* 1998;**42**:511–9.
14. Green EC, Gkoutos GV, Lad HV, *et al.* EMPReSS: European mouse phenotyping resource for standardized screens. *Bioinformatics* 2005;**21**:2930–1.
15. Gailus-Durner V, Fuchs H, Becker L, *et al.* Introducing the German Mouse Clinic: open access platform for standardized phenotyping. *Nat Methods* 2005;**2**:403–4.
16. Qiu J. Mighty mouse. *Nature* 2006;**444**:814–6.
17. European Mouse Disease Clinic. <http://www.eumodic.org> (23 May 2007, date last accessed).
18. Chen W, Kelly MA, Opitz-Araya X, *et al.* Exocrine gland dysfunction in MC5-R-deficient mice: evidence for coordinated regulation of exocrine gland function by melanocortin peptides. *Cell* 1997;**91**:789–98.
19. Pechnick RN, Chesnokova VM, Kariagina A, *et al.* Reduced immobility in the forced swim test in mice with a targeted deletion of the leukemia inhibitory factor (LIF) gene. *Neuropsychopharmacol* 2004;**29**:770–6.
20. Schofield G, Kimber SJ. Leukocyte subpopulations in the uteri of leukemia inhibitory factor knockout mice during early pregnancy. *Biol Reprod* 2005;**72**:872–8.
21. Thatcher JW, Shaw JM, Dickinson WJ. Marginal fitness contributions of nonessential genes in yeast. *Proc Natl Acad Sci USA* 1998;**95**:253–7.
22. Tautz D. A genetic uncertainty problem. *Trends Genet* 2000;**16**:475–7.
23. Mouse Phenome Database. Category: Haematology. <http://phenome.jax.org/pub-cgi/phenome/mpdcgi?rtm=assays/group&reqheading=hematology> (23 May 2007, date last accessed).
24. Power Calculation Comparing Independent Samples. <http://stat.ubc.ca/~rollin/stats/ssize/n2.html> (23 May 2007, date last accessed).
25. Meyer CW, Elvert R, Scherag A, *et al.* Power matters in closing the phenotyping gap. *Naturwissenschaften* 2007;**94**:401–406.

26. Hermisson J, Wagner GP. Evolution of phenotypic robustness. In: Jen E (ed). *Robust Design: A Repertoire from Biology, Ecology, and Engineering*. Oxford: Oxford University Press, 2004:47–70.
27. Maltsev N, Glass EM, Ovchinnikova G, et al. Molecular mechanisms involved in robustness of yeast central metabolism against null mutations. *J Biochem (Tokyo)* 2005; **137**:177–87.
28. Wagner A, Wright J. Alternative routes and mutational robustness in complex regulatory networks. *Biosystems* 2007; **88**:163–72.
29. de Visser JA, Hermisson J, Wagner GP, et al. Perspective: evolution and detection of genetic robustness. *Evolution Int J Org Evolution* 2003; **57**:1959–72.
30. Gu X. Evolution of duplicate genes versus genetic robustness against null mutations. *Trends Genet* 2003; **19**: 354–6.
31. Waterston RH, Linblad-Toh K, Birney E, et al. Initial sequencing and comparative analysis of the mouse genome. *Nature* 2002; **420**:520–62.
32. Hatinen T, Holm L, Airaksinen MS. Loss of neurturin in frog—comparative genomics study of GDNF family ligand-receptor pairs. *Mol Cell Neurosci* 2007; **34**:155–67.
33. Braun T, Rudnicki MA, Arnold HH, et al. Targeted inactivation of the muscle regulatory gene *myf-5* results in abnormal rib development and perinatal death. *Cell* 1992; **71**:369–82.
34. Rudnicki MA, Braun T, Hinuma S, et al. Inactivation of *MyoD* in mice leads to up-regulation of the myogenic HLH gene *Myf-5* and results in apparently normal muscle development. *Cell* 1992; **71**:383–90.
35. Rudnicki MA, Schnegelsberg PN, Stead RH, et al. *MyoD* or *Myf-5* is required for the formation of skeletal muscle. *Cell* 1993; **75**:1351–9.
36. Wellik DM, Capecchi MR. *Hox10* and *Hox11* genes are required to globally pattern the mammalian skeleton. *Science* 2003; **301**:363–7.
37. Thomas JH. Thinking about genetic redundancy. *Trends Genet* 1993; **9**:395–9.
38. Kablar B, Krastel K, Ying C, et al. *MyoD* and *Myf-5* differentially regulate the development of limb versus trunk skeletal muscle. *Development* 1997; **124**:4729–38.
39. Hannon K, Smith CK 2nd, Bales KR, et al. Temporal and quantitative analysis of myogenic regulatory and growth factor gene expression in the developing mouse embryo. *Dev Biol* 1992; **151**:137–44.
40. Wang Y, Schnegelsberg PN, Dausman J, et al. Functional redundancy of the muscle-specific transcription factors *Myf5* and *myogenin*. *Nature* 1996; **379**:823–5.
41. Janne PA, Suchy SF, Bernard D, et al. Functional overlap between murine *Inpp5b* and *Ocr1l* may explain why deficiency of the murine ortholog for *OCRL1* does not cause Lowe syndrome in mice. *J Clin Invest* 1998; **101**: 2042–53.
42. Wagner A. Asymmetric functional divergence of duplicate genes in yeast. *Mol Biol Evol* 2002; **19**:1760–8.
43. Briggs GC, Osmond KS, Shindo C, et al. Unequal genetic redundancies in *Arabidopsis*—a neglected phenomenon? *Trends Plant Sci* 2006; **11**:492–8.
44. Varfolomeev EE, Schuchmann M, Luria V, et al. Targeted disruption of the mouse Caspase 8 gene ablates cell death induction by the TNF receptors, Fas/Apo1, and DR3 and is lethal prenatally. *Immunity* 1998; **9**:267–76.
45. Kuida K, Haydar TF, Kuan CY, et al. Reduced apoptosis and cytochrome c-mediated caspase activation in mice lacking caspase 9. *Cell* 1998; **94**:325–37.
46. Kuida K, Lippke JA, Ku G, et al. Altered cytokine export and apoptosis in mice deficient in interleukin-1 beta converting enzyme. *Science* 1995; **267**:2000–3.
47. Li P, Allen H, Banerjee S, et al. Mice deficient in IL-1 beta-converting enzyme are defective in production of mature IL-1 beta and resistant to endotoxic shock. *Cell* 1995; **80**: 401–11.
48. Mouse Genome Informatics, The Jackson Laboratories. URL: [www.informatics.jax.org](http://www.informatics.jax.org) (23 May 2007, date last accessed).
49. Martin JF, Bradley A, Olson EN. The paired-like homeo box gene *MHox* is required for early events of skeletogenesis in multiple lineages. *Genes Dev* 1995; **9**: 1237–49.
50. Ten Berge D, Brouwer A, Korving J, et al. *Prx1* and *Prx2* in skeletogenesis: roles in the craniofacial region of inner ear and limbs. *Development* 1998; **125**:4811–8.
51. Suzuki A, de la Pompa JL, Stambolic V, et al. High cancer susceptibility and embryonic lethality associated with mutation of the PTEN tumor suppressor gene in mice. *Curr Biol* 1998; **8**:1169–78.
52. Suzuki Y, Han M. Genetic redundancy masks diverse functions of the tumor suppressor gene PTEN during *C. elegans* development. *Genes Dev* 2006; **20**:423–8.
53. Pasek S, Risler J-L, Brezellec P. The role of domain redundancy in genetic robustness against null mutations. *J Mol Biol* 2006; **362**:184–91.
54. Rutherford SL, Lindquist S. *Hsp90* as a capacitor for morphological evolution. *Nature* 1998; **396**:336–42.
55. He X, Zhang J. Why do hubs tend to be essential in protein networks? *PLOS Genetics* 2006; **2**:826–34.
56. Jeong H, Mason SP, Barabasi A-L, Oltvai ZN. Lethality and centrality in protein networks. *Nature* 2001; **411**:41–2.
57. Wuchty S. Interaction and domain networks of yeast. *Proteomics* 2002; **2**:1715–23.
58. Yu H, Greenbaum D, Xin Lu H, et al. Genomic analysis of essentiality within protein networks. *Trends Genet* 2004; **20**: 227–31.
59. Hahn MW, Kern AD. Comparative genomics of centrality and essentiality in three eukaryotic protein-interaction networks. *Mol Biol Evol* 2005; **22**:803–6.
60. Coulomb S, Bauer M, Bernard D, Marsolier-Kergoat MC. Gene essentiality and the topology of protein interaction networks. *Proc Biol Sci* 2005; **272**:1721–5.
61. Cooper TF, Morby AP, Gunn A, Schneider D. Effect of random and hub gene disruptions on environmental and mutational robustness in *Escherichia coli*. *BMC Genomics* 2006; **7**:237–47.
62. Tsankov AM, Brown CR, Yu MC, et al. Communication between levels of transcriptional control improves robustness and adaptivity. *Mol Syst Biol* 2006; **2**:1–11.
63. Pearson H. Surviving a knockout blow. *Nature* 2002; **415**: 8–9.
64. Mouse Phenome Database. URL: <http://phenome.jax.org/pub/cgi/phenome/mpdcgi?rtn=docs/home> (23 May 2007, date last accessed).

65. Ikeda A, Zheng QY, Zuberi AR, *et al.* Microtubule-associated protein 1A is a modifier of tubby hearing (moth1). *Nat Genet* 2002;**30**:401–5.
66. Dietrich WF, Lander ES, Smith JS, *et al.* Genetic identification of Mom-1, a major modifier locus affecting Min-induced intestinal neoplasia in the mouse. *Cell* 1993;**75**:631–9.
67. Cormier RT, Hong KH, Halberg RB, *et al.* Secretory phospholipase Pla2g2a confers resistance to intestinal tumorigenesis. *Nat Genet* 1997;**17**:88–91.
68. Carpinelli MR, Hilton DJ, Metcalf D, *et al.* Suppressor screen in Mpl<sup>-/-</sup> mice: c-Myb mutation causes supraphysiological production of platelets in the absence of thrombopoietin signaling. *Proc Natl Acad Sci USA* 2004;**101**:6553–8.
69. Dobzhansky T. Nothing in biology makes sense except in the light of evolution. *American Biology Teacher* 1973;**35**:125–9.
70. Nowak MA, Boerlijst MC, Cooke J, *et al.* Evolution of genetic redundancy. *Nature* 1997;**388**:167–71.
71. Lynch M, Conery JS. The evolutionary fate and consequences of duplicate genes. *Science* 2000;**290**:1151–5.
72. Kitano H. Biological robustness. *Nat Rev Genet* 2004;**5**:826–37.
73. Kafri R, Bar-Even A, Pilpel Y. Transcription control reprogramming in genetic backup circuits. *Nat Genet* 2005;**37**:295–9.
74. Rutherford SL. Between genotype and phenotype: protein chaperones and evolvability. *Nat Rev Genet* 2003;**4**:263–74.
75. Kuehn MR, Bradley A, Robertson EJ, *et al.* A potential animal model for Lesch-Nyhan syndrome through introduction of HPRT mutations into mice. *Nature* 1987;**326**:295–8.
76. Finger S, Heavens RP, Sirinathsinghji DJ, *et al.* Behavioral and neurochemical evaluation of a transgenic mouse model of Lesch-Nyhan syndrome. *J Neurol Sci* 1988;**86**:203–13.
77. Wu CL, Melton DW. Production of a model for Lesch-Nyhan syndrome in hypoxanthine phosphoribosyltransferase-deficient mice. *Nat Genet* 1993;**3**:235–40.
78. Engle SJ, Womer DE, Davies PM, *et al.* HPRT-APRT-deficient mice are not a model for Lesch-Nyhan syndrome. *Hum Mol Genet* 1996;**5**:1607–10.
79. Wu X, Wakamiya M, Vaishnav S, *et al.* Hyperuricemia and urate nephropathy in urate oxidase-deficient mice. *Proc Natl Acad Sci USA* 1994;**91**:742–6.