

Gene duplication as a mechanism of genomic adaptation to a changing environment

Fyodor A. Kondrashov

Proc. R. Soc. B published online 12 September 2012
doi: 10.1098/rsob.2012.1108

References

[This article cites 143 articles, 51 of which can be accessed free](#)

<http://rsob.royalsocietypublishing.org/content/early/2012/09/05/rsob.2012.1108.full.html#ref-list-1>

P<P

Published online 12 September 2012 in advance of the print journal.

EXiS Open Choice

This article is free to access

Subject collections

Articles on similar topics can be found in the following collections

[ecology](#) (1161 articles)
[evolution](#) (1287 articles)
[genomics](#) (17 articles)

Email alerting service

Receive free email alerts when new articles cite this article - sign up in the box at the top right-hand corner of the article or click [here](#)

Advance online articles have been peer reviewed and accepted for publication but have not yet appeared in the paper journal (edited, typeset versions may be posted when available prior to final publication). Advance online articles are citable and establish publication priority; they are indexed by PubMed from initial publication. Citations to Advance online articles must include the digital object identifier (DOIs) and date of initial publication.

*Review***Gene duplication as a mechanism of genomic adaptation to a changing environment****Fyodor A. Kondrashov****Institució Catalana de Recerca i Estudis Avançats (ICREA), Centre for Genomic Regulation (CRG) and Universitat Pompeu Fabra (UPF), 88 Dr Aiguader, Barcelona 08003, Spain*

A subject of extensive study in evolutionary theory has been the issue of how neutral, redundant copies can be maintained in the genome for long periods of time. Concurrently, examples of adaptive gene duplications to various environmental conditions in different species have been described. At this point, it is too early to tell whether or not a substantial fraction of gene copies have initially achieved fixation by positive selection for increased dosage. Nevertheless, enough examples have accumulated in the literature that such a possibility should be considered. Here, I review the recent examples of adaptive gene duplications and make an attempt to draw generalizations on what types of genes may be particularly prone to be selected for under certain environmental conditions. The identification of copy-number variation in ecological field studies of species adapting to stressful or novel environmental conditions may improve our understanding of gene duplications as a mechanism of adaptation and its relevance to the long-term persistence of gene duplications.

Keywords: gene duplication; environmental adaptation; genomics**1. BACKGROUND****(a) Motivation**

Few dispute that gene duplication is the main source of functional diversity on the genotype level. The wealth of sequence data to this effect [1–3] has been preceded by theoretical considerations that it is far easier to create new functions from pre-existing ones rather than from scratch. At the heart of the interest towards gene duplications is the principle that after a duplication ‘each copy can evolve independently and diversify their effects’ [4, p. 64] leading to functional novelty [2,3,5]. As any mutation, a duplication event by itself may also have consequences on organism’s fitness. However, two factors overshadowed the study of the short-term immediate fitness effects of gene duplication. First, the contrast in the difficulties in studying copy-number variation (CNV) that persist to this day [6], and the abundance of long-term evolutionary data on paralogous sequence divergence [1–3,7] shifted the attention to where the data were. Second, and perhaps crucially, the conceptual appeal of gene duplications leading to novel functions was strong enough to overshadow that of the short-term implications of duplications.

The trend of focusing on the long-term implication to the detriment of the study of short-term effects was initiated by Ohno [5] when he proposed that extra gene copies are redundant. His reasoning was that since one copy already performs the necessary function extra copies

are redundant and free from selection, as they do not add anything to the organism’s capacity in performing this function. As it became clearer that gene copies of various degree of divergence are present in genomes, the main question in the field became how can completely redundant gene copies be maintained in the genome long enough to evolve into a new function without being eliminated by mutation. Thus, the theoretical community made extensive use of the redundancy hypothesis with a string of papers looking at this question in some detail [8–15] with many models not formally taking account the process of fixation of a segregating gene duplication.

At the same time, data from the microbiological community were being published suggesting adaptive impacts of gene duplications under certain environmental conditions (see [16–18] for review); however, these data remain largely anecdotal and have not led to a genome-wide survey of selection on gene copy number. The diversification of the community into those that study the short-term versus the long-term effects of gene duplication may have had a linguistic component as some researchers tended to discuss ‘gene duplication’, while others studied ‘gene amplification’ with the relevant papers rarely being cross-cited. However, the fascination with the long-term consequences and the simplicity of the complete redundancy hypothesis, which essentially created the neutral theory of gene duplications, resonated in the community at the time when the neutral theory was popular and few authors took into account the adaptive point of view.

Although it is clear that some gene duplications have been fixed in the course of evolution by positive selection

*fyodor.kondrashov@crg.es

One contribution to a Special Feature ‘Genomics of adaptation’.

[16–26], whether or not it plays a significant role in the fixation of an appreciable fraction of gene duplications has yet to be comprehensively addressed at a genome-wide scale. The purpose of this review is to highlight recently described examples of adaptive short-term effects of gene duplication. Given recent reviews on whole genome duplication [27], negative impact of gene duplications [28–30] and various theories and models on gene duplications [31,32], I will avoid these issues here.

(b) *Duplication and dosage*

The main issue behind whether or not positive selection for an environmental adaptation can be the driving force behind the fixation of gene duplications is whether or not gene duplications are truly redundant. Under the verbal model commonly attributed to Ohno, the variable that affects fitness is an abstract, qualitative function, such that a gene either performs a function or not [5]. By contrast, the adaptive hypothesis holds that function is a quantitative measure of gene action that can be influenced by the amount of gene product in the cell, which in turn can be influenced by the gene copy number [31,32]. However, even if the dosage of the product in the cell influences function, it is possible that a gene duplication is still quantitatively redundant if the product dosage is tightly regulated by negative feedback loops that keep the product dosage constant against gene copy number.

In the last few years, data accumulated demonstrating that while a gene duplication may not necessarily double gene dosage, it still generally leads to its increase [33–35]. It thus seems that one of the ways in which a new gene copy number may affect function is through a quantitative gene dosage effect. The complete redundancy model may still stand based on the possibility that even if dosage is increased the increase does not have an effect on fitness. The evidence against this notion is patchy and not genome-wide. However, several authors have made the claim that many CNVs are selected against in the genome owing to an increase in gene dosage, including cases of CNVs contributing to disease [28–30,36,37] and, therefore, will not be discussed here. At the very least, the data seem to show that situations when a gene duplication affects gene dosage and this, in turn, affects fitness are common. Other mechanisms that can lead to an adaptive response of a gene duplication have been described [32]; however, the dosage response is conceptually the simplest and for the purposes of this review it is assumed that the fitness effect of a gene duplication, when present, lies primarily in that the increased copy number causes an increase in protein dosage.

(c) *Adaptive gene duplications*

Three genomic approaches may detect the action of selection on a specific gene duplication. Unfortunately, all three of these approaches have their limitations. First, one can study whether or not recent copies with equal function are maintained by selection by measuring the ratio of non-synonymous to synonymous evolution (dN/dS) between diverging sequences. If both are maintained by negative selection, then their initial emergence may have occurred through the action of positive selection. The logic is based on the argument that if a substitution $A > G$ is deleterious then reverse $G > A$ substitution is

beneficial. Similarly, if in a gene copy $dN/dS < 1$, indicating that it is currently under negative selection and that its deletion is likely to be deleterious, then contingent on the frequency of gain-of-function mutations and lack of a new function its emergence may have been beneficial. However, when comparing very closely related gene copies with low sequence divergence, which is necessary to reduce the likelihood of this copy already performing a new function, it is not possible to measure a statistically significant departure from neutrality based on just a few substitutions, rendering the dN/dS measure relatively useless. Additionally, when comparing the dN/dS between diverged gene copies, it is possible that a measurement of $dN/dS < 1$ reflects selection maintaining a novel function that emerged in the process of the duplication divergence rather than selection maintaining two copies with identical function. Thus, while many studies of dN/dS between gene copies purport to reveal the action of one mechanism or other in the early evolution of gene duplications [1] such data do not necessarily provide convincing evidence for or against the action of positive selection in the fixation of such copies. Of course, it may be possible to test whether or not very recent gene copies are under selection by looking at the ratio of non-synonymous-to-synonymous polymorphisms in the same manner as dN/dS looks at the ratio of substitutions. However, in order to apply this approach, it is necessary to sequence the gene copies independently of each other, which is not possible with currently available sequencing technologies [6].

The second approach is to measure the levels of variability around the emerging gene copy and search for traces of hitchhiking effects. Only one study in *Arabidopsis thaliana* successfully applied this approach and found evidence of hitchhiking around recent gene copies [38]. However, this remains to my knowledge the only study of this type, possibly owing to the difficulties associated with assembling very recent gene duplications in whole genomes, which is generally necessary for such a study.

Finally, given enough data on fixed gene duplications and segregating CNVs, it may be possible to perform a McDonald–Kreitman test [39] to try to quantify the fraction of gene copies fixed by positive selection. This was crudely done several years ago [36], and the results remain unconvincing owing to the complications in finding a genome-wide set of accurately annotated polymorphic gene duplications to apply the test. It is likely that different genes duplicate at different rates due to difference in length [36] and the influence of short repeats that may substantially increase the rate of duplication of a DNA segment located between them [40]. This issue is coupled to the relative difficulty in identifying polymorphic versus fixed gene copies in genomes [6]. Thus, picking the right set of duplicated and polymorphic gene copies for a non-biased analysis may not be feasible at this point.

In sum, because it does not yet appear possible to systematically test the possibility of positive selection driving the fixation of gene duplications, we may study only specific cases and attempt to make generalizations based on them.

2. EXAMPLES OF ADAPTIVE GENE DUPLICATIONS

(a) *Transport of nutrients*

Nutrient limitation has been observed in many species under different conditions and has been reviewed

previously [18–21]. A clear example of a gene duplication conferring an adaptive response to nutrient limitation is that of the yeast hexose transporter. Under growth conditions with low glucose, the appearance of a new hybrid copy from two closely related paralogues, HXT6 and HXT7, increases the level of expression of the hexose transporter and, crucially, the rate of glucose transport into the cell [41]. Furthermore, the authors have shown in competition experiments that the strain with the gene duplication outcompetes the parental strain. The reason why this case is particularly exemplary is because the duplicated HXT6 and HXT7 genes are recent gene duplications themselves, with several more distantly related paralogues in the genome [18]. It is therefore just a simple step from the experiments of Brown *et al.* [41] to the hypothesis that the paralogues HXT6 and HXT7 have appeared as a result of an adaptive duplication owing to selection favouring extra dosage of the transporters of glucose under various stressful or low-glucose environmental conditions [18,42]. Extensive experiments carried out by Gresham *et al.* [43] also show that many strains evolve amplifications either of the HXT6 or the sHXT7 gene in glucose-limited populations, although some strains adapted to starvation by different mechanisms.

(b) Protection from heat

Adaptation to heat stress has been shown to occur through gene duplication of several stress-related genes in *Escherichia coli* [44]. Some but not all of the genes contained in the duplicated regions showed increased levels of expression and the duplication events coincided with substantial increases in fitness, although not explaining all of it. Furthermore, the duplication of *evgA*, a master transcriptional regulator that is a part of the signal transduction system and upregulates 37 genes, allowed *E. coli* to withstand temperatures of over 50°C [45]. The molecular mechanisms behind such robustness to heat remain unclear. Similar observations have been made in yeast and *Arabidopsis*, such that strains subjected to selection in high temperatures showed an increase propensity for chromosomal segmental duplications [46,47].

(c) Protection from cold

A study of the genome-wide expression levels and the corresponding copy numbers provides convincing evidence for cold adaptations in Antarctic cod. In addition to the expansion of the antifreeze glycoprotein (AFGP) gene family in this species [48] a comparison of the levels of expression between the Antarctic cod and its warm water dwelling relatives revealed 177 genes with substantial overexpression in the Antarctic cod [49]. Through DNA hybridization, it was shown that 118 genes, many of which were from the same set with upregulated expression, have been duplicated in the Antarctic cod, some of them hundreds of times. Duplications were 10 times more common in Antarctic cod suggesting that upregulation through gene duplication of many different functions, including the well-characterized AFGP genes and FBP32II, an F-type lectin that is considered to be a close relative of the fish type II antifreeze proteins, played a role in cold adaptation of this species [49].

(d) Dosage balance and restoration of fitness

Genes rarely act in isolation and it is thought that an optimal dosage of interacting proteins must be maintained for maximal fitness. This logic has been applied to explain why interacting proteins are more likely to be retained together after the whole genome duplication in yeast [50], however, it is also possible that a gene duplication may be favourable because it restores the correct dosage balance in a dosage sensitive system. Two such examples are found in the literature. In yeast, there are two pairs of closely related paralogous histone H2A and H2B proteins, coded by the HTA1–HTB1 and HTA2–HTB2 tandem genes pairs, respectively. When the dosage of the H2A and H2B drops owing to the deletion of the HTA1–HTB1 locus, one of the compensatory mechanisms restoring normal dosage and phenotype is a duplication of the paralogous HTA2–HTB2 locus [51].

Another interesting case was described by Pranting & Andersson [52]. In *Salmonella typhimurium*, a mutation in a haem-biosynthesis enzyme (*hemC*) increases resistance of the bacterium to protamine, however, this resistance comes at a cost of the reduction of growth. In the course of a laboratory evolution experiment, the first step of a series of events after the fixation of the resistance mutation was the amplification of the mutant *hemC* gene that restored some of the fitness [52]. Once the gene accumulated other compensatory point mutations that created a *hemC* gene that conferred resistance to protamine, but without the reduction in the rate of growth, the gene copy number was restored to normal. A similar adaptive response to the reduced fitness of a mutation leading to antibiotic resistance and a subsequent transitory compensatory increase in gene copy number of the mutated gene was found in *Salmonella enterica* [53].

(e) Protection from salt

Adaptation to stress by high salt content may also occur through gene duplication. In one selection experiment, an increase in expression of several stress-response genes was observed, however, this increase does not appear to have anything to do with individual gene duplications as such were generally not observed in the course of the experiment [54]. However, a consistent polyploidy of the yeast strains was observed that may play a role in the adaptation [54,55]. Similarly, polyploidy has been linked to resistance to high salt concentrations in citrus [56] and sorghum [57] suggesting that polyploidy may be a general physiological adaptive response to osmotic stress [54].

(f) Heavy metals

Duplication-induced metal resistance in different species seems to be related mainly to the export of the cations outside the cell [18,36]. Among recent publications, a study showed that *Ralstonia pickettii* adapted to high copper concentrations in copper-contaminated lake sediment tends to accumulate the copper in the outer membrane, and the genome revealed the duplication of a region with several metal resistance and transporter genes [58]. A detailed analysis of genome evolution of *Cupriavidus metallidurans*, a species particularly renowned for its tolerance of heavy metals, revealed two primary mechanisms, the increase in copy number of genes responsible for metal efflux, through

HGT and duplication, and a general decrease in metal uptake [59]. A particular strain of the fungal pathogen *Cryptococcus neoformans*, var. *grubii* subclade VNI A5 MLST, was found to contain a tandem array of several copies of the arsenite efflux transporter that confers copy-number-correlated levels of arsenite resistance [60].

(g) *Antibiotics and drugs*

Many bacteria amplify genes as an adaptive response to antibiotic treatment [36,61,62]. Similarly, it is commonly acknowledged that gene amplifications are known to occur in cancer tumours in response to various drug treatments (see Kondrashov & Kondrashov [36] for review). However, the amplification of genes in response to various drug treatments is not limited to somatic cells and microbes. In the last several years, abundant data have been collected on the amplification of genes in response to various treatments of *Leishmania* [63] and malaria [23]. The *Plasmodium falciparum* multidrug resistance gene (*pfmdr1*) is a target of adaptive evolution in nature in response to the widespread use of chloroquine and other anti-malarial drugs. The *pfmdr1* protein is an ABC transporter involved in the transport of chloroquine outside of the cell, however, it is more directly involved in the adaptation of the malaria parasite to mefloquine, a different anti-malarial drug. Interestingly, *pfmdr1* is a homologue of the human multiple drug resistance (*mdr*) protein that in tumours has a crucial role of expulsion of different drugs from the cell [64,65]. The *pfmdr1* gene is amplified in *P. falciparum* in response to anti-malarial drug treatment, which can confer resistance to mefloquine and other anti-malarial drugs [66–68]. It is now understood that *pfmdr1* gene duplication occurred independently in nature multiple times [66,69], and malaria with increased resistance to different drugs is found throughout the world from Africa [70] to Asia [71,72] and South America [73]. Finally, at least one amplification event of the *pfmdr1* gene shows evidence of having occurred through the action of positive selection [74] and adaptive amplification in *P. falciparum* has been shown in at least one other gene [75], GTP-cyclohydrolase I (*gch1*), which is involved in the synthesis of substrates upstream of other enzymes that are commonly targeted by antifolate drugs.

Gene amplification as a mechanism of adaptation to drugs is also common in *Leishmania* [63]. The amplifications of the genes generally occur as extra-chromosomal circular DNA units [63,76], possibly due to the relatively minor role of transcription initiation in this organism [76]. This may be the reason why gene amplification and the subsequent increase in gene dosage seem to be a common basis for *Leishmania* being resistant to different drugs [77]. A genome-wide assay of *E. coli* strain in 237 toxic environments, mostly antibiotics, found 115 genes an artificial amplification of which leads to the increased fitness in the toxic environments [78]. As with the case of metal toxicity, the function that most widely corresponded to an increase in fitness under the toxic condition was that of efflux pumps and transporters closely followed by genes with regulatory functions and many of the identified genes have not been previously implicated in toxic response. Furthermore, the amplification of several genes conferred resistance in more than

one toxic environment suggesting a wider role of promiscuous functions in the evolution of resistance.

(h) *Pesticides and complex organic compounds*

The application of pesticides and herbicides promotes the adaptive duplication of enzymes that digest the chemical ([79] and see [25] for review). A well-studied example is that of the mosquito *Culex pipiens* that was subject to pesticide treatment on the southern coast of France. Adaptation to pesticide treatment evolved by the duplication of two non-specific esterases (*Est-2* and *Est-3*), and *Ace.1*, a locus encoding acetylcholinesterase that is the main target of the applied organo-phosphate insecticides [36,80]. However, the duplication of *Ace.1* comes at a substantial fitness cost [80], which has led to intricate evolutionary dynamics in the evolution of this locus with subsequent duplications being maintained in heterozygous states across the population owing to a low fitness of the duplications in homozygous state [81]. Furthermore, similar to the strong selection pressure of the anti-malarial drugs leading to independent multiple duplications of the *pfmdr1* gene, the *Ace.1* resistance duplication also appeared multiple times independently around the world [82,83]. The combination of independent origin and the persistence of the duplication in heterozygous form despite the low fitness of homozygous states indicate positive selection for the original emergence of the heterozygous duplication.

(i) *Adaptation to domestication*

Different cow breeds have been analysed for differences in CNVs with different functional families featuring prominently in the differences. First, CNVs of immune-related genes coding for UL16-binding proteins have been described in different breeds of cattle [84,85] and fixed copies of these genes have been shown to be duplicated in the recent ancestor of the *Bos* lineage [86]. Furthermore, recently duplicated genes have been associated with recent domestication-related phenotypes [86] such as milk proteins and proteins related to ruminant digestion. A recent genome-wide study of three cattle breeds identified approximately 400 genes located in CNVs, many of which were related to different environmental interaction and domestication related functions, including parasite and pathogen resistance, lipid transport and metabolism [85]. In a few cases, the types of gene function could be linked to breed-specific phenotypes, suggesting that several aspects of cattle health and productivity have been selected through CNV in the course of cattle domestication.

(j) *Adaptive duplications in humans*

In our own species, several examples have been described of a polymorphic duplication conferring either a distinct phenotype or a fitness advantage. Thus, individuals with more copies of the *CCL3L1* gene, a ligand for the HIV-suppressive CC chemokine receptor, have lower susceptibility to HIV infection and slower progression to AIDS after infection [87]. A widely reviewed example is that of the salivary amylase gene (*AMY1*) that plays a role in the initiation of starch digestion. Individuals from populations with higher starch content in their traditional diets have a higher copy number of this protein providing an apparent fitness benefit [88]. The high prevalence of

CNV in the human genome coupled with data that many of them may be adaptive [37,89–91], suggest that more examples of adaptive gene duplication in the human population are yet to be described.

(k) Generalizations

Adaptive duplications in different species continue to be described in the literature (see table 1 for additional examples) implying that gene duplication as a form of adaptation to various environmental conditions is not a rare mechanism, although by far not an exclusive one. Perhaps several generalizations can be made regarding when an adaptive duplication may be expected to play a role in various adaptations in nature. First, in instances when the protein products of the genes physically interact with molecules associated with a variable environment such as toxins or nutrients. Second, when the proteins coded by the duplicated genes function in the same pathway as those that physically interact with such molecules. Finally, adaptive duplication seems to involve those genes which product needs to be produced rapidly or constantly at a high level, such as antifreeze or storage proteins.

The main logic for hypothesizing that fixed gene duplications played an adaptive role in dosage response to stressful environment has been the functions of gene duplications with characterized adaptive role and the functions of fixed gene copies that are observed in the genome [16–18,21,22,36,104–116]. However, it may be instructive to reverse the logic and predict the types of genes that may confer an adaptation in some environmental conditions based on the functional repertoire of the gene duplications observed in the genome.

One of the main duplicated gene families are the olfactory receptor proteins [18,117–119] so perhaps their duplication may lead to an increase in sensitivity to a particular odour may be adaptive under certain conditions. The match in the sensitivity range of duplicated opsin genes and coloration [120,121] suggests that the duplication of an opsin gene may cause an adaptive increase in sensitivity to light of a certain wavelength. Duplications of genes involved in pathogen resistance [105,108–111,122,123] and pathogenicity [110,124] suggest that gene duplications may have a role in rapid coevolution between host and pathogen or symbiont. Recent gene duplications in drought-resistant plants [125] suggest that genes involved in osmotic stress response are also good candidates. Duplication of the globin genes [107,126] may be adaptive in low-oxygen environments, such as high altitude, by optimizing oxygen transport. Common duplications of different defence toxins and venoms [116,127–132] suggest that toxin dosage and their effectiveness may be regulated through gene duplication. Similarly, the duplication of proteins involved in the degradation of toxins, for example enzymes that neutralize pyrrolizidine alkaloids [133], may confer a fitness advantage to a herbivore feeding on a plant with high concentration of specific toxins. The duplication of developmentally related genes [134] may confer an adaptive advantage in the anatomical modification of the body plan for a new environment. The production of carotenoids [135] or light emission by luciferase [136] may be dosage-mediated by gene duplication leading to extra pigment or higher intensity of light. The duplication of

Table 1. Recent examples of adaptive gene duplication not discussed in the main text.

gene	function	species	environmental condition	evidence	references
Lac operon	lactose metabolism	<i>Escherichia coli</i>	starvation	laboratory experiments	[92–94]
SUL1	high affinity sulphur transporter	<i>Saccharomyces cerevisiae</i>	low-sulphur environment	laboratory experiments	[43]
prolamine-coding genes	storage proteins	Poaceae	nutrient storage in seeds	indirect inference from large number of copies	[95]
AFP	antifreeze protein	<i>Anarhichas lupus</i>	low temperature	higher number of copies and higher expression in a cold-tolerant species	[96]
IRI-like gene family	ice re-crystallisation inhibition	Pooideae	low temperature	gene family expansion in cold-tolerant lineage	[97]
COR15	cold tolerance-related protein	Brassicaceae	low temperature	several independent duplications	[98]
ribosomal protein coding genes	protein translation	<i>Salmonella typhimurium</i>	horizontal gene transfer	laboratory experiment	[99]
EPSPS	5-enolpyruvylshikimate-3-phosphate synthase	<i>Amaranthus palmeri</i> , <i>Medicago sativa</i> , <i>Glycine max</i> , <i>Nicotiana tabacum</i> , other species	glyphosate	field studies, experimental evidence	[26,100–102]
Gpdh	glycerol-3-phosphate dehydrogenase	<i>Drosophila melanogaster</i>	ethanol	laboratory experiments with artificial amplification	[103]

proteases expressed in the female reproductive tract of *Drosophila mojavensis* [137] and sperm-related proteins in *Caenorhabditis elegans* [18] suggests that dosage increase through gene duplication may be adaptive as a strategy in sexual conflict.

3. DISCUSSION

The development of the theory of gene duplications reflects that of the neutral theory of molecular evolution. The strong claim of neutrality of gene duplications made by Ohno [5] reflected a point of view that all gene duplications are neutral. Most theoretical models that followed maintained that the redundancy of the gene duplication leads to its neutrality. In the last few years, it became evident that many gene duplications are deleterious from the moment of their origin [28–30,138]. However, this can be easily accommodated in the redundancy-based theories on long-term evolution of gene duplication in the form of a weak claim of neutrality: those gene duplications that are fixed are neutral. Because no quantitative genome-wide estimates of the fraction of gene duplications that are fixed by positive selection have been made, population genetic theories explaining the evolution of gene duplication have not incorporated selection on gene copy number in the models [31,32].

At the present moment, the literature reflects a point of view that gene duplications are more neutral than amino acid substitutions. A large fraction of amino acid substitutions is thought to be fixed by positive selection [139], yet most authors continue to explain the origin and maintenance of gene duplications by referring to theories of their maintenance that assume complete redundancy and a neutral path to fixation. However, presumably a gene duplication is mutation of stronger impact than a point mutation and, therefore, it should on average have a more profound impact on fitness than a substitution [19,20]. This somewhat paradoxical situation is reflected in the literature on CNVs where authors avoid reference to the existing theories on gene duplication [28–30,36,37].

Owing to the fascination with the idea of complete genetic redundancy, classical population genetic and evolutionary models of gene duplications generally failed to be useful in the description of CNV in natural population (but see [140–142]). Perhaps the same fate of irrelevance awaits the theoretical field once ecological genomics starts addressing the question of adaptation to stressful environmental conditions on a genome-wide scale in multiple non-model organisms. To model such adaptation, the necessary theoretical models may be more complicated than those simply incorporating a quantitative measure of genetic redundancy as they may also have to be time-dependent; in several cases, it appears that a gene duplication that is adaptive under a stressful condition comes at a fitness cost in a benign environment [36,52,80,143,144].

The cost and utility of sequencing technology makes it possible to obtain a genome sequence of reasonable quality for all but the most complicated genomes. The application of such technology to natural populations is bound to reveal new mechanisms of ecological adaptation. Given the abundance of examples of adaptive gene duplications that emerge in the course of adaptation to stressful environments, researchers studying the

adaptation of species to novel or stressful environments would be well advised to consider taking a closer look at recent gene duplications and copy-number polymorphisms. Meanwhile, the relative role of positive selection versus drift in fixation of gene duplications remains an open question.

The work was supported by a Plan Nacional grant no. BFU2009-09271 from the Spanish Ministry of Science and Innovation. The author is a European Molecular Biology Organization Young Investigator and Howard Hughes Medical Institute International Early Career Scientist.

REFERENCES

- Lynch, M. & Conery, J. S. 2000 The evolutionary fate and consequences of duplicate genes. *Science* **290**, 1151–1155. (doi:10.1126/science.290.5494.1151)
- Ponting, C. P. 2008 The functional repertoires of metazoan genomes. *Nat. Rev. Genet.* **9**, 689–698. (doi:10.1038/nrg2413)
- Conant, G. C. & Wolfe, K. H. 2008 Turning a hobby into a job: how duplicated genes find new functions. *Nat. Rev. Genet.* **9**, 938–950. (doi:10.1038/nrg2482)
- Bridges, C. A. 1935 Salivary chromosome maps. *J. Hered.* **26**, 60–64.
- Ohno, S. 1970 *Evolution by gene duplication*. Berlin, Germany: Springer.
- Alkan, C., Coe, B. P. & Eichler, E. E. 2011 Genome structural variation discovery and genotyping. *Nat. Rev. Genet.* **12**, 363–376. (doi:10.1038/nrg2958)
- Demuth, J. P. & Hahn, M. W. 2009 The life and death of gene families. *Bioessays* **31**, 29–39. (doi:10.1002/bies.080085)
- Kimura, M. & King, J. L. 1979 Fixation of a deleterious allele at one of two 'duplicate' loci by mutation pressure and random drift. *Proc. Natl Acad. Sci. USA* **76**, 2858–2861. (doi:10.1073/pnas.76.6.2858)
- Li, W. H. 1980 Rate of gene silencing at duplicate loci: a theoretical study and interpretation of data from tetraploid fishes. *Genetics* **95**, 237–258.
- Ohta, T. 1990 How gene families evolve? *Theor. Popul. Biol.* **37**, 213–219. (doi:10.1016/0040-5809(90)90036-U)
- Walsh, J. B. 1995 How often do duplicated genes evolve new functions? *Genetics* **139**, 421–428.
- Wagner, A. 1998 The fate of duplicated genes: loss or new function? *BioEssays* **20**, 785–788. (doi:10.1002/(SICI)1521-1878(199810)20:10<785::AID-BIES2>3.0.CO;2-M)
- Stoltzfus, A. 1999 On the possibility of constructive neutral evolution. *J. Mol. Evol.* **49**, 169–181. (doi:10.1007/PL00006540)
- Force, A., Lynch, M., Pickett, F. B., Amores, A., Yan, Y. L. & Postlethwait, J. 1999 Preservation of duplicate genes by complementary, degenerative mutations. *Genetics* **151**, 1531–1545.
- Lynch, M. & Force, A. 2000 The probability of duplicate gene preservation by subfunctionalization. *Genetics* **154**, 459–473.
- Velkov, V. V. 1982 Gene amplification in prokaryotic and eukaryotic systems. *Genetika* **18**, 529–543.
- Romero, D. & Palacios, R. 1997 Gene amplification and genomic plasticity in prokaryotes. *Annu. Rev. Genet.* **31**, 91–111. (doi:10.1146/annurev.genet.31.1.91)
- Kondrashov, F. A., Rogozin, I. B., Wolf, Y. I. & Koonin, E. V. 2002 Selection in the evolution of gene duplications. *Genome Biol.* **3**, research0008. (doi:10.1186/gb-2002-3-2-research0008)
- Rapoport, I. A. 1940 Mnogokratnye linejnye povtoreniya uchastkov khromosom i ikh evolyucionnoe

- znachenie. [Multiple linear repeats of chromosome segments and their evolutionary significance]. *Zh. Obshchej Biologii* **1**, 235–270. [In Russian]
- 20 Sebat, J. *et al.* 2004 Large-scale copy number polymorphism in the human genome. *Science* **305**, 525–528. (doi:10.1126/science.1098918)
 - 21 Hastings, P. J. 2007 Adaptive amplification. *Crit. Rev. Biochem. Mol. Biol.* **42**, 271–283. (doi:10.1080/10409230701507757)
 - 22 Flagel, L. E. & Wendel, J. F. 2009 Gene duplication and evolutionary novelty in plants. *New Phytol.* **183**, 557–564. (doi:10.1111/j.1469-8137.2009.02923.x)
 - 23 Templeton, T. J. 2009 The varieties of gene amplification, diversification and hypervariability in the human malaria parasite, *Plasmodium falciparum*. *Mol. Biochem. Parasitol.* **166**, 109–116. (doi:10.1016/j.molbiopara.2009.04.003)
 - 24 Nascimento, N. C. & Fett-Neto, A. G. 2010 Plant secondary metabolism and challenges in modifying its operation: an overview. *Methods Mol. Biol.* **643**, 1–13. (doi:10.1007/978-1-60761-723-5-1)
 - 25 Bass, C. & Field, L. M. 2011 Gene amplification and insecticide resistance. *Pest Manag. Sci.* **67**, 886–890. (doi:10.1002/ps.2189)
 - 26 Shaner, D. L., Lindenmeyer, R. B. & Ostlie, M. H. 2012 What have the mechanisms of resistance to glyphosate taught us? *Pest Manag. Sci.* **68**, 3–9. (doi:10.1002/ps.2261)
 - 27 Van de Peer, Y., Maere, S. & Meyer, A. 2009 The evolutionary significance of ancient genome duplications. *Nat. Rev. Genet.* **10**, 725–732. (doi:10.1038/nrg2600)
 - 28 de Smith, A. J., Walters, R. G., Froguel, P. & Blake-more, A. I. 2008 Human genes involved in copy number variation: mechanisms of origin, functional effects and implications for disease. *Cytogenet. Genome Res.* **123**, 17–26. (doi:10.1159/000184688)
 - 29 Henrichsen, C. N., Chaignat, E. & Reymond, A. 2009 Copy number variants, diseases and gene expression. *Hum. Mol. Genet.* **18**, R1–R8. (doi:10.1093/hmg/ddp011)
 - 30 Zhang, F., Gu, W., Hurles, M. E. & Lupski, J. R. 2009 Copy number variation in human health, disease, and evolution. *Annu. Rev. Genomics Hum. Genet.* **10**, 451–481. (doi:10.1146/annurev.genom.9.081307.164217)
 - 31 Hahn, M. W. 2009 Distinguishing among evolutionary models for the maintenance of gene duplicates. *J. Hered.* **100**, 605–617. (doi:10.1093/jhered/esp047)
 - 32 Innan, H. & Kondrashov, F. 2010 The evolution of gene duplications: classifying and distinguishing between models. *Nat. Rev. Genet.* **11**, 97–108. (doi:10.1038/nrg2689)
 - 33 Schuster-Böckler, B., Conrad, D. & Bateman, A. 2010 Dosage sensitivity shapes the evolution of copy-number varied regions. *PLoS ONE* **5**, e9474. (doi:10.1371/journal.pone.0009474)
 - 34 Stranger, B. E. *et al.* 2007 Relative impact of nucleotide and copy number variation on gene expression phenotypes. *Science* **315**, 848–853. (doi:10.1126/science.1136678)
 - 35 Zhou, J., Lemos, B., Dopman, E. B. & Hartl, D. L. 2011 Copy-number variation: the balance between gene dosage and expression in *Drosophila melanogaster*. *Genome Biol. Evol.* **3**, 1014–1024. (doi:10.1093/gbe/evr023)
 - 36 Kondrashov, F. A. & Kondrashov, A. S. 2006 Role of selection in fixation of gene duplications. *J. Theor. Biol.* **239**, 141–151. (doi:10.1016/j.jtbi.2005.08.033)
 - 37 Lupski, J. R. 2007 An evolution revolution provides further revelation. *Bioessays* **29**, 1182–1184. (doi:10.1002/bies.20686)
 - 38 Moore, R. C. & Purugganan, M. D. 2003 The early stages of duplicate gene evolution. *Proc. Natl Acad. Sci. USA* **100**, 15 682–15 687. (doi:10.1073/pnas.2535513100)
 - 39 McDonald, J. H. & Kreitman, M. 1991 Adaptive protein evolution at the Adh locus in *Drosophila*. *Nature* **351**, 652–654. (doi:10.1038/351652a0)
 - 40 Shaw, C. J. & Lupski, J. R. 2004 Implications of human genome architecture for rearrangement-based disorders: the genomic basis of disease. *Hum. Mol. Genet.* **13**, R57–R64. (doi:10.1093/hmg/ddh073)
 - 41 Brown, C. J., Todd, K. M. & Rosenzweig, R. F. 1998 Multiple duplications of yeast hexose transport genes in response to selection in a glucose-limited environment. *Mol. Biol. Evol.* **15**, 931–942. (doi:10.1093/oxfordjournals.molbev.a026009)
 - 42 Lin, Z. & Li, W. H. 2011 Expansion of hexose transporter genes was associated with the evolution of aerobic fermentation in yeasts. *Mol. Biol. Evol.* **28**, 131–142. (doi:10.1093/molbev/msq184)
 - 43 Gresham, D., Desai, M. M., Tucker, C. M., Jenq, H. T., Pai, D. A., Ward, A., DeSevo, C. G., Botstein, D. & Dunham, M. J. 2008 The repertoire and dynamics of evolutionary adaptations to controlled nutrient-limited environments in yeast. *PLoS Genet.* **4**, e1000303. (doi:10.1371/journal.pgen.1000303)
 - 44 Riehle, M. M., Bennett, A. F. & Long, A. D. 2001 Genetic architecture of thermal adaptation in *Escherichia coli*. *Proc. Natl Acad. Sci. USA* **98**, 525–530. (doi:10.1073/pnas.98.2.525)
 - 45 Christ, D. & Chin, J. W. 2008 Engineering *Escherichia coli* heat-resistance by synthetic gene amplification. *Protein Eng. Des. Sel.* **21**, 121–125. (doi:10.1093/protein/gzm085)
 - 46 James, T. C., Usher, J., Campbell, S. & Bond, U. 2008 Lager yeasts possess dynamic genomes that undergo rearrangements and gene amplification in response to stress. *Curr. Genet.* **53**, 139–152. (doi:10.1007/s00294-007-0172-8)
 - 47 DeBolt, S. 2010 Copy number variation shapes genome diversity in *Arabidopsis* over immediate family generational scales. *Genome Biol. Evol.* **2**, 441–453. (doi:10.1093/gbe/evq033)
 - 48 Chen, L., DeVries, A. L. & Cheng, C. H. 1997 Evolution of antifreeze glycoprotein gene from a trypsinogen gene in Antarctic notothenioid fish. *Proc. Natl Acad. Sci. USA* **94**, 3811–3816. (doi:10.1073/pnas.94.8.3811)
 - 49 Chen, Z. *et al.* 2008 Transcriptomic and genomic evolution under constant cold in Antarctic notothenioid fish. *Proc. Natl Acad. Sci. USA* **105**, 12 944–12 949. (doi:10.1073/pnas.0802432105)
 - 50 Papp, B., Pál, C. & Hurst, L. D. 2003 Dosage sensitivity and the evolution of gene families in yeast. *Nature* **424**, 194–197. (doi:10.1038/nature01771)
 - 51 Libuda, D. E. & Winston, F. 2006 Amplification of histone genes by circular chromosome formation in *Saccharomyces cerevisiae*. *Nature* **443**, 1003–1007. (doi:10.1038/nature05205)
 - 52 Pránting, M. & Andersson, D. I. 2011 Escape from growth restriction in small colony variants of *Salmonella typhimurium* by gene amplification and mutation. *Mol. Microbiol.* **79**, 305–315. (doi:10.1111/j.1365-2958.2010.07458.x)
 - 53 Paulander, W., Andersson, D. I. & Maisnier-Patin, S. 2010 Amplification of the gene for isoleucyl-tRNA synthetase facilitates adaptation to the fitness cost of mupirocin resistance in *Salmonella enterica*. *Genetics* **185**, 305–312. (doi:10.1534/genetics.109.113514)
 - 54 Dhar, R., Sägeser, R., Weikert, C., Yuan, J. & Wagner, A. 2011 Adaptation of *Saccharomyces cerevisiae* to saline stress through laboratory evolution. *J. Evol. Biol.* **24**, 1135–1153. (doi:10.1111/j.1420-9101.2011.02249.x)

- 55 Gerstein, A. C., Chun, H. J., Grant, A. & Otto, S. P. 2006 Genomic convergence toward diploidy in *Saccharomyces cerevisiae*. *PLoS Genet.* **2**, e145. (doi:10.1371/journal.pgen.0020145)
- 56 Saleh, B., Allario, T., Dambier, D., Ollitrault, P. & Morillon, R. 2008 Tetraploid citrus rootstocks are more tolerant to salt stress than diploid. *C. R. Biol.* **331**, 703–710. (doi:10.1016/j.crv.2008.06.007)
- 57 Ceccarelli, M., Santantonio, E., Marmottini, F., Amzallag, G. N. & Cionini, P. G. 2006 Chromosome endoreduplication as a factor of salt adaptation in *Sorghum bicolor*. *Protoplasma* **227**, 113–118. (doi:10.1007/s00709-005-0144-0)
- 58 Yang, F., Pecina, D. A., Kelly, S. D., Kim, S. H., Kemner, K. M., Long, D. T. & Marsh, T. L. 2010 Biosequestration via cooperative binding of copper by *Ralstonia pickettii*. *Environ. Technol.* **31**, 1045–1060. (doi:10.1080/09593330.2010.487290)
- 59 von Rozyczki, T. & Nies, D. H. 2009 *Cupriavidus metallidurans*: evolution of a metal-resistant bacterium. *Antonie Van Leeuwenhoek* **96**, 115–139. (doi:10.1007/s10482-008-9284-5)
- 60 Chow, E. W., Morrow, C. A., Djordjevic, J. T., Wood, I. A. & Fraser, J. A. 2012 Microevolution of *Cryptococcus neoformans* driven by massive tandem gene amplification. *Mol. Biol. Evol.* **29**, 1987–2000. (doi:10.1093/molbev/mss066)
- 61 Craven, S. H. & Neidle, E. L. 2007 Double trouble: medical implications of genetic duplication and amplification in bacteria. *Future Microbiol.* **2**, 309–321. (doi:10.2217/17460913.2.3.309)
- 62 Sandegren, L. & Andersson, D. I. 2009 Bacterial gene amplification: implications for the evolution of antibiotic resistance. *Nat. Rev. Microbiol.* **7**, 578–588. (doi:10.1038/nrmicro2174)
- 63 Segovia, M. 1994 *Leishmania* gene amplification: a mechanism of drug resistance. *Ann. Trop. Med. Parasitol.* **88**, 123–130.
- 64 Foote, S. J., Thompson, J. K., Cowman, A. F. & Kemp, D. J. 1989 Amplification of the multidrug resistance gene in some chloroquine-resistant isolates of *P. falciparum*. *Cell* **57**, 921–930. (doi:10.1016/0092-8674(89)90330-9)
- 65 Foote, S. J., Kyle, D. E., Martin, R. K., Oduola, A. M., Forsyth, K., Kemp, D. J. & Cowman, A. F. 1990 Several alleles of the multidrug-resistance gene are closely linked to chloroquine resistance in *Plasmodium falciparum*. *Nature* **345**, 255–258. (doi:10.1038/345255a0)
- 66 Duraisingh, M. T. & Cowman, A. F. 2005 Contribution of the pfm^{dr}1 gene to antimalarial drug-resistance. *Acta Trop.* **94**, 181–190. (doi:10.1016/j.actatropica.2005.04.008)
- 67 Preechapornkul, P., Imwong, M., Chotivanich, K., Pongtavornpinyo, W., Dondorp, A. M., Day, N. P., White, N. J. & Pukrittayakamee, S. 2009 *Plasmodium falciparum* pfm^{dr}1 amplification, mefloquine resistance, and parasite fitness. *Antimicrob. Agents Chemother.* **53**, 1509–1515. (doi:10.1128/AAC.00241-08)
- 68 Chavchich, M., Gerena, L., Peters, J., Chen, N., Cheng, Q. & Kyle, D. E. 2010 Role of pfm^{dr}1 amplification and expression in induction of resistance to artemisinin derivatives in *Plasmodium falciparum*. *Antimicrob. Agents Chemother.* **54**, 2455–2464. (doi:10.1128/AAC.00947-09)
- 69 Triglia, T., Foote, S. J., Kemp, D. J. & Cowman, A. F. 1991 Amplification of the multidrug resistance gene pfm^{dr}1 in *Plasmodium falciparum* has arisen as multiple independent events. *Mol. Cell. Biol.* **11**, 5244–5250.
- 70 Gadalla, N. B. *et al.* 2011 Increased pfm^{dr}1 copy number and sequence polymorphisms in *Plasmodium falciparum* isolates from Sudanese malaria patients treated with artemether–lumefantrine. *Antimicrob. Agents Chemother.* **55**, 5408–5411. (doi:10.1128/AAC.05102-11)
- 71 Imwong, M. *et al.* 2008 Gene amplification of the multidrug resistance 1 gene of *Plasmodium vivax* isolates from Thailand, Laos, and Myanmar. *Antimicrob. Agents Chemother.* **52**, 2657–2659. (doi:10.1128/AAC.01459-07)
- 72 Meng, H., Zhang, R., Yang, H., Fan, Q., Su, X., Miao, J., Cui, L. & Yang, Z. 2010 *In vitro* sensitivity of *Plasmodium falciparum* clinical isolates from the China–Myanmar border area to quinine and association with polymorphism in the Na⁺/H⁺ exchanger. *Antimicrob. Agents Chemother.* **54**, 4306–4313. (doi:10.1128/AAC.00321-10)
- 73 Griffing, S. *et al.* 2010 pfm^{dr}1 amplification and fixation of pf^{cr}t chloroquine resistance alleles in *Plasmodium falciparum* in Venezuela. *Antimicrob. Agents Chemother.* **54**, 1572–1579. (doi:10.1128/AAC.01243-09)
- 74 Nair, S., Nash, D., Sudimack, D., Jaidee, A., Barends, M., Uhlemann, A. C., Krishna, S., Nosten, F. & Anderson, T. J. 2007 Recurrent gene amplification and soft selective sweeps during evolution of multidrug resistance in malaria parasites. *Mol. Biol. Evol.* **24**, 562–573. (doi:10.1093/molbev/msl185)
- 75 Nair, S. *et al.* 2008 Adaptive copy number evolution in malaria parasites. *PLoS Genet.* **4**, e1000243. (doi:10.1371/journal.pgen.1000243)
- 76 Ubeda, J. M. *et al.* 2008 Modulation of gene expression in drug resistant *Leishmania* is associated with gene amplification, gene deletion and chromosome aneuploidy. *Genome Biol.* **9**, R115. (doi:10.1186/gb-2008-9-7-r115)
- 77 Mary, C., Faraut, F., Deniau, M., Dereure, J., Aoun, K., Ranque, S. & Piarroux, R. 2010 Frequency of drug resistance gene amplification in clinical *Leishmania* strains. *Int. J. Microbiol.* **2010**, 819060. (doi:10.1155/2010/819060)
- 78 Soo, V. W., Hanson-Manful, P. & Patrick, W. M. 2011 Artificial gene amplification reveals an abundance of promiscuous resistance determinants in *Escherichia coli*. *Proc. Natl Acad. Sci. USA* **108**, 1484–1489. (doi:10.1073/pnas.1012108108)
- 79 Taylor, M. & Feyereisen, R. 1996 Molecular biology and evolution of resistance of toxicants. *Mol. Biol. Evol.* **13**, 719–734. (doi:10.1093/oxfordjournals.molbev.a025633)
- 80 Lenormand, T., Guillemaud, T., Bourguet, D. & Raymond, M. 1998 Appearance and sweep of a gene duplication: adaptive response and potential for new functions in the mosquito *Culex pipiens*. *Evolution* **52**, 1705–1712. (doi:10.2307/2411343)
- 81 Labbé, P., Berticat, C., Berthomieu, A., Unal, S., Bernard, C., Weill, M. & Lenormand, T. 2007 Forty years of erratic insecticide resistance evolution in the mosquito *Culex pipiens*. *PLoS Genet.* **3**, e205. (doi:10.1371/journal.pgen.0030205)
- 82 Labbé, P., Berthomieu, A., Berticat, C., Alout, H., Raymond, M., Lenormand, T. & Weill, M. 2007 Independent duplications of the acetylcholinesterase gene conferring insecticide resistance in the mosquito *Culex pipiens*. *Mol. Biol. Evol.* **24**, 1056–1067. (doi:10.1093/molbev/msm025)
- 83 Alout, H., Labbé, P., Pasteur, N. & Weill, M. 2011 High incidence of ace-1 duplicated haplotypes in resistant *Culex pipiens* mosquitoes from Algeria. *Insect Biochem. Mol. Biol.* **41**, 29–35. (doi:10.1016/j.ibmb.2010.09.009)
- 84 Larson, J. H., Marron, B. M., Beever, J. E., Roe, B. A. & Lewin, H. A. 2006 Genomic organization and evolution of the ULBP genes in cattle. *BMC Genomics* **7**, 227. (doi:10.1186/1471-2164-7-227)
- 85 Bickhart, D. M. *et al.* 2012 Copy number variation of individual cattle genomes using next-generation sequencing. *Genome Res.* **22**, 778–790. (doi:10.1101/gr.133967.111)

- 86 Liu, G. E., Ventura, M., Cellamare, A., Chen, L., Cheng, Z., Zhu, B., Li, C., Song, J. & Eichler, E. E. 2009 Analysis of recent segmental duplications in the bovine genome. *BMC Genomics* **10**, 571. (doi:10.1186/1471-2164-10-571)
- 87 Gonzalez, E. *et al.* 2005 The influence of CCL3L1 gene-containing segmental duplications on HIV-1/AIDS susceptibility. *Science* **307**, 1434–1440. (doi:10.1126/science.1101160)
- 88 Perry, G. H. *et al.* 2007 Diet and the evolution of human amylase gene copy number variation. *Nat. Genet.* **39**, 1256–1260. (doi:10.1038/ng2123)
- 89 Cooper, G. M., Nickerson, D. A. & Eichler, E. E. 2007 Mutational and selective effects on copy-number variants in the human genome. *Nat. Genet.* **39**, S22–S29. (doi:10.1038/ng2054)
- 90 Emerson, J. J., Cardoso-Moreira, M., Borevitz, J. O. & Long, M. 2008 Natural selection shapes genome-wide patterns of copy-number polymorphism in *Drosophila melanogaster*. *Science* **320**, 1629–1631. (doi:10.1126/science.1158078)
- 91 Gazave, E. *et al.* 2011 Copy number variation analysis in the great apes reveals species-specific patterns of structural variation. *Genome Res.* **21**, 1626–1639. (doi:10.1101/gr.117242.110)
- 92 Hastings, P. J., Bull, H. J., Klump, J. R. & Rosenberg, S. M. 2000 Adaptive amplification: an inducible chromosomal instability mechanism. *Cell* **103**, 723–731. (doi:10.1016/S0092-8674(00)00176-8)
- 93 Zhong, S., Khodursky, A., Dykhuizen, D. E. & Dean, A. M. 2004 Evolutionary genomics of ecological specialization. *Proc. Natl Acad. Sci. USA* **101**, 11 719–11 724. (doi:10.1073/pnas.0404397101)
- 94 Lin, D., Gibson, I. B., Moore, J. M., Thornton, P. C., Leal, S. M. & Hastings, P. J. 2011 Global chromosomal structural instability in a subpopulation of starving *Escherichia coli* cells. *PLoS Genet.* **7**, e1002223. (doi:10.1371/journal.pgen.1002223)
- 95 Xu, J. H. & Messing, J. 2009 Amplification of prolamin storage protein genes in different subfamilies of the Poaceae. *Theor. Appl. Genet.* **119**, 1397–1412. (doi:10.1007/s00122-009-1143-x)
- 96 Desjardins, M., Graham, L. A., Davies, P. L. & Fletcher, G. L. 2012 Antifreeze protein gene amplification facilitated niche exploitation and speciation in wolffish. *FEBS J.* **279**, 2215–2230. (doi:10.1111/j.1742-4658.2012.08605.x)
- 97 Sandve, S. R., Rudi, H., Asp, T. & Rognli, O. A. 2008 Tracking the evolution of a cold stress associated gene family in cold tolerant grasses. *BMC Evol. Biol.* **8**, 245. (doi:10.1186/1471-2148-8-245)
- 98 Zhou, D. *et al.* 2009 Duplication and adaptive evolution of the COR15 genes within the highly cold-tolerant *Draba* lineage (Brassicaceae). *Gene* **441**, 36–44. (doi:10.1016/j.gene.2008.06.024)
- 99 Lind, P. A., Tobin, C., Berg, O. G., Kurland, C. G. & Andersson, D. I. 2010 Compensatory gene amplification restores fitness after inter-species gene replacements. *Mol. Microbiol.* **75**, 1078–1089. (doi:10.1111/j.1365-2958.2009.07030.x)
- 100 Widholm, J. M., Chinnala, A. R., Ryu, J. H., Song, H. S., Eggett, T. & Brotherton, J. E. 2001 Glyphosate selection of gene amplification in suspension cultures of 3 plant species. *Physiol. Plant* **112**, 540–545. (doi:10.1034/j.1399-3054.2001.1120411.x)
- 101 Gaines, T. A. *et al.* 2010 Gene amplification confers glyphosate resistance in *Amaranthus palmeri*. *Proc. Natl Acad. Sci. USA* **107**, 1029–1034. (doi:10.1073/pnas.0906649107)
- 102 Gaines, T. A., Shaner, D. L., Ward, S. M., Leach, J. E., Preston, C. & Westra, P. 2011 Mechanism of resistance of evolved glyphosate-resistant *Palmer amaranth* (*Amaranthus palmeri*). *J. Agric. Food. Chem.* **59**, 5886–5889. (doi:10.1021/jf104719k)
- 103 Eanes, W. F., Merritt, T. J., Flowers, J. M., Kumagai, S. & Zhu, C. T. 2009 Direct evidence that genetic variation in glycerol-3-phosphate and malate dehydrogenase genes (*Gpdh* and *Mdh1*) affects adult ethanol tolerance in *Drosophila melanogaster*. *Genetics* **181**, 607–614. (doi:10.1534/genetics.108.089383)
- 104 Wapinski, I., Pfeffer, A., Friedman, N. & Regev, A. 2007 Natural history and evolutionary principles of gene duplication in fungi. *Nature* **449**, 54–61. (doi:10.1038/nature06107)
- 105 Hanada, K., Zou, C., Lehti-Shiu, M. D., Shinozaki, K. & Shiu, S. H. 2008 Importance of lineage-specific expansion of plant tandem duplicates in the adaptive response to environmental stimuli. *Plant Physiol.* **148**, 993–1003. (doi:10.1104/pp.108.122457)
- 106 Kuo, C. H. & Kissinger, J. C. 2008 Consistent and contrasting properties of lineage-specific genes in the apicomplexan parasites *Plasmodium* and *Theileria*. *BMC Evol. Biol.* **8**, 108. (doi:10.1186/1471-2148-8-108)
- 107 Hoffmann, F. G., Opazo, J. C. & Storz, J. F. 2008 Rapid rates of lineage-specific gene duplication and deletion in the alpha-globin gene family. *Mol. Biol. Evol.* **25**, 591–602. (doi:10.1093/molbev/msn004)
- 108 Cornman, R. S. & Willis, J. H. 2008 Extensive gene amplification and concerted evolution within the CPR family of cuticular proteins in mosquitoes. *Insect Biochem. Mol. Biol.* **38**, 661–676. (doi:10.1016/j.ibmb.2008.04.001)
- 109 Lehti-Shiu, M. D., Zou, C., Hanada, K. & Shiu, S. H. 2009 Evolutionary history and stress regulation of plant receptor-like kinase/pelle genes. *Plant Physiol.* **150**, 12–26. (doi:10.1104/pp.108.134353)
- 110 Jiang, S. Y., Ma, Z. & Ramachandran, S. 2010 Evolutionary history and stress regulation of the lectin superfamily in higher plants. *BMC Evol. Biol.* **10**, 79. (doi:10.1186/1471-2148-10-79)
- 111 Gong, D., Wilson, P. W., Bain, M. M., McDade, K., Kalina, J., Hervé-Grépinet, V., Nys, Y. & Dunn, I. C. 2010 Gallin; an antimicrobial peptide member of a new avian defensin family, the ovodefensins, has been subject to recent gene duplication. *BMC Immunol.* **11**, 12. (doi:10.1186/1471-2172-11-12)
- 112 Ames, R. M., Rash, B. M., Hentges, K. E., Robertson, D. L., Delneri, D. & Lovell, S. C. 2010 Gene duplication and environmental adaptation within yeast populations. *Genome Biol. Evol.* **2**, 591–601. (doi:10.1093/gbe/evq043)
- 113 Miller, S. R., Wood, A. M., Blankenship, R. E., Kim, M. & Ferreira, S. 2011 Dynamics of gene duplication in the genomes of chlorophyll d-producing cyanobacteria: implications for the ecological niche. *Genome Biol. Evol.* **3**, 601–613. (doi:10.1093/gbe/evr060)
- 114 Colbourne, J. K. *et al.* 2011 The coresponsive genome of *Daphnia pulex*. *Science* **331**, 555–561. (doi:10.1126/science.1197761)
- 115 Xu, J. H., Bennetzen, J. L. & Messing, J. 2012 Dynamic gene copy number variation in collinear regions of grass genomes. *Mol. Biol. Evol.* **29**, 861–871. (doi:10.1093/molbev/msr261)
- 116 Wong, E. S. & Belov, K. 2012 Venom evolution through gene duplications. *Gene* **496**, 1–7. (doi:10.1016/j.gene.2012.01.009)
- 117 Nei, M., Niimura, Y. & Nozawa, M. 2008 The evolution of animal chemosensory receptor gene

- repertoires: roles of chance and necessity. *Nat. Rev. Genet.* **9**, 951–963. (doi:10.1038/nrg2480)
- 118 Niimura, Y. 2009 Evolutionary dynamics of olfactory receptor genes in chordates: interaction between environments and genomic contents. *Hum. Genomics* **4**, 107–118. (doi:10.1186/1479-7364-4-2-107)
- 119 Lavagnino, N., Serra, F., Arbiza, L., Dopazo, H. & Hasson, E. 2012 Evolutionary genomics of genes involved in olfactory behavior in the *Drosophila melanogaster* species group. *Evol. Bioinform. Online* **8**, 89–104. (doi:10.4137/EBO.S8484)
- 120 Sison-Mangus, M. P., Bernard, G. D., Lampel, J. & Briscoe, A. D. 2006 Beauty in the eye of the beholder: the two blue opsins of lycaenid butterflies and the opsin gene-driven evolution of sexually dimorphic eyes. *J. Exp. Biol.* **209**, 3079–3090. (doi:10.1242/jeb.02360)
- 121 Gojobori, J. & Innan, H. 2009 Potential of fish opsin gene duplications to evolve new adaptive functions. *Trends Genet.* **25**, 198–202. (doi:10.1016/j.tig.2009.03.008)
- 122 Schulenburg, H. & Boehnisch, C. 2008 Diversification and adaptive sequence evolution of *Caenorhabditis* lysozymes (Nematoda: Rhabditidae). *BMC Evol. Biol.* **8**, 114. (doi:10.1186/1471-2148-8-114)
- 123 Neiman, M., Olson, M. S. & Tiffin, P. 2009 Selective histories of poplar protease inhibitors: elevated polymorphism, purifying selection, and positive selection driving divergence of recent duplicates. *New Phytol.* **183**, 740–750. (doi:10.1111/j.1469-8137.2009.02936.x)
- 124 Powell, A. J., Conant, G. C., Brown, D. E., Carbone, I. & Dean, R. A. 2008 Altered patterns of gene duplication and differential gene gain and loss in fungal pathogens. *BMC Genomics* **9**, 147. (doi:10.1186/1471-2164-9-147)
- 125 Fischer, I., Camus-Kulandaivelu, L., Allal, F. & Stephan, W. 2011 Adaptation to drought in two wild tomato species: the evolution of the *Asr* gene family. *New Phytol.* **190**, 1032–1044. (doi:10.1111/j.1469-8137.2011.03648.x)
- 126 Storz, J. F., Runck, A. M., Sabatino, S. J., Kelly, J. K., Ferrand, N., Moriyama, H., Weber, R. E. & Fago, A. 2009 Evolutionary and functional insights into the mechanism underlying high-altitude adaptation of deer mouse hemoglobin. *Proc. Natl Acad. Sci. USA* **106**, 14 450–14 455. (doi:10.1073/pnas.0905224106)
- 127 Duda Jr, T. F. & Palumbi, S. R. 1999 Molecular genetics of ecological diversification: duplication and rapid evolution of toxin genes of the venomous gastropod *Conus*. *Proc. Natl Acad. Sci. USA* **96**, 6820–6823. (doi:10.1073/pnas.96.12.6820)
- 128 Moran, Y. & Gurevitz, M. 2006 When positive selection of neurotoxin genes is missing. The riddle of the sea anemone *Nematostella vectensis*. *FEBS J.* **273**, 3886–3892. (doi:10.1111/j.1742-4658.2006.05397.x)
- 129 Lynch, V. J. 2007 Inventing an arsenal: adaptive evolution and neofunctionalization of snake venom phospholipase A2 genes. *BMC Evol. Biol.* **7**, 2. (doi:10.1186/1471-2148-7-2)
- 130 Murray, S. A., Mihali, T. K. & Neilan, B. A. 2011 Extraordinary conservation, gene loss, and positive selection in the evolution of an ancient neurotoxin. *Mol. Biol. Evol.* **28**, 1173–1182. (doi:10.1093/molbev/msq295)
- 131 Ma, Y., He, Y., Zhao, R., Wu, Y., Li, W. & Cao, Z. 2012 Extreme diversity of scorpion venom peptides and proteins revealed by transcriptomic analysis: implication for proteome evolution of scorpion venom arsenal. *J. Proteomics* **75**, 1563–1576. (doi:10.1016/j.jprot.2011.11.029)
- 132 Chang, D. & Duda Jr, T. F. 2012 Extensive and continuous duplication facilitates rapid evolution and diversification of gene families. *Mol. Biol. Evol.* **29**, 2019–2029. (doi:10.1093/molbev/mss068)
- 133 Wang, L., Beuerle, T., Timbilla, J. & Ober, D. 2012 Independent recruitment of a flavin-dependent mono-oxygenase for safe accumulation of sequestered pyrrolizidine alkaloids in grasshoppers and moths. *PLoS ONE* **7**, e31796. (doi:10.1371/journal.pone.0031796)
- 134 Shigenobu, S. *et al.* 2010 Comprehensive survey of developmental genes in the pea aphid, *Acyrtosiphon pisum*: frequent lineage-specific duplications and losses of developmental genes. *Insect Mol. Biol.* **19**(Suppl. 2), 47–62. (doi:10.1111/j.1365-2583.2009.00944.x)
- 135 Moran, N. A. & Jarvik, T. 2010 Lateral transfer of genes from fungi underlies carotenoid production in aphids. *Science* **328**, 624–627. (doi:10.1126/science.1187113)
- 136 Day, J. C., Goodall, T. I. & Bailey, M. J. 2009 The evolution of the adenylate-forming protein family in beetles: multiple luciferase gene paralogues in fireflies and glow-worms. *Mol. Phylogenet. Evol.* **50**, 93–101. (doi:10.1016/j.ympev.2008.09.026)
- 137 Kelleher, E. S. & Markow, T. A. 2009 Duplication, selection and gene conversion in a *Drosophila mojavensis* female reproductive protein family. *Genetics* **181**, 1451–1465. (doi:10.1534/genetics.108.099044)
- 138 Beckmann, J. S., Estivill, X. & Antonarakis, S. E. 2007 Copy number variants and genetic traits: closer to the resolution of phenotypic to genotypic variability. *Nat. Rev. Genet.* **8**, 639–646. (doi:10.1038/nrg2149)
- 139 Andolfatto, P. 2005 Adaptive evolution of non-coding DNA in *Drosophila*. *Nature* **437**, 1149–1152. (doi:10.1038/nature04107)
- 140 Bergthorsson, U., Andersson, D. I. & Roth, J. R. 2007 Ohno's dilemma: evolution of new genes under continuous selection. *Proc. Natl Acad. Sci. USA* **104**, 17 004–17 009. (doi:10.1073/pnas.0707158104)
- 141 Andersson, D. I. & Hughes, D. 2009 Gene amplification and adaptive evolution in bacteria. *Annu. Rev. Genet.* **43**, 167–195. (doi:10.1146/annurev-genet-102108-134805)
- 142 Levasseur, A. & Pontarotti, P. 2011 The role of duplications in the evolution of genomes highlights the need for evolutionary-based approaches in comparative genomics. *Biol. Direct.* **6**, 11. (doi:10.1186/1745-6150-6-11)
- 143 Holloway, A. K., Palzkill, T. & Bull, J. J. 2007 Experimental evolution of gene duplicates in a bacterial plasmid model. *J. Mol. Evol.* **64**, 215–222. (doi:10.1007/s00239-006-0087-x)
- 144 Reams, A. B., Kofoid, E., Savageau, M. & Roth, J. R. 2010 Duplication frequency in a population of *Salmonella enterica* rapidly approaches steady state with or without recombination. *Genetics* **184**, 1077–1094. (doi:10.1534/genetics.109.111963)