need to be borne in mind (Box 1), a quantitative genetic framework has much to offer in terms of our understanding of similar problems in evolutionary ecology, because it offers a simple, coherent way to combine measured selection and expected response. This is likely to prove even more the case as the influence of genomics is felt ever more strongly in organismic biology (e.g. [4,16]). For example, the question of genetic linkage between male trait, female choice and fitness in wild populations could be addressed using existing genetic mapping approaches, which have already been applied, at least as proof of concept, to wild birds [17,18].

Acknowledgements

We thank Anna Qvarnström and Jon Brommer for discussion and four anonymous referees for their helpful comments; A.C. was funded by a Biotechnology and Biological Sciences Research Council (BBSRC) grant to B.C.S. and L.E.B. Kruuk.

References

1 Andersson, M. (1994) *Sexual Selection,* Princeton University Press
2 Kokko, H. *et al.* (2003) The evolution of mate choice and mating biases. *Proc. R. Soc. B* 270, 653–664
3 Kirkpatrick, M. and Ryan, M.J. (1991) The evolution of mating preferences and the paradox of the lek. *Nature* 350, 33–38
4 Andersson, M. and Simmons, L.W. (2006) Sexual selection and mate choice. *Trends Ecol. Evol.* 21, 296–302
5 Head, M.L. *et al.* (2005) The indirect benefits of mating with attractive males outweigh the direct costs. *PloS Biol.* 3, 289–294
6 Kirkpatrick, M. and Barton, N.H. (1997) The strength of indirect selection on female mating preferences. *Proc. Natl. Acad. Sci. U. S. A.* 94, 1282–1286
7 Arnqvist, G. and Kirkpatrick, M. (2005) The evolution of infidelity in socially monogamous passerines: the strength of direct and indirect selection on extrapair copulation behavior in females. *Am. Nat.* 165, S26–S37
8 Qvarnström, A. *et al.* (2006) Testing the genetics underlying the co-evolution of mate choice and ornament in the wild. *Nature* 441, 84–86
9 Hadfield, J. *et al.* (2006) Direct versus indirect sexual selection: genetic basis of colour, size and recruitment in a wild bird. *Proc. R. Soc. B* 273, 1347–1353
10 Griffith, S.C. *et al.* (2002) Extra pair paternity in birds: a review of interspecific variation and adaptive function. *Mol. Ecol.* 11, 2195–2212
11 Kempenaers, B. and Sheldon, B.C. (1997) Studying paternity and paternal care: pitfalls and problems. *Anim. Behav.* 53, 423–427
12 Arnqvist, G. and Rowe, L. (2005) *Sexual Conflict,* Princeton University Press
13 Whittingham, L. and Dunn, P.O. (2005) Effects of extra-pair and within-pair reproductive success on the opportunity for selection in birds. *Behav. Ecol.* 16, 138–144
14 Steppan, S.J. *et al.* (2002) Comparative quantitative genetics: evolution of the G matrix. *Trends Ecol. Evol.* 17, 320–327
15 Merilä, J. *et al.* (2001) Explaining stasis: microevolutionary studies of natural populations. *Genetica* 112, 199–222
16 Colosimo, P.F. *et al.* (2005) Widespread parallel evolution in sticklebacks by repeated fixation of ectodysplasin alleles. *Science* 307, 1928–1933
17 Hansson, B. *et al.* (2006) Linkage mapping reveals sex-dimorphic map distances in a passerine bird. *Proc. R. Soc. B* 272, 2289–2298
18 Dawson, D.A. *et al.* (2006) A predicted microsatellite map of the passerine genome based on chicken–passerine sequence similarity. *Mol. Ecol.* 15, 1299–1320
19 Owens, I.P.F. (2006) Where is behavioural ecology going? *Trends Ecol. Evol.* 21, 356–361

0169-5347/$ – see front matter doi:10.1016/j.tree.2006.06.001

Are network motifs the spandrels of cellular complexity?

Ricard V. Solé[1,2] and Sergi Valverde[1]

[1] ICREA-Complex Systems Lab, Universitat Pompeu Fabra, Dr. Aiguader 80, 08003 Barcelona, Spain
[2] Santa Fe Institute, 1399 Hyde Park Road, Santa Fe, NM 87501, USA

Cellular networks display modular organization at different levels, from small sets of genes exchanging signals in morphogenesis to large groups of proteins involved in cell death. At the smallest scale, minute groups of interacting proteins or genes, so-called 'network motifs', have been suggested to be the functional building blocks of network biology. In this context, the relative abundance of a network motif would reflect its adaptive value. However, although the over-abundance of motifs is non-random, recent studies by Mazurie *et al.* and by Kuo *et al.* show that motif abundance does not reflect their true adaptive value. Just as some architectural components emerge as a byproduct of a prior decision, common motifs might be a side effect of inevitable rules of genome growth and change.

Corresponding author: Solé, R.V. (ricard.sole@upf.edu).
Available online 9 June 2006

Motifs and modules

Life is modular [1] and selection for function has shaped cellular networks, such as metabolic pathways or gene regulatory webs, to perform efficiently. Modular structures, where proteins, genes or metabolites can interact, have been identified at different scales: for example, groups of two or three cellular components, particularly proteins, can enable the implementation of well-defined gene circuits [2]. The intrinsic properties of these basic modules are suitable for theoretical analysis and many insights have been obtained from such systems biology approaches [3].

In this context, identifying modular structures with no previous knowledge of their exact function has become a hot topic over the past decade. One popular approach is the so-called 'network motif analysis', in which complex networks are dissected into small subgraphs comprising three

or four interacting units, or so-called 'motifs' [4,5]. Broadly speaking, motifs are small, repeated and conserved biological units ranging from molecular domains to small reaction networks [3]. A remarkable trait of such motifs is their frequency in a given network: some are very common, whereas others are much less so than would be expected by chance [4,5]. A functional, adaptive interpretation for these deviations seems to be a reasonable conclusion. But is this the case? Recent work using available gene databases [6] and *in silico* gene network models [7] appear to suggest a different interpretation.

Selection versus constraints

If selection proceeds by choosing specific motifs to improve or achieve a given function, then we should expect the abundance of a motif in a network to reflect its adaptive value, at least when compared with randomly wired networks with the same number of elements and links. Random networks can be obtained by linking any pair of elements (e.g. genes or proteins) with some (small) given probability. However, cellular networks are not generated through such a process. Instead, they grow and change through the duplication of elements already present or, as François Jacob pointed out, as a result of 'tinkering'. After duplication, mutations within the network (affecting, for example, the binding of one protein to another) cause it to change, enabling new functions to evolve. Such mutations affect the links among elements of the network, either creating or deleting them (Figure 1). But how many of the deviations in motif abundances (compared with random models) are associated with the rules of network growth?

In a recent study [7], Mazurie *et al.* performed phylogenetic comparative analyses of the genes within motifs, and concluded that motifs cannot be isolated from the cellular context because they are usually part of larger sub-webs that include many components. This work shows that network motifs are not subject to any particular evolutionary pressure to be preserved. Such a result is consistent with other studies indicating that higher-order interconnection patterns, integrating different regulatory mechanisms, should be considered to address functionally meaningful cellular properties [8,9]. In this context, all the examples studied by Mazurie *et al.* reveal that the receiving and processing of cellular signals cannot be reduced to simple protein–protein interactions.

Neutral models

An alternative test to this non-functional interpretation of motif abundances can be obtained by simulating the evolution of regulatory networks by gene duplication and re-wiring [6]. Kuo *et al.* recently explored this approach in detail by extending previous work on artificial transcriptional regulatory networks lacking real functionality [10]. In their model, a linear string of bits (the genome) is formed by several finite chains: the genes. These genes encode proteins that recognize specific 'regulatory' sites. Random duplications and mutations enable this genome to grow and genes to interact in a changing network. Although no true functionality is introduced, the resulting motif abundances match those found in real cells.

Network spandrels

Kuo *et al.*'s results indicate that the abundance of network motifs should be seen as a byproduct of the network construction process. When looking at the network organization, we are likely to recognize some forms

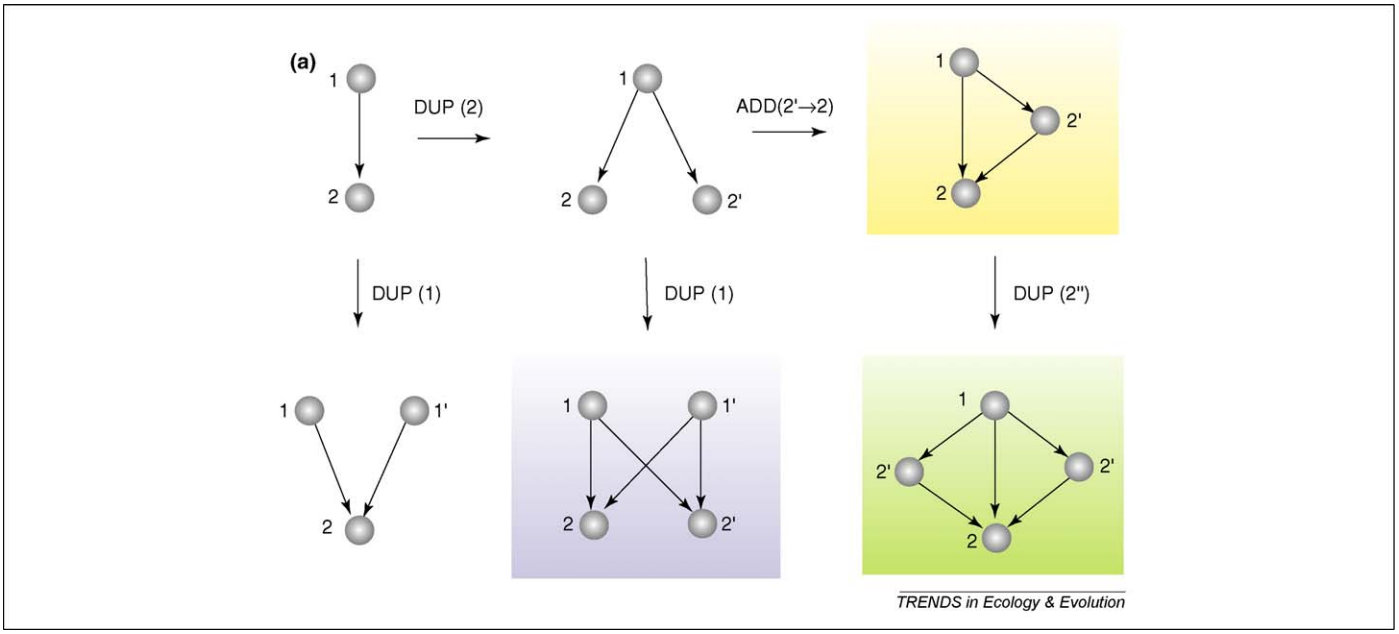

Figure 1. Network motifs from 'tinkered' evolution. Motifs are identified as small groups of interacting elements with a well-defined arrangement of links. These links can correspond to physical interactions among proteins or to regulatory links among genes. In this figure, grey balls and arrows indicate units and their interactions, respectively. Starting from the simplest graph **(a)** we can generate different motifs by gene duplication (DUP) or link addition (ADD). Three common motifs found in cellular networks are highlighted by the colored boxes.

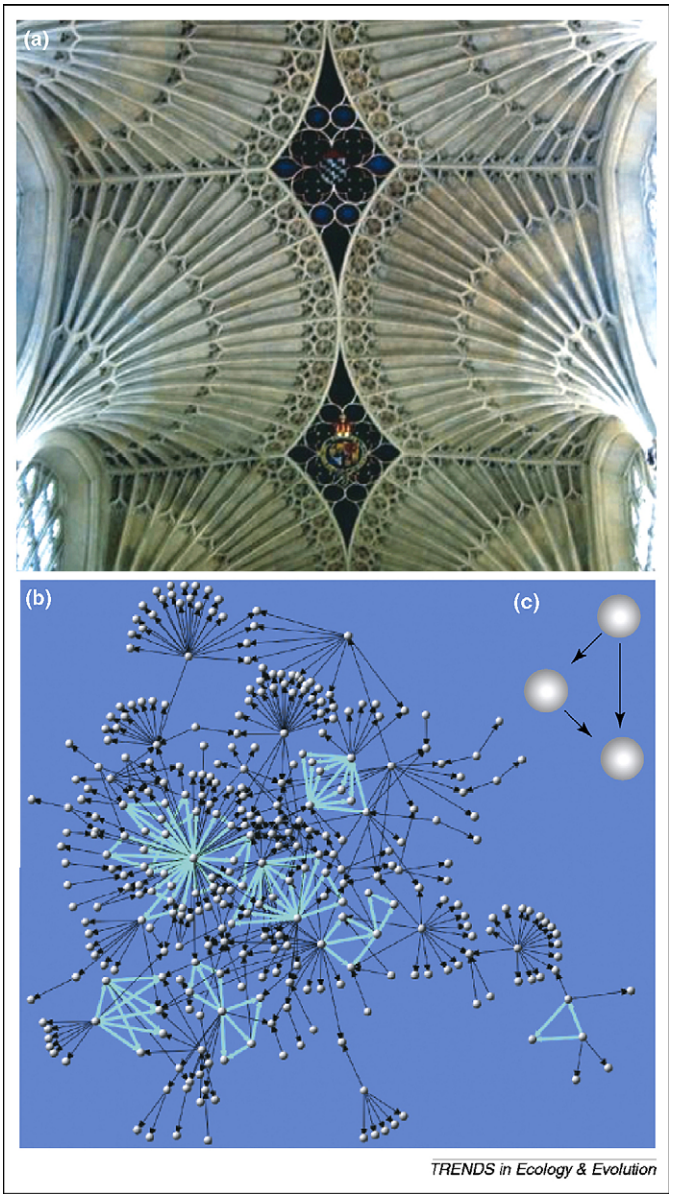

Figure 2. Architectural and evolved spandrels. **(a)** The ceiling of Bath Abbey in the UK. Spandrels are the darker, profusely decorated structures in the middle of four rounded arches. The network shown in **(b)** is the gene regulatory map of *Escherichia coli*. Here, balls indicate genes, and arrows point from a regulatory to a regulated gene. In **(c)**, a specific motif is shown that is present within the network at multiple locations, indicated in (b) by means of lighter, thicker arrows. Most appear to be aggregated, forming larger structures.

and shapes that can be interpreted as resulting from selective pressures. These are examples of what Stephen Jay Gould and Richard Lewontin [11] called 'spandrels'. Architecturally, a spandrel is the space between two arches or between an arch and a rectangular enclosure (Figure 2a). In evolutionary biology, a spandrel is a phenotypic characteristic that evolved as a side effect of a true adaptation. We can summarize the features of evolutionary spandrels as follows: (i) they are the byproduct of building rules; (ii) they have intrinsic, well-defined, non-random features; and (iii) their structure reveals some of the underlying rules of construction.

In Figure 2a, we can see that architectural spandrels are well-defined, non-random structures that are the result of a prior decision, namely to mount a dome on four rounded arches, so arrayed [12]. Moreover, their geometric shape is fully constrained by the dominant arches. Similarly, the 'tinkered' evolution of cellular networks generates well-defined structures as a result of the underlying growth rules [13]. Thus, motif abundances are the spandrels of network biocomplexity, because: (i) their abundance is matched by *in silico* models lacking real functionality, and are thus a by-product of the network building rules; (ii) they exhibit highly non-random features at several scales, which are particularly obvious when looking at the way in which motifs form clusters (Figure 2b); and (iii) the patterns of motif aggregates indicate that duplication-rewiring processes, which are known to generate the structure of the whole web, are also responsible for the presence of such clusters. The internal organization of these clusters has special features (Figure 2b) that are easily interpreted under the light of the duplication mechanisms shown in Figure 1. Thus, motif abundances are the byproduct of extensive re-use of existing motifs and their frequency and distribution within a given network have no adaptive meaning per se. These results are consistent with previous work on both biological and technological network evolution [14–16].

Conclusions

Networks pervade biology at multiple scales [17]. Although the observation of common subgraphs within networks is interesting, it should be viewed in the appropriate context: selection might have taken advantage of the network structure *a posteriori*. As Gould said: 'although spandrels must originate as necessary side-consequences of an architectural decision, and not as forms explicitly chosen to serve a purpose, they still exist in undeniable abundance, and can therefore be secondarily used in important and interesting ways' [12]. The reliable nature of cellular circuits might well be an example of how some abundant motifs can be used to process information under the presence of noise [18]. In this context, new motifs emerging as a consequence of random gene duplications (such as those highlighted in Figure 1) are known to reduce efficiently internal fluctuations and delays in cellular responses.

Acknowledgments

We thank David Hales for helpful discussions. This work was supported by the Santa Fe Institute and grants FIS2004-2154, and EU DELIS grant under contract 001907.

References

1 Hartwell, L.H. *et al.* (1999) From molecular to modular cell biology. *Nature* 402, C47–C52
2 Wolf, D.M. and Arkin, A.P. (2003) Motifs, modules and games in bacteria. *Curr. Opin. Microbiol.* 6, 125–134
3 Bintu, L. *et al.* (2005) Transcriptional regulation by the numbers: models. *Curr. Opin. Genet. Dev.* 15, 116–124
4 Milo, R. *et al.* (2002) Network motifs: simple building blocks of complex networks. *Science* 298, 824–827
5 Milo, R. *et al.* (2004) Superfamilies of designed and evolved networks. *Science* 303, 1538–1542

6 Kuo, P.D. *et al.* Network topology and the evolution of dynamics in an artificial regulatory network model created by whole genome duplication and divergence. *Biosystems* (in press).
7 Mazurie, A. *et al.* (2005) An evolutionary and functional assessment of regulatory network motifs. *Genome Biol.* 6, R35
8 Dobrin, R. *et al.* (2004) Aggregation of topological motifs in the *Escherichia coli* transcriptional regulatory network. *BMC Bioinform.* 5, 10
9 Zhang, L.V. *et al.* (2005) Motifs, themes and thematic maps of an integrated *Saccharomyces cerevisiae* interaction network. *J. Biol.* 4, 6
10 Banzhaf, W. and Kuo, P.D. (2004) Network motifs in natural and artificial transcriptional regulatory networks. *J. Biol. Phys. Chem.* 4, 85–92
11 Gould, S.J. and Lewontin, R.C. (1979) The Spandrels of San Marco and the Panglossian Paradigm: a critique of the Adaptationist Programme. *Proc. R. Soc. B* 205, 581–598
12 Gould, S.J. (2002) *The Structure of Evolutionary Theory,* Harvard University Press
13 Solé, R.V. *et al.* (2002) Selection, tinkering and emergence in complex networks. *Complexity* 8, 20–33
14 Wagner, A. (2003) Does natural selection mold molecular networks? *Science STKE* 202, 41–43
15 Solé, R.V. *et al.* (2002) A model of large-scale proteome evolution. *Adv. Complex Syst.* 5, 43–54
16 Valverde, S. and Solé, R.V. (2005) Network motifs in computational graphs: a case study in software architecture. *Phys. Rev. E* 72, 26107
17 Proulx, S.R. *et al.* (2005) Network thinking in ecology and evolution. *Trends Ecol. Evol.* 20, 345–353
18 Klemm, K. and Bornholdt, S. (2005) Topology of biological networks and reliability of information processing. *Proc. Natl. Acad. Sci. U. S. A.* 102, 18414–18419

0169-5347/$ – see front matter doi:10.1016/j.tree.2006.05.013

Letters

The *h* index: playing the numbers game

Andy Purvis

Division of Biology, Imperial College London, Silwood Park campus, Ascot, UK, SL5 7PY

The '*h* index' was developed recently as a measure of research performance [1]: a researcher's *h* is the number of his or her papers that have been cited at least *h* times. In their thoughtful critique of the index, Kelly and Jennions [2] point out many ways in which *h* is no better than 'traditional' bibliometrics, such as total citation counts. However, there is one way in which, for researchers, it could be very much better, especially if (as Hirsch suggests [1]) it is to inform hiring and promotion decisions. The skewed nature of the distribution of citations among publications means that most researchers have several papers that nearly but not quite count. Consequently, *h* can be distorted much more easily than can total citation count just by finding a subtle way to cite one's own papers that are 'bubbling under'. Incidentally, bats show broadly the same life-history allometries as other mammalian clades [3].

References

1 Hirsch, J.E. (2005) An index to quantify an individual's scientific research output. *Proc. Natl. Acad. Sci. U. S. A.* 102, 16569–16572
2 Kelly, C.D. and Jennions, M.D. (2006) The *h* index and career assessment by numbers. *Trends Ecol. Evol.* 21, 167–170
3 Jones, K.E. and Purvis, A. (1997) An optimal body size for mammals? Comparative evidence from bats *Funct. Ecol.* 11, 751–756

0169-5347/$ – see front matter doi:10.1016/j.tree.2006.05.014

Corresponding author: Purvis, A. (a.purvis@imperial.ac.uk).
Available online 14 June 2006