FEBS Letters 582 (2008) 1225–1230

Minireview

Co-evolution and co-adaptation in protein networks

David Juan[a,1], Florencio Pazos[b,2], Alfonso Valencia[a,*]

[a] Structural Bioinformatics Group, Spanish National Cancer Research Centre (CNIO), C/Melchor Fernández Almagro 3, 28029 Madrid, Spain
[b] Computational Systems Biology Group, National Centre for Biotechnology (CNB-CSIC), C/Darwin 3, Cantoblanco, 28049 Madrid, Spain

Received 9 January 2008; accepted 8 February 2008

Available online 20 February 2008

Edited by Robert B. Russell and Patrick Aloy

Abstract **Interacting or functionally related proteins have been repeatedly shown to have similar phylogenetic trees. Two main hypotheses have been proposed to explain this fact. One involves compensatory changes between the two protein families (co-adaptation). The other states that the tree similarity may be an indirect consequence of the involvement of the two proteins in similar cellular process, which in turn would be reflected by similar evolutionary pressure on the corresponding sequences. There are published data supporting both propositions, and currently the available information is compatible with both hypotheses being true, in an scenario in which both sets of forces are shaping the tree similarity at different levels.**

Keywords: Co-evolution; Co-adaptation; Protein interaction; Mirrortree

1. Introduction

Proteins rarely act in isolation and their biological roles can only be fully understood in the context of their complex interactions with others. One of the prototypic elements studied in Systems Biology is the "interactome" [1–4], the network of interactions and functional relationships between the components of a proteome. The study of the interactome from a systemic ("top-down") perspective has provided important biological information not evident in the properties of the individual proteins [5–8].

One of the most important, yet still poorly understood, phenomenon related to protein interactions is the similar evolutionary paths typically followed by interacting and/or functionally related proteins (co-evolution[3]). This is an interesting theoretical problem that, as argued in the last section of this review, also has important practical consequences for the planning of mutagenesis experiments and for the design of protein interaction prediction algorithms.

Two general hypotheses have been proposed to explain the similarity observed in the evolutionary history of interacting proteins. One states that the observed co-evolution is a consequence of the similar evolutionary pressure exerted on interacting and functionally related proteins due to the similar control mechanisms that act on them, for example concerted transcription and regulation. According to this hypothesis, the observed similarity of evolutionary histories deduced from the corresponding sequences would not be a direct consequence of a specific physical interaction between these proteins. Thus, it may in principle be similar for groups of proteins that share a functional relationship, such as those involved in the same biochemical pathway or in the same macromolecular complex.

The alternative hypothesis is that the observed co-evolution is directly related to the co-adaptation of interacting proteins. A physical model that might underlie this process could imply that changes which decrease a proteins stability or capacity to fold correctly are compensated by changes in the interacting partner in order to maintain the complex functional. Or, more properly expressed, complexes that are functional are selected if deleterious mutations have been properly compensated (Fig. 1). Indeed, this model is related to the proposed co-variation model of compensation [10].

In this minireview, we will summarize the bibliography on co-evolution at the molecular level, and the studies that support one or the other hypotheses. We will also discuss the kind of efforts that would be required to carry out conclusive experiments, and the potential consequences of resolving the contribution of the physical versus general functional constraints in the evolution of protein complexes and interaction networks.

2. Co-evolution and protein interactions

The co-evolution of certain features in the sequences of proteins that interact at the functional and/or physical level is well established. Co-evolution at the molecular level has been repeatedly demonstrated in a variety of scenarios by different authors, including studies of molecular co-evolution at the residue and complete protein level. Note that, in this section, we describe the co-evolutionary features observed without going into their possible causes, which will be discussed in the next section.

*Corresponding author. Fax: +34 912246980.
E-mail addresses: dadejuan@cnio.es (D. Juan), pazos@cnb.uam.es (F. Pazos), valencia@cnio.es (A. Valencia).

[1]Fax: +34 912246980.
[2]Fax: +34 915854506.

[3]Here we will use the term "co-evolution" to refer to the similarity of evolutionary histories, which is an observable and can be quantified, while we will refer to "co-adaptation" as a possible explanation for the observed co-evolutionary changes. We use this nomenclature because it is widely accepted in the field of molecular evolution, even if it differs from that used in fields such as Ecology where "co-evolution" involves adaptive compensatory changes [9].

0014-5793/$34.00
doi:10.1016/j.febslet.2008.02.017

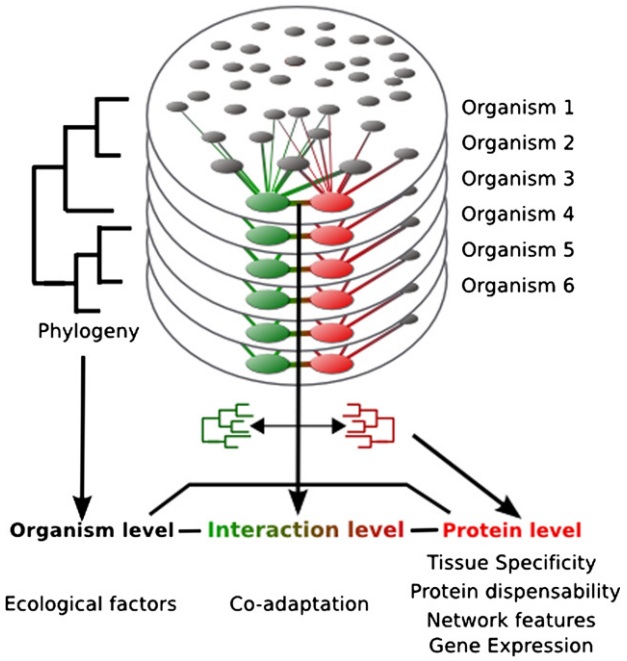

Fig. 1. Factors affecting the similarity of the evolutionary histories of two interacting proteins. Many factors acting at different levels may be responsible for the observed similarity between the phylogenetic trees of interacting proteins. While several factors can affect the evolutionary rate of both proteins to a similar degree, at the sub-protein level co-adaptation may also be at play. Even at the organism level, the underlying speciation process affects the observed tree similarity by adding a background resemblance to any pair of trees.

2.1. Co-evolutionary features at the residue level

In multiple sequence alignments, some pairs of positions show concerted mutational behaviour, such that changes in one of the positions seem not to be independent but rather related to those at the other [11]. Correlations between intra-protein pairs have been shown to be weakly yet consistently related to the closeness between the corresponding residues in the three-dimensional structure of the protein [11], as well as to other functional and structural characteristics of the involved positions [12]. Inter-protein correlated pairs, those in which the two positions belong to different protein families, have also been shown to be closer than the average inter-protein pairs [13]. Even if these pairs are not in direct contact in most cases, that is they are not actually at the protein–protein interface [14], the fact that they are closer than average means it may often be possible to use them as constraints to select the right complex between two protein chains [13,15]. The accumulation of correlated mutations between two proteins has also been used to predict whether these two proteins interact or not. Such predictions are based on the idea that pairs of interacting proteins should present an enrichment of these correlated mutations [16].

2.2. Co-evolutionary features at the whole-chain level

Pairs of interacting or functionally related proteins have been shown to co-evolve at different levels. Indeed, many methods for detecting protein interactions from genomic features, sometimes called "context-based methods" [8,17–19], actually have a co-evolutionary base. The two approaches where this co-evolutionary base is more evident are the "phylogenetic profiling" and the "mirrortree" methods.

A "phylogenetic profile" is a pattern of presence/absence of a given protein family in a set of genomes. Proteins with similar phylogenetic profiles, that is, with the same species distribution, have been shown to be functionally or physically interacting in many cases [20,21]. A possible explanation for this observation is that related proteins (those that need each other to perform a given function) must necessarily be present in the same genomes and, since one cannot work without the other, they never appear alone. In fact, we could say that the "existences" of such proteins are not independent but related to each other, and hence they co-evolve.

The "*mirrortree*" approach is based on the observation that interacting or functionally related proteins tend to have phylogenetic trees with similar shapes. This observation was first made qualitatively for some specific cases [22,23] and later, it was quantified and statistically evaluated in large datasets [24,25]. Since then, this methodology has been followed by many researches, who have improved it in different ways (i.e. [26–33]). The relationship between protein interaction and similarity of phylogenetic trees has been used not only to predict whether two families interact or not, but also to predict the mapping between the members of two families that are known to interact, that is, the individual connections between the leaves of their trees [34–38]. Indeed, the *mirrortree* quantification of the similarity between trees has not only been used to infer protein interactions in large datasets, but also to get a deeper insight into the co-evolution and function of particular interacting families [39–42].

The *mirrortree* methodology might be the one which more intuitively illustrates the relationship between protein co-evolution and interactions, since a phylogenetic tree encompasses global information on the evolution of a given protein family. Still, it is clear that "*mirrortree*" and "*phylogenetic profiling*" are conceptually related, since interacting or functionally related proteins that co-evolve can have similar trees and, ultimately, they might concurrently lose their corresponding genes. In this sense, "phylogenetic profiling" detects cases of extreme co-evolution, in which not only do sequence features co-evolve, but also the existence of the proteins themselves.

3. Co-evolution and co-adaptation

Having dealt with the evidence supporting the co-evolution of interacting proteins, we shall now look at the possible causes of such co-evolution. We will review the evidence supporting each of the two alternative hypothesis, the one stating that it is a direct consequence of a physical co-adaptative process, and the alternative one proposing that it is an indirect consequence of the similarity of their environments ultimately reflected in evolutionary rates.

At the residue level, there is plenty of evidence indicating that physical co-adaptation causes the observed intra-protein co-evolution [43]. Compensatory mutations at interfaces have been reported, particularly in fast-evolving systems such as RNA viruses [44,45]. In these cases, a destabilising mutation at the interface of one of the interacting partners is compensated by a mutation in the other partner, which restores stabil-

ity. These compensatory mutations have been found both, in natural sequences, and as a result of an introduced artificial mutation. Compensatory mutations have also been proposed as an explanation for mutations that are pathogenic in an organism and neutral in others [46–48]. For many of these cases, it has been argued that the second (compensatory) mutation may explain the neutral effect of the (otherwise pathogenic) first mutation. Most of these disease-avoiding compensatory mutations are within the same protein, although a number of inter-protein compensatory mutations have also been found [47]. Furthermore, compensatory mutations have not only been reported between interacting proteins but also between proteins and DNA [49] and protein–RNA [50].

Since co-adaptation at the residue level has been observed and it has a conceivable biophysical interpretation, it makes sense to think that the observations of co-evolution at "sub-protein" levels (i.e., segments or regions of the proteins) could also be the result of physical compensation to some extent, following the co-adaptation model. Indeed, the *mirrortree* approach was applied to protein domains showing that the domains responsible for the interaction (within two interacting proteins) co-evolve [33]. This sub-protein co-evolutionary behaviour, quantified as in *mirrortree*, has also been found between protein segments with high sequence conservation [32], as well as between the protein surfaces of obliged complexes [51].

Compensatory changes have also been proposed as the best mechanism to explain some cases of observed co-evolution where protein families are evolving very fast while having to maintain highly specific interactions without crosstalk [52–56]. In these cases, intra-organism interactions are conserved while divergence between orthologues leads to an absence of inter-organism interactions. The observations show that a strong evolutionary pressure acts on those protein interactions, both to maintain the intra-organism interactions and to avoid the inter-organism interactions. Given that these interactions derive from a common ancestral one, each interacting pair probably evolved in a co-ordinated fashion, introducing mutations that are compensated in the interacting partner in a organism-specific manner. Therefore, compensatory changes are apparently the simplest way to explain these cases and, in some of these, it was indeed possible to track the inter-protein compensatory changes down to the residue level [56].

Even if compensatory changes might be an influential factor, it is obvious that there are many other factors that can affect the evolution of interacting proteins, thereby contributing to the similarities of the corresponding phylogenetic trees. Indeed, two families that display similar evolutionary rates across all taxa would have similar trees, since the changes that occur in both families and that are responsible for shaping the tree are of a similar magnitude. A direct relationship has been reported between similarity of evolutionary rates and interaction [57,58], which could explain the similarities in the trees of interacting proteins without the strict requirement for co-adaptation. Apart from this direct relationship, protein interactions and evolutionary rates are indirectly related through a number of factors. The mRNA expression of interacting proteins or proteins involved in the same cellular process are often similar, possibly due to a similar transcriptional control [59]. Moreover, the expression levels of interacting proteins has been shown to co-evolve, in the sense that changes in the expression levels of one of the proteins from one organism to another are related to changes in the expression of the partner, as inferred from the "codon adaptation index" of the corresponding proteins [60,61]. Closing this circle, there is the relationship between the level of expression and the evolutionary rate, whereby highly expressed proteins tend to evolve more slowly [62–64]. This is typically explained by the fact that the mutational possibilities of important proteins (usually "hubs" in interaction networks) are constrained, as is their expression, because changes in any of these two elements would effect many other proteins. The evolutionary rate has also been related to protein dispensability, in the sense that essential genes evolve more slowly [62,65–67]. Essentiality has in turn been related to protein interactions: essential proteins tend to be highly connected ("hubs") in protein interaction networks [6]. A direct relationship between protein connectivity and evolutionary rate has also been found: hubs evolve more slowly [57,68,69]. This may be explained by mutations in hubs being highly constrained since they affect many interacting proteins. For this reason they would tend to be more conserved. In summary, the evolutionary rate is related to protein interactions through many different direct and indirect pathways.

Another factor that could shed some light on the causes of the observed co-evolution of interacting proteins is the specificity of that co-evolution. Co-evolution particular to a pair of proteins could be interpreted as a sign of co-adaptation, while it makes more sense to explain broad co-evolutionary trends involving many proteins by a general similarity of evolutionary rates. In this sense, a number of methods are able to detect specific co-evolution excluding global co-evolutionary trends, such as the one due to the underlying speciation process [26,27]. Other methods, using a partial correlation formulation, are directly able to quantify to which extent a co-evolutionary signal is particular to a given pair of proteins [70]. Even if these initial studies point towards co-adaptation as the cause of co-evolution, they do not provide direct evidence and more work is needed in this respect.

In a recent study, Hakes et al. [58] have tried to rule out the co-adaptation hypothesis by showing that residues in protein interfaces, that in a simple model are expected to be implicated in physical co-adaptation, do not show strong signals of co-evolution, i.e., similar trees. These results are controversial since previous detailed studies showed that there was co-evolution between interfaces in obliged complexes [51]. It is also worth to point out that even if the actual residues in the interfaces do not co-mutate, compensatory effects of the mutations can still occur over relatively large distances. Indeed, it was previously shown that inter-protein correlated mutations tend to be closer than average [13], but not necessarily in direct contact [14]. In our opinion, even if co-evolution is not at play at strict interfaces, it is still not be possible to rule out physical co-adaptation at longer distances, possibly via "allosteric" effects.

4. Conclusions

The co-evolution between interacting protein could be due to the accumulation of compensatory changes at the residue level or to similar evolutionary rates that globally affect the two

protein families. There are results that support both hypotheses, and it is conceivable that both forces are playing a role in different degrees, at different levels, and for different cases.

It will be useful to imagine what might be the ideal experiment to discriminate between these two hypotheses. One can imagine that one such experiment could involve the detailed comparison of the energetic contribution associated to each mutational path in a family of proteins, possibly established by reconstructing the ancestral sequences. This experiment would have to be carried out for families of proteins with known structure to make it possible to perform sufficiently detailed energetic calculations. This is an obviously difficult scenario since it requires detailed reconstruction of mutation pathways, modelling of structures and biophysical characterization. However, it is possible that this is the closest we could come to completely solving this problem.

On a practical sense, even if the predictive power of *mirrortree* and related methods is totally independent of any of these hypothesis, since it only depends on the observed relationship between co-evolution (i.e., tree similarity) and protein interactions, discerning the origin of protein co-evolution has important theoretical and practical consequences. For example, knowing what is behind the process would speed up the development and improvement of the co-evolution based methods for predicting interactions, since they would be designed in a more rational way taking this information into account. It would also help modify, ''tune'' and redesign the specificity of protein interactions. Finally, such information could have profound implications in Systems Biology since it would help to discern the evolutionary scenario that led to the complex structure of current interactomes. In this sense, co-evolutionary forces might be an important factor to consider in the models of interactome evolution [71]. None of these models is able to explain all the observed topological features of the interactomes. Maybe co-evolutionary forces have to be considered in their contribution to the topological characteristics of the protein interaction networks, since it lead to differential connectivity degree on various regions of the interactome.

Finally, it is important to consider that the two hypotheses to explain the similarities between trees are not mutually exclusive and both together could shape the trees of interacting proteins at different scales and to different degrees. Compensatory changes at the residue level have been found in many pairs of interacting proteins, but it is difficult to consider that they alone could be responsible for the global similarity in trees. Indeed, if this were the case, a large number of changes would be necessary to produce observable changes in the tree topology. In the limit, tree similarity due to a large number of compensatory changes spread throughout the length of the protein would be indistinguishable from a tree similarity due to similar evolutionary rates. A feasible hypothesis that is compatible with all the available data is that the observed tree similarity between related proteins is mainly due to similar evolutionary rates (which are in turn related to similar expression patterns, etc.), and that compensatory changes are acting locally, shaping the details of the interacting regions.

Acknowledgements: This work was funded in part by the Grants BIO2006-15318 and PIE 200620I240 from the Spanish Ministry for Education and Science, and the Grants LSHG-CT-2003-503265 (BioSapiens) and LSHG-CT-2004-503567 (ENFIN) from the EC.

References

[1] Van Regenmortel, M.H. (2004) Reductionism and complexity in molecular biology. Scientists now have the tools to unravel biological and overcome the limitations of reductionism. EMBO Rep. 5, 1016–1020.
[2] Nurse, P. (2003) Systems biology: understanding cells. Nature 424, 883.
[3] Hood, L. (2003) Systems biology: integrating technology, biology, and computation. Mech. Ageing Dev. 124, 9–16.
[4] Kitano, H. (2002) Systems biology: a brief overview. Science 295, 1662–1664.
[5] Uetz, P. and Finley Jr., R.L. (2005) From protein networks to biological systems. FEBS Lett. 579, 1821–1827.
[6] Jeong, H., Mason, S.P., Barabási, A.L. and Oltvai, Z.N. (2001) Lethality and centrality in protein networks. Nature 411, 41–42.
[7] Alm, E. and Arkin, A.P. (2003) Biological networks. Curr. Opin. Struct. Biol. 13, 193–202.
[8] Xia, Y., Yu, H., Jansen, R., Seringhaus, M., Baxter, S., Greenbaum, D., Zhao, H. and Gerstein, M. (2004) Analyzing cellular biochemistry in terms of molecular networks. Annu. Rev. Biochem. 73, 1051–1087.
[9] Thompson, J.N. (1994) The Coevolutionary Process, University of Chicago Press, Chicago.
[10] Miyamoto, M.M. and Fitch, W.M. (1995) Testing the covarion hypothesis of molecular evolution. Mol. Biol. Evol. 12, 503–513.
[11] Göbel, U., Sander, C., Schneider, R. and Valencia, A. (1994) Correlated mutations and residue contacts in proteins. Proteins 18, 309–317.
[12] Socolich, M., Lockless, S.W., Russ, W.P., Lee, H., Gardner, K.H. and Ranganathan, R. (2005) Evolutionary information for specifying a protein fold. Nature 437, 512–518.
[13] Pazos, F., Helmer-Citterich, M., Ausiello, G. and Valencia, A. (1997) Correlated mutations contain information about protein–protein interaction. J. Mol. Biol. 271, 511–523.
[14] Halperin, I., Wolfson, H. and Nussinov, R. (2006) Correlated mutations: advances and limitations. A study on fusion proteins and on the cohesin–dockerin families. Proteins 63, 832–845.
[15] Tress, M., Juan, D., Grana, O., Gomez, M.J., Gomez-Puertas, P., Gonzalez, J.M., Lopez, G. and Valencia, A. (2005) Scoring docking models with evolutionary information. Proteins 60, 275–280.
[16] Pazos, F. and Valencia, A. (2002) In silico two-hybrid system for the selection of physically interacting protein pairs. Proteins 47, 219–227.
[17] Valencia, A. and Pazos, F. (2002) Computational methods for the prediction of protein interactions. Curr. Opin. Struct. Biol. 12, 368–373.
[18] Shoemaker, B.A. and Panchenko, A.R. (2007) Deciphering protein–protein interactions. Part II. Computational methods to predict protein and domain interaction partners. PLoS Comput. Biol. 3, e43.
[19] Salwinski, L. and Eisenberg, D. (2003) Computational methods of analysis of protein–protein interactions. Curr. Opin. Struct. Biol. 13, 377–382.
[20] Pellegrini, M., Marcotte, E.M., Thompson, M.J., Eisenberg, D. and Yeates, T.O. (1999) Assigning protein functions by comparative genome analysis: protein pylogenetic profiles. Proc. Natl. Acad. Sci. USA 96, 4285–4288.
[21] Marcotte, E.M., Pellegrini, M., Thompson, M.J., Yeates, T.O. and Eisenberg, D. (1999) A combined algorithm for genome-wide prediction of protein function. Nature 402, 83–86.
[22] Fryxell, K.J. (1996) The coevolution of gene family trees. Trend Genet. 12, 364–369.
[23] Pages, S., Belaich, A., Belaich, J.P., Morag, E., Lamed, R., Shoham, Y. and Bayer, E.A. (1997) Species-specificity of the

cohesin–dockerin interaction between *Clostridium thermocellum* and *Clostridium cellulolyticum*: prediction of specificity determinants of the dockerin domain. Proteins 29, 517–527.
[24] Pazos, F. and Valencia, A. (2001) Similarity of phylogenetic trees as indicator of protein–protein interaction. Protein Eng. 14, 609–614.
[25] Goh, C.-S., Bogan, A.A., Joachimiak, M., Walther, D. and Cohen, F.E. (2000) Co-evolution of proteins with their interaction partners. J. Mol. Biol. 299, 283–293.
[26] Pazos, F., Ranea, J.A.G., Juan, D. and Sternberg, M.J.E. (2005) Assessing protein co-evolution in the context of the tree of life assists in the prediction of the interactome. J. Mol. Biol. 352, 1002–1015.
[27] Sato, T., Yamanishi, Y., Kanehisa, M. and Toh, H. (2005) The inference of protein–protein interactions by co-evolutionary analysis is improved by excluding the information about the phylogenetic relationships. Bioinformatics 21, 3482–3489.
[28] Sato, T., Yamanishi, Y., Horimoto, K., Kanehisa, M. and Toh, H. (2006) Partial correlation coefficient between distance matrices as a new indicator of protein–protein interactions. Bioinformatics 22, 2488–2492.
[29] Kim, W.K., Bolser, D.M. and Park, J.H. (2004) Large-scale co-evolution analysis of protein structural interlogues using the global protein structural interactome map (PSIMAP). Bioinformatics 20, 1138–1150.
[30] Tan, S., Zhang, Z. and Ng, S. (2004) ADVICE: automated detection and validation of interaction by co-evolution. Nucleic Acid Res. 32, W69–W72.
[31] Waddell, P.J., Kishino, H. and Ota, R. (2007) Phylogenetic methodology for detecting protein interactions. Mol. Biol. Evol. 24, 650–659.
[32] Kann, M.G., Jothi, R., Cherukuri, P.F. and Przytycka, T.M. (2007) Predicting protein domain interactions from coevolution of conserved regions. Proteins 67, 811–820.
[33] Jothi, R., Cherukuri, P.F., Tasneem, A. and Przytycka, T.M. (2006) Co-evolutionary analysis of domains in interacting proteins reveals insights into domain–domain interactions mediating protein–protein interactions. J. Mol. Biol. 362, 861–875.
[34] Gertz, J., Elfond, G., Shustrova, A., Weisinger, M., Pellegrini, M., Cokus, S. and Rothschild, B. (2003) Inferring protein interactions from phylogenetic distance matrices. Bioinformatics 19, 2039–2045.
[35] Ramani, A.K. and Marcotte, E.M. (2003) Exploiting the co-evolution of interacting proteins to discover interaction specificity. J. Mol. Biol. 327, 273–284.
[36] Jothi, R., Kann, M.G. and Przytycka, T.M. (2005) Predicting protein–protein interaction by searching evolutionary tree automorphism space. Bioinformatics 21, i241–i250.
[37] Izarzugaza, J.M., Juan, D., Pons, C., Ranea, J.A., Valencia, A. and Pazos, F. (2006) TSEMA: interactive prediction of protein pairings between interacting families. Nucleic Acid Res. 34, W315–W319.
[38] Tillier, E.R., Biro, L., Li, G. and Tillo, D. (2006) Codep: maximizing co-evolutionary interdependencies to discover interacting proteins. Proteins 63, 822–831.
[39] Labedan, B., Xu, Y., Naumoff, D.G. and Glansdorff, N. (2004) Using quaternary structures to assess the evolutionary history of proteins: the case of the aspartate carbamoyltransferase. Mol. Biol. Evol. 21, 364–373.
[40] Devoto, A., Hartmann, H.A., Piffanelli, P., Elliott, C., Simmons, C., Taramino, G., Goh, C.S., Cohen, F.E., Emerson, B.C., Schulze-Lefert, P. and Panstruga, R. (2003) Molecular phylogeny and evolution of the plant-specific seven-transmembrane MLO family. J. Mol. Evol. 56, 77–88.
[41] Dou, T., Ji, C., Gu, S., Xu, J., Ying, K., Xie, Y. and Mao, Y. (2006) Co-evolutionary analysis of insulin/insulin like growth factor 1 signal pathway in vertebrate species. Front Biosci. 11, 380–388.
[42] McPartland, J.M., Norris, R.W. and Kilpatrick, C.W. (2007) Coevolution between cannabinoid receptors and endocannabinoid ligands. Gene 397, 126–135.
[43] Shim Choi, S., Li, W. and Lahn, B.T. (2005) Robust signals of coevolution of interacting residues in mammalian proteomes identified by phylogeny-aided structural analysis. Nat. Genet. 37, 1367–1371.
[44] Mateu, M.G. and Fersht, A.R. (1999) Mutually compensatory mutations during evolution of the tetramerization domain of tumor suppressor p53 lead to impaired hetero-oligomerization. Proc. Natl. Acad. Sci. USA 96, 3595–3599.
[45] del Alamo, M. and Mateu, M.G. (2005) Electrostatic repulsion, compensatory mutations, and long-range non-additive effects at the dimerization interface of the HIV capsid protein. J. Mol. Biol. 345, 893–906.
[46] Kondrashov, A.S., Sunyaev, S. and Kondrashov, F.A. (2002) Dobzhansky–Muller incompatibilities in protein evolution. Proc. Natl. Acad. Sci. USA 99, 14878–14883.
[47] Ferrer-Costa, C., Orozco, M. and Cruz, X. (2007) Characterization of compensated mutations in terms of structural and physicochemical properties. J. Mol. Biol. 365, 249–256.
[48] Kulathinal, R.J., Bettencourt, B.R. and Hartl, D.L. (2004) Compensated deleterious mutations in insect genomes. Science 306, 1553–1554.
[49] Athanikar, J.N. and Osborne, T.F. (1998) Specificity in cholesterol regulation of gene expression by coevolution of sterol regulatory DNA element and its binding protein. Proc. Natl. Acad. Sci. USA 95, 4935–4940.
[50] Metzenberg, S., Joblet, C., Verspieren, P. and Agabian, N. (1993) Ribosomal protein L25 from *Trypanosoma brucei*: phylogeny and molecular co-evolution of an rRNA-binding protein and its rRNA binding site. Nucleic Acid Res. 21, 4936–4940.
[51] Mintseris, J. and Weng, Z. (2005) Structure, function, and evolution of transient and obligate protein–protein interactions. Proc. Natl. Acad. Sci. USA 102, 10930–10935.
[52] Wang, S. and Kimble, J. (2001) The TRA-1 transcription factor binds TRA-2 to regulate sexual fates in *Caenorhabditis elegans*. EMBO J. 20, 1363–1372.
[53] Watanabe, M., Ito, A., Takada, Y., Ninomiya, C., Kakizaki, T., Takahata, Y., Hatakeyama, K., Hinata, K., Suzuki, G., Takasaki, T., Satta, Y., Shiba, H., Takayama, S. and Isogai, A. (2000) Highly divergent sequences of the pollen self-incompatibility (S) gene in class-I S haplotypes of *Brassica campestris* (syn. rapa) L. FEBS Lett. 473, 139–144.
[54] Kachroo, A., Schopfer, C.R., Nasrallah, M.E. and Nasrallah, J.B. (2001) Allele-specific receptor–ligand interactions in *Brassica* self-incompatibility. Science 293, 1824–1826.
[55] Haag, E.S., Wang, S. and Kimble, J. (2002) Rapid coevolution of the nematode sex-determining genes fem-3 and tra-2. Curr. Biol. 12, 2035–2041.
[56] Liu, J.-C., Makova, K.D., Adkins, R.M., Gibson, S. and Li, W.-H. (2001) Episodic evolution of growth hormone in primates and emergence of the species specificity of human growth hormone receptor. Mol. Biol. Evol. 18, 945–953.
[57] Fraser, H.B., Hirsh, A.E., Steinmetz, L.M., Scharfe, C. and Feldman, M.W. (2002) Evolutionary rate in the protein interaction network. Science 296, 750–752.
[58] Hakes, L., Lovell, S., Oliver, S.G. and Robertson, D.L. (2007) Specificity in protein interactions and its relationship with sequence diversity and coevolution. Proc. Natl. Acad. Sci. USA 104, 7999–8004.
[59] Eisen, M.B., Spellman, P.T., Brown, P.O. and Botstein, D. (1998) Cluster analysis and display of genome-wide expression patterns. Proc. Natl. Acad. Sci. USA 95, 14863–14868.
[60] Fraser, H.B., Hirsh, A.E., Wall, D.P. and Eisen, M.B. (2004) Coevolution of gene expression among interacting proteins. Proc. Natl. Acad. Sci. USA 101, 9033–9038.
[61] Chen, Y. and Dokholyan, N.V. (2006) The coordinated evolution of yeast proteins is constrained by functional modularity. Trend Genet. 22, 416–419.
[62] Pal, C., Papp, B. and Hurst, L.D. (2001) Highly expressed genes in yeast evolve slowly. Genetics 158, 927–931.
[63] Subramanian, S. and Kumar, S. (2004) Gene expression intensity shapes evolutionary rates of the proteins encoded by the vertebrate genome. Genetics 168, 373–381.
[64] Drummond, D.A., Raval, A. and Wilke, C.O. (2006) A single determinant dominates the rate of yeast protein evolution. Mol. Biol. Evol. 23, 327–337.
[65] Hirsh, A.E. and Fraser, H.B. (2001) Protein dispensability and rate of evolution. Nature 411, 1046–1049.

[66] Jordan, I.K., Rogozin, I.B., Wolf, Y.I. and Koonin, E.V. (2002) Essential genes are more evolutionarily conserved than are nonessential genes in bacteria. Genome Res. 12, 962–968.
[67] Zhang, L. and Li, W.-H. (2004) Mammalian housekeeping genes evolve more slowly than tissue-specific genes. Mol. Biol. Evol. 21, 236–239.
[68] Fraser, H.B. (2005) Modularity and evolutionary constraint on proteins. Nat. Genet. 6, 6.
[69] Kim, P., Lu, L., Xia, Y. and Gerstein, M. (2006) Relating three-dimensional structures to protein networks provides evolutionary insights. Science 314, 1938–1941.
[70] Juan, D., Pazos, F. and Valencia, A. (2008) High-confidence prediction of global interactomes based on genome-wide coevolutionary networks. Proc. Natl. Acad. Sci. USA 105, 934–939.
[71] Barabasi, A.L. and Oltvai, Z.N. (2004) Network biology: understanding the cell's functional organization. Nat. Rev. Genet. 5, 101–113.