JOURNAL OF EXPERIMENTAL ZOOLOGY (MOL DEV EVOL) 308B:97–103 (2007)

Erosion of Interaction Networks in Reduced and Degraded Genomes

HOWARD OCHMAN[1]*, RENYI LIU[1], AND EDUARDO P.C. ROCHA[2]
[1]*Department of Biochemistry and Molecular Biophysics, University of Arizona, Tucson, Arizona 85721*
[2]*Unite Genetique des Genomes Bacteriens, Institut Pasteur, Paris, France*

ABSTRACT Unlike eukaryotes, which often recruit duplicated genes into existing protein–protein interaction (PPI) networks, the low levels of gene duplication coupled with the high probability of lateral transfer of novel genes alters the manner in which PPI networks can evolve in bacteria. By inferring the PPIs present in the ancestor to contemporary *Gammaproteobacteria*, we were able to trace the changes in gene repertoires, and their consequences on PPI network evolution, in several bacterial lineages that have independently undergone reductions in genome size and genome contents. As genomes degrade, virtually all multi-partner proteins have lost interactors; however, the overall average number of connections increases due to the preferential elimination of proteins that interact with only one other protein partner. We also studied the effect of lateral gene transfer on PPI network evolution by analyzing the connectivity of genes that have been gained along the *Escherichia coli* lineage, as well as those acquired genes subsequently silenced in *Shigella flexneri*, since diverging from the gammaproteobacterial ancestor. The situation in PPI networks, in which newly acquired genes preferentially attach to the hubs of the network, contrasts that observed in metabolic networks, which evolve by the peripheral gain and loss of genes, and in regulatory networks, in which high connectivity increases the propensity of loss. *J. Exp. Zool. (Mol. Dev. Evol.) 308B:97–103, 2007.*

How to cite this article: Ochman H, Liu R, Rocha EPC. 2007. Erosion of interaction networks in reduced and degraded genomes. J. Exp. Zool. (Mol. Dev. Evol.) 308B: 97–103.

Genome-scale analyses of the functional links among proteins have revealed that interaction networks are essentially scale-free, with relatively few proteins (termed ''hubs'') that are highly connected and interacting with numerous other proteins but with the majority having few interacting partners (Albert et al., 2000; Jeong et al., 2001; Wuchty, 2004). This topology causes biological networks to be tolerant of perturbation, such that the loss or inactivation of most genes will have little effect on organismal fitness. The protein–protein interaction (PPI) networks resolved for yeast, worm and *Drosophila* are organized similarly; however, there are cumulatively many differences among these organisms in the numbers and connections of homologous proteins (Matthews et al., 2001; Hahn and Kern, 2005; Sharan et al., 2005; Gandhi et al., 2006; Li et al., 2006). These differences and the robustness of interaction networks have been ascribed to the combination of two factors: the inherent nature of a scale-free network topology, which is robust relative to the elimination of most individual proteins (Albert et al., 2000; Jeong et al., 2001; Maslov and Sneppen, 2002), and the recruitment of paralogs, which are common in most eukaryotic genomes (Wagner, 2000; Gu et al., 2003; Pereira-Leal et al., 2006).

Unfortunately, this model of PPI network restructuring cannot account for the major trends observed in the evolution of bacterial genomes. Most bacterial genomes contain few paralogs (Hooper and Berg, 2003; Lerat et al., 2005), which might limit the extent to which newly generated proteins can join modules or displace existing

*Correspondence to: H. Ochman, Department of Biochemistry and Molecular Biophysics, University of Arizona, Tucson, AZ 85721, USA. E-mail: hochman@email.arizona.edu
Received 2 August 2006; Revised 10 October 2006; Accepted 6 November 2006
Published online 11 January 2007 in Wiley InterScience (www.interscience.wiley.com). DOI: 10.1002/jez.b.21147.

proteins. Additionally, lateral gene transfer can enrich the proteome with proteins having no functional, structural or sequence similarity to those already encoded by the genome (Daubin and Ochman, 2004; Pál et al., 2005). Moreover, many bacterial lineages have undergone a massive reduction in genome size caused by deletions that can sometimes span hundreds of genes and result in large variation in the genome contents of related strains (Moran, 2002; Nilsson et al., 2005). These features imply that the numbers and types of interactions, although encoded by homologous sets of genes, will be inconsistent across bacterial genomes and can change in ways that differ from those observed in eukaryotic genomes.

In this paper, we ask how the large-scale degradation of bacterial genomes has affected the architecture of PPI networks. To monitor and understand the process by which these networks adjust to severe changes in the numbers of potential protein interactors, we focus on the *Gammaproteobacteria* because this class boasts some of the most extreme reductions in genome size and gene contents, with sequenced members ranging from 160 to over 7,000 kb. Phylogenetic analyses and comparisons of gene repertoires both indicate that genome size reduction occurred independently in several proteobacterial lineages (Ochman, 2005). In addition, complete sequences are available for genomes at more intermediate stages of degradation, in which genomes are less reduced in size and can also harbor high numbers of pseudogenes. Our results show that PPI networks in bacteria evolve in a way that is distinct from both the PPI networks in eukaryotes and the metabolic and regulatory networks in bacteria.

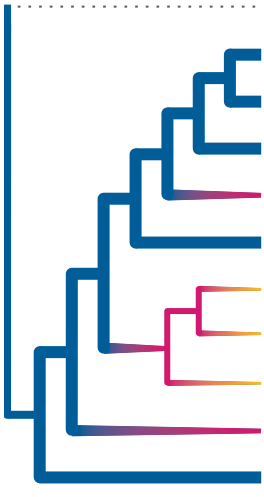
MATERIALS AND METHODS

Sequence data

We used the complete genome sequences available in GenBank (ftp://ftp.ncbi.nih.gov/genomes/Bacteria/) of the following bacteria (for accession numbers, see Fig. 1): *Buchnera aphidicola* Sg, *Candidatus Blochmannia floridanus*, *Escherichia coli* MG1655, *Haemophilus influenzae* Rd, *Shigella flexneri* 2a_311, *Sodalis glossinidius* morsitans, *Vibrio cholerae* N16961, *Salmonella enterica* Typhimurium LT2, *Wigglesworthia glossinidia* brevipalpis, *Yersinia pestis* CO92. Pseudogenes in *Sodalis* were identified by Toh et al. (2006), and those in *Shigella* were identified by Jin et al. (2002) and by Lerat and Ochman (2004).

Identification of orthologs

All-vs.-all BLAST searches were performed with annotated proteins in *E. coli* and each of the other genomes listed above. A pair of genes was regarded as a pair of putative orthologs if they were reciprocal best hits with more than 40% similarity



Species	Genbank #	Genome Size (kb)	Genes	Genes in PPI	Interactions	Genes with *n* Interactors 1	2-5	6-10	11-15	>15
Ancestor			2,425	942	5,047	345	264	107	43	183
Escherichia coli	NC_000913	4,640	4,288	1,339	5,868	612	338	135	59	195
Shigella flexneri	NC_004337	4,607	4,434	861	4,347	312	250	99	45	155
Salmonella enterica	NC_003197	4,857	4,450	875	4,727	309	253	95	48	170
Sodalis glossinidius	NC_007712	4,171	2,432	614	3,763	169	191	74	35	145
Yersinia pestis	AL590842	4,654	4,012	814	4,445	277	243	88	46	160
Blochmannia floridanus	NC_005061	706	625	290	1,751	61	82	40	29	78
Wigglesworthia glossinidia	NC_004344	698	621	303	1,789	65	88	42	30	78
Buchnera aphidicola	NC_004061	641	545	297	1,850	56	98	38	26	79
Haemophilus influenzae	NC_000907	1,830	1,742	540	3,406	145	151	70	43	131
Vibrio cholerae	NC_002505 & 002506	4,033	3,885	705	4,070	219	207	85	41	153

Fig. 1. Characteristics of gammaproteobacterial genomes and of the reconstructed ancestor of these species. Tree topology adapted from those presented in Gil et al. (2003) and Herbeck et al. (2005).

in sequence in a global alignment and if the difference in their lengths was less than 20% of the overall size of the gene. The gene contents of the gammaproteobacterial ancestor follows that reconstructed and described by Moran and Mira (2001). We determined which ancestral genes were retained in each of the genomes considered based on their orthology to the corresponding gene in *E. coli* MG1655.

Network data

Interaction networks were based solely on the PPIs determined for *E. coli* by Butland et al. (2005). The address of each gene/protein assayed was identified in the latest release of the *E. coli* chromosome available in Genbank, and subsequently mapped onto the reconstructed ancestral genome. All redundant links from the list of pairwise PPIs were removed, such that the interactions A to B and B to A were only considered once. Naturally, it is assumed that if the orthologs of *E. coli* genes are present in the ancestral or contemporary genomes that the interactions among orthologs are the same as those determined experimentally for *E. coli*. Given that orthologs largely maintain the same function in very divergent organisms, the interactions among conserved proteins are very likely to be conserved. Data on regulatory networks were derived from RegulonDB (Salgado et al., 2006; http://regulondb.ccg.unam.mx/index.html), which contains the transcription factors/gene interaction network assembled for *E. coli*. Because regulatory networks are directed graphs with regulators as source nodes, we were able to distinguish regulators from regulated elements. Data analyses and graphing were performed with in-house scripts and *Pajek* (de Nooy et al., 2005; http://vlado.fmf.uni-lj.si/pub/networks/pajek/). All lists of the orthologs detected in each genome as well as information on interaction networks are available upon request from the authors.

RESULTS AND DISCUSSION

A gene's propensity for loss is dependent on its connectivity

The connections among proteins encoded by two slightly reduced (~4,000 genes; *Shigella flexneri*, *Y. pestis*), two moderately reduced (~2,000 genes; *Sodalis glossinidia*, *H. influenzae*) and three highly reduced endosymbiont (<1,000 genes; *Buchnera aphidicola*, *Blochmannia floridanus*, *W. glossinidia*) genomes (Fig. 1) were derived from the protein interactions resolved experimentally for *E. coli* by tandem affinity purification (Butland et al., 2005). Because contemporary *E. coli* is not the ancestor to these species, we consider only those proteins inferred to be present in the common ancestor (Moran and Mira, 2001), thereby excluding those *E. coli* genes acquired after these lineages diverged. The genome of the reconstructed ancestor genome contains 57% of the 4,288 genes; but because the proteins analyzed were not selected at random (Butland et al., 2005), 70% of the 1,339 interacting proteins and 86% of the 5,868 interactions (including self interactions but removing redundant links) detected in *E. coli* are present in the ancestor.

The endosymbiont genomes are approximately 85% smaller than that of *E. coli*, and their current gene repertoires have been static for millions of years due to their long-term relationships with particular insect hosts (Tamas et al., 2002). Each of the endosymbionts has retained about 30%, albeit different sets, of the tested genes present in the ancestor, and there are significantly higher numbers of connections per protein in endosymbionts (11.4 vs. 10.1) despite the reduced number of potential interactors. Although the most highly connected proteins have lost substantial numbers of links in these reduced genomes, this difference is due primarily to the removal of genes that encode proteins with only one interacting partner (Fig. 1, Fig. 2d,e). For example, in *Buchnera*, single-partner proteins are five times less likely to be retained than are proteins with more than 15 partners.

Only 17 (5%) of the single-partner proteins remain in all three endosymbiont genomes. Although this seems to be a low number, it is more than the 11 expected by chance if genome reduction occurred independently in each lineage. The retention of this particular set of single-partner proteins is most likely due to selection for their individual functions: in all but two cases (*acpS, dut*), the genes are maintained in all other gammaproteobacterial genomes considered, and 10 have been shown to be essential in *E. coli* (Gerdes et al., 2003). In contrast to the situation with single-partner proteins, higher proportions of multiple-partner proteins—even those with only two partners—are retained, supporting the view, originally developed for eukaryotes (Jeong et al., 2001; Krylov et al., 2003; Wuchty, 2004; Campillos et al., 2006), that the contribution of a protein to fitness and its evolutionary conservation increase

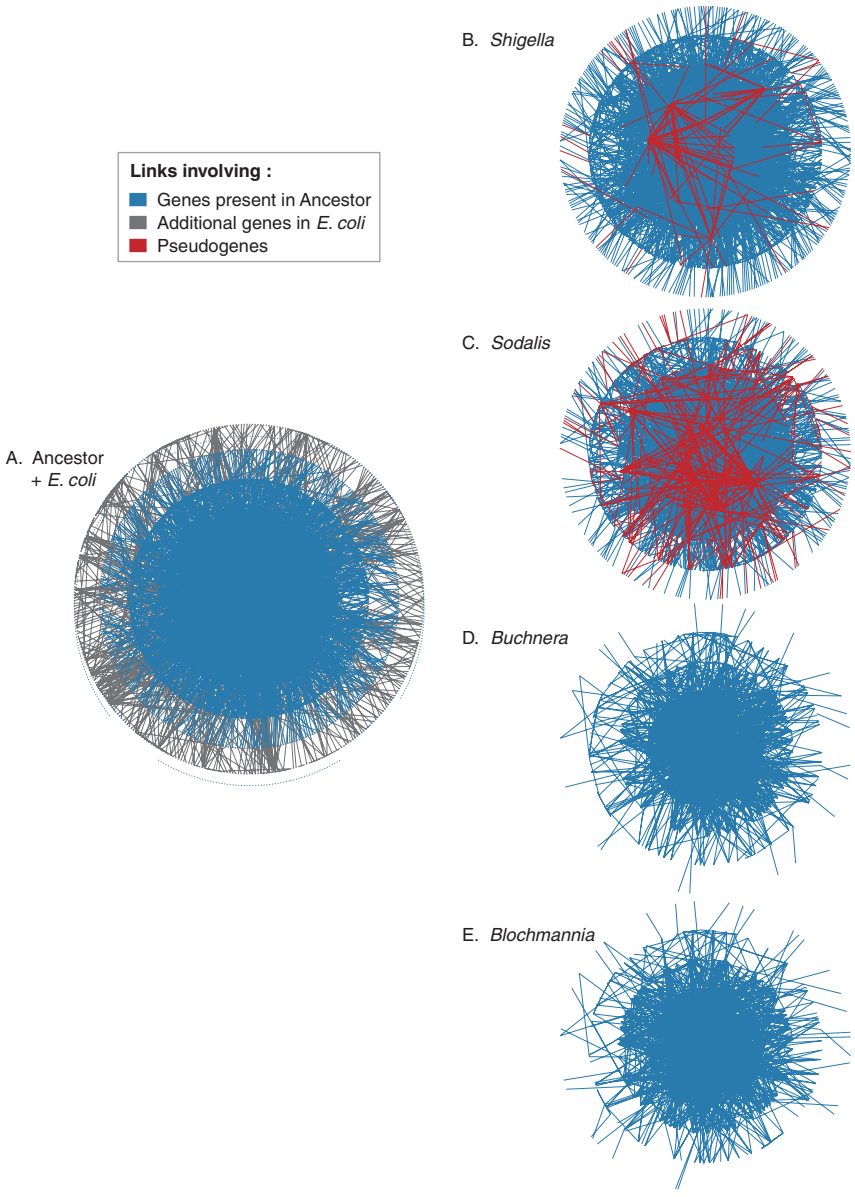


Fig. 2. Protein–protein interaction (PPI) networks in reduced and degraded gammaproteobacterial genomes. (**A**) Blue edges denote interactions among proteins present in the reconstructed ancestor. Proteins are arranged into five rings according to their numbers of interactions, with the innermost ring containing those proteins with the most connections and the outermost, the least. Gray edges denote interactions encoded by genes acquired by *E. coli* since it diverged from the ancestor: these constitute the ring outside of the network of the ancestor. (**B**, **C**) PPI networks in the reduced genomes of *Shigella flexneri* and *Sodalis glossinidia*, respectively. Red edges denote interactions involving a pseudogene. (**D**, **E**) PPI networks of the highly degraded genomes of the endosymbionts *Buchnera aphidicola* and *Blochmannia floridanus*, respectively.

with its number of interacting partners. Still, over half of the proteins with >20 partners have been eliminated from at least one symbiont genome (74/142), and 24% are absent from all endosymbionts.

When considering the highly reduced genomes, virtually all modules of all sizes within the PPI interaction network have lost constituents. As might be expected, proteins that interact with larger numbers of proteins in the ancestor still tend to be those with more partners in reduced genomes. However, those proteins with >10 interactions lose, on average, a significantly greater percentage of partners than do proteins with 5–10 partners (39% vs. 28%), implying that highly connected proteins are better buffered against the loss of individual protein partners.

How gene inactivation trims PPI networks

Whereas the current genomes of obligate endosymbionts are well-honed and stable, glimpses into the process of genome reduction are captured in the genomes of the more recent associates of eukaryotes. Among the species considered, an extreme case of genome degradation is provided by the facultative symbiont *Sodalis glossinidia*, in which a quarter of its 4.2 Mb genome consists of pseudogenes (Toh et al., 2006). The current population of *Sodalis* pseudogenes corresponds to proteins that have, on average, only 3.7 partners in the ancestor, and significantly more ancestral proteins with single partners have been inactivated. When compared to those proteins already jettisoned from the *Sodalis* genome, these pseudogenes average 15% fewer links, and a lower proportion are single-partnered. This decrease in the number of pseudogene interactors results from the changes in the number of potential interactors as the genome shrinks. Although genome reduction has occurred independently in *Sodalis* and the three endosymbionts, virtually all of the genes that are removed or inactivated in *Sodalis* have also been eliminated from each of the endosymbiont genomes.

The effects of lateral gene transfer on PPI networks

In contrast to the attrition of interaction networks resulting from gene degradation and loss, examining the features of the *E. coli* proteins gained since its divergence from the reconstructed ancestor provides insights into the manner in which bacterial proteins join interaction networks. In eukaryotes, new proteins arise primarily through gene duplications and immediately obtain the same interacting partners as their paralog upon origination. But because most new genes in bacterial genomes originate by lateral transfer and produce proteins that are largely unique to the genome, there are usually no pre-established interactions that can be readily recruited from a paralog at the moment of transfer.

Of the nearly 400 *E. coli* proteins for which PI's were analyzed and that are not present in the ancestor, 70% (277/397) have only one interacting partner, nearly double the proportion of single-partner proteins in the ancestor. This situation is very different from the one in eukaryotes where gene duplication leads to groups of paralogs that have higher connectivity than average genes (Wagner, 2001). The majority of these single-partnered acquired proteins interact directly with existing hubs, i.e., ancestral proteins with over 10 interactions, indicating that they are most often integrated into established networks as opposed to offering entirely new functions. This result parallels that observed for the recruitment of acquired genes into *E. coli* metabolic networks (Light et al., 2005; Pál et al., 2005). However, metabolic networks evolve by attachment to their periphery, whereas new PPIs evolve by preferential attachment to hubs. Indeed, only about 20% (64/277) of the new single-partner proteins interact only with another acquired protein, but surprisingly, the proportion of links formed by new proteins with other acquired proteins increases with the number of interactions (28%, 32% and 50%, for the two-, three- and four-partnered proteins, respectively). At first glance, this seems counterintuitive; however, new proteins with multiple partners are often acquired as part of a cluster of functionally related genes (e.g., operons) whose products physically interact (Dandekar et al., '98).

Losing acquired genes

Having shared much of its evolutionary history with *E. coli*, *Shigella flexneri* has a very similar gene repertoire but contains hundreds of recently derived pseudogenes that are still functional in *E. coli* (Jin et al., 2002). These nonfunctional proteins in *Shigella* can provide insights into (1) the initial stages of genome degradation and (2) the incorporation and removal of acquired genes from PPI networks. Among those proteins acquired by *E. coli*, since it diverged from the reconstructed ancestor, over 10% are presently encoded by pseudogenes in *Shigella*. With regard to their numbers of interaction partners, the genes that are silenced in *Shigella* constitute a random sample of those acquired; i.e., most (32/46) encode single-partner proteins whose sole partners possess a large number of links ($\bar{x} = 26$). This result runs counter to the prediction that introduced genes acquiring large numbers of protein interactions are more likely to be retained.

Since acquired sequences rarely encode housekeeping functions, pseudogenes are expected to predominate among acquired sequences. However, 30 of the 943 ancestral genes tested for interactions have also become pseudogenes in *Shigella* (and are still functional and were tested for PPI in *E. coli*). The ancestral genes that became inactivated in *Shigella* do not represent a random subset of the proteins present in the ancestor: on

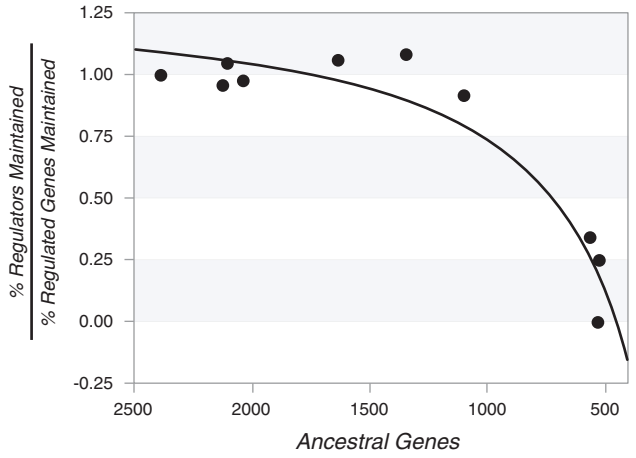


Fig. 3. Degradation of regulatory networks in gammaproteobacterial genomes. The proportion of ancestral regulators over the proportion of regulated genes maintained in a genome (y-axis) is plotted according to the total number of ancestral genes in the corresponding genome (x-axis). The steep decline associated with reductions in genome size indicates that highly connected regulators are lost more quickly than are the lowly connected regulated genes.

average, they have significantly fewer interacting partners (4.4 vs. 10.1), and each is missing in at least one of the other gammaproteobacterial genomes considered.

Expansion and contraction of bacterial interaction networks

The genome variation within the *Gammaproteobacteria*, as caused by high levels of gene acquisition in some taxa and massive gene erosion in others, has allowed us to monitor the process by which PPI interaction networks grow and shrink over evolutionary timescales. Although all biological networks display a similar scale-free topology (Barabási and Albert, '99; Wagner and Fell, 2001; Albert, 2005), changes to the contents and structure of PPI networks are not the same as those observed in other types of networks. In contrast to PPI networks, the components of bacterial metabolic networks tend to attach and detach from the periphery of networks, and essential metabolic functions are not more highly connected and do not have larger numbers of interactors (Pál et al., 2005, 2006; Vitkup et al., 2006).

In contract to PPI and metabolic network, the changes within regulatory networks in response to genome degradation are even more radical. We reconstructed the regulatory network of the gammaproteobacterial ancestor (using *E. coli* homologs of regulatory information in RegulonDB (Salgado et al., 2006)) and carried out analyses analogous to those we performed for PPI. In the most reduced symbiont genomes, virtually no regulatory network remains (Fig. 3), possibly because precise modulation of gene expression is not needed when residing in a stable host environment (Wilcox et al., 2003). Of the 84 regulators present in the ancestor, only four remain in *Buchnera* ($n_{\text{expected}} = 18$; $P < 0.05$), and their connectivity is not significantly different from the average regulator. Regulated elements, which are the least connected elements in the regulatory network, are more conserved than regulators (19% vs. 5%; $P < 0.001$). Whereas higher connectivity within PPI networks protects a gene from being eliminated, high connectivity in the regulatory network increases the gene's susceptibility to loss. Hence, in spite of the apparent topological similarities of biological networks, our analyses reveal that process of genome erosion has altered different types of networks in distinctive ways.

LITERATURE CITED

Albert R. 2005. Scale-free networks in cell biology. J Cell Sci 118:4947–4957.

Albert R, Jeong H, Barabasi AL. 2000. Error and attack tolerance of complex networks. Nature 406:378–382.

Barabási A, Albert R. 1999. Emergence of scaling in random networks. Science 286:509–512.

Butland G, Peregrin-Alvarez JM, Li J, Yang W, Yang X, Canadien V, Starostine A, Richards D, Beattie B, Krogan N, Davey M, Parkinson J, Greenblatt J, Emili A. 2005. Interaction network containing conserved and essential protein complexes in *E. coli*. Nature 433:531–537.

Campillos M, von Mering C, Jensen LJ, Bork P. 2006. Identification and analysis of evolutionarily cohesive functional modules in protein networks. Genome Res 16: 374–382.

Dandekar T, Snel B, Huynen M, Bork P. 1998. Conservation of gene order: a fingerprint of proteins that physically interact. Trends Biochem Sci 23:324–328.

Daubin V, Ochman H. 2004. Bacterial genomes as new gene homes: the genealogy of ORFans in *E. coli*. Genome Res 14: 1036–1042.

de Nooy W, Mrvar A, Batagelj V. 2005. Exploratory social network analysis with Pajek. New York: Cambridge University Press.

Gandhi TK, Zhong J, Mathivanan S, Karthick L, Chandrika KN, Mohan SS, Sharma S, Pinkert S, Nagaraju S, Periaswamy B, Mishra G, Nandakumar K, Shen B, Deshpande N, Nayak R, Sarker M, Boeke JD, Parmigiani G, Schultz J, Bader JS, Pandey A. 2006. Analysis of the human protein interactome and comparison with yeast, worm and fly interaction datasets. Nat Genet 38:285–293.

Gerdes SY, Scholle MD, Campbell JW, Balazsi G, Ravasz E, Daugherty MD, Somera AL, Kyrpides NC, Anderson I, Gelfand MS, Bhattacharya A, Kapatral V, D'Souza M, Baev

MV, Grechkin Y, Mseeh F, Fonstein MY, Overbeek R, Barabási AL, Oltvai ZN, Osterman AL. 2003. Experimental determination and system level analysis of essential genes in *E. coli* MG1655. J Bacteriol 185:5673–5684.

Gil R, Silva FJ, Zientz E, Delmotte F, Gonzalez-Candelas F, Latorre A, Rausell C, Kamerbeek J, Gadau J, Holldobler B, van Ham RC, Gross R, Moya A. 2003. The genome sequence of *Blochmannia floridanus*: comparative analysis of reduced genomes. Proc Natl Acad Sci USA 100:9388–9393.

Gu Z, Steinmetz LM, Gu X, Scharfe C, Davis RW, Li WH. 2003. Role of duplicate genes in genetic robustness against null mutations. Nature 421:63–66.

Hahn MW, Kern AD. 2005. Comparative genomics of centrality and essentiality in three eukaryotic protein-interaction networks. Mol Biol Evol 22:803–806.

Herbeck JT, Degnan PH, Wernegreen JJ. 2005. Nonhomogeneous model of sequence evolution indicates independent origins of primary endosymbionts within the enterobacteriales (G*ammaproteobacteria*). Mol Biol Evol 22:520–532.

Hooper SD, Berg OG. 2003. Duplication is more common among laterally transferred genes than among indigenous genes. Genome Biol 4:R48.

Jeong H, Mason SP, Barabasi AL, Oltvai ZN. 2001. Lethality and centrality in protein networks. Nature 411:41–42.

Jin Q, Yuan Z, Xu J, Wang Y, Shen Y, Lu W, Wang J, Liu H, Yang J, Yang F, Zhang X, Zhang J, Yang G, Wu H, Qu D, Dong J, Sun L, Xue Y, Zhao A, Gao Y, Zhu J, Kan B, Ding K, Chen S, Cheng H, Yao Z, He B, Chen R, Ma D, Qiang B, Wen Y, Hou Y, Yu J. 2002. Genome sequence of *Shigella flexneri* 2a: insights into pathogenicity through comparison with genomes of *E. coli* K12 and O157. Nucleic Acids Res 30: 4432–4441.

Krylov DM, Wolf YI, Rogozin IB, Koonin EV. 2003. Gene loss, protein sequence divergence, gene dispensability, expression level, and interactivity are correlated in eukaryotic evolution. Genome Res 13:2229–2235.

Lerat E, Ochman H. 2004. ψ-φ: exploring the outer limits of bacterial pseudogenes. Genome Res 14:2273–2278.

Lerat E, Daubin V, Ochman H, Moran NA. 2005. Evolutionary origins of genomic repertoires in bacteria. PLoS Biol 3:e130.

Li D, Li J, Ouyang S, Wang J, Wu S, Wan P, Zhu Y, Xu X, He F. 2006. Protein interaction networks of *Saccharomyces cerevisiae*, *Caenorrhabditis elegans* and *Drosophila melanogaster*: large-scale organization and robustness. Proteomics 6:456–461.

Light S, Kraulis P, Elofsson A. 2005. Preferential attachment in the evolution of metabolic networks. BMC Genomics 6:159.

Maslov S, Sneppen K. 2002. Specificity and stability in topology of protein networks. Science 296:910–913.

Matthews LR, Vaglio P, Reboul J, Ge H, Davis BP, Garrels J, Vincent S, Vidal M. 2001. Identification of potential interaction networks using sequence-based searches for conserved protein–protein interactions or ''interlogs''. Genome Res 11:2120–2126.

Moran NA. 2002. Microbial minimalism: genome reduction in bacterial pathogens. Cell 108:583–586.

Moran NA, Mira A. 2001. The process of genome shrinkage in the obligate symbiont *Buchnera aphidicola*. Genome Biol 2:R54.

Nilsson AI, Koskiniemi S, Eriksson S, Kugelberg E, Hinton JC, Andersson DI. 2005. Bacterial genome size reduction by experimental evolution. Proc Natl Acad Sci USA 102: 12112–12116.

Ochman H. 2005. Genomes on the shrink. Proc Natl Acad Sci USA 102:11959–11960.

Pál C, Papp B, Lercher MJ. 2005. Adaptive evolution of bacterial metabolic networks by horizontal gene transfer. Nat Genet 37:1372–1375.

Pál C, Papp B, Lercher MJ, Csermely P, Oliver SG, Hurst LD. 2006. Chance and necessity in the evolution of minimal metabolic networks. Nature 440:667–670.

Pereira-Leal JB, Levy ED, Teichmann SA. 2006. The origins and evolution of functional modules: lessons from protein complexes. Philos Trans R Soc Lond B 361:507–517.

Salgado H, Santos-Zavaleta A, Gama-Castro S, Peralta-Gil M, Penaloza-Spinola MI, Martinez-Antonio A, Karp PD, Collado-Vides J. 2006. The comprehensive updated regulatory network of *E. coli* K-12. BMC Bioinform 7:5.

Sharan R, Suthram S, Kelley RM, Kuhn T, McCuine S, Uetz P, Sittler T, Karp RM, Ideker T. 2005. Conserved patterns of protein interaction in multiple species. Proc Natl Acad Sci USA 102:1974–1979.

Tamas I, Klasson L, Canback B, Naslund AK, Eriksson AS, Wernegreen JJ, Sandstrom JP, Moran NA, Andersson SGE. 2002. 50 million years of genomic stasis in endosymbiotic bacteria. Science 296:2376–2379.

Toh H, Weiss BL, Perkin SA, Yamashita A, Oshima K, Hattori M, Aksoy S. 2006. Massive genome erosion and functional adaptations provide insights into the symbiotic lifestyle of *Sodalis glossinidius* in the tsetse host. Genome Res 16:149–156.

Vitkup D, Kharchenko P, Wagner A. 2006. Influence of metabolic network structure and function on enzyme evolution. Genome Biol 7:R39.

Wagner A. 2000. Robustness against mutations in genetic networks of yeast. Nat Genet 24:355–361.

Wagner A. 2001. The yeast protein interaction network evolves rapidly and contains few redundant duplicate genes. Mol Boil Evol 18:1283–1292.

Wagner A, Fell DA. 2001. The small world inside large metabolic networks. Proc Biol Sci 268:1803–1810.

Wilcox JL, Dunbar HE, Wolfinger RD, Moran NA. 2003. Consequences of reductive evolution for gene expression in an obligate endosymbiont. Mol Microbiol 48:1491–1500.

Wuchty S. 2004. Evolution and topology in the yeast protein interaction network. Genome Res 14:1310–1314.