Available online at www.sciencedirect.com
ScienceDirect
Current Opinion in Biotechnology

Gene dispensability

Ryszard Korona

Genome-wide mutagenesis studies indicate that up to about 90% of genes in bacteria and 80% in eukaryotes can be inactivated individually leaving an organism viable, often seemingly unaffected. Several strategies are used to learn what these apparently dispensable genes contribute to fitness. Assays of growth under hundreds of physical and chemical stresses are among the most effective experimental approaches. Comparative studies of genomic DNA sequences continue to be valuable in discriminating between the core bacterial genome and the more variable niche-specific genes. The concept of the core genome appears currently unfeasible for eukaryotes but progress has been made in understanding why they contain numerous gene duplicates.

Address
Institute of Environmental Sciences, Jagiellonian University, Gronostajowa 7, 30-387 Krakow, Poland

Corresponding author: Korona, Ryszard (ryszard.korona@uj.edu.pl)

Current Opinion in Biotechnology 2011, **22**:547–551

This review comes from a themed issue on
Systems biology
Edited by Roy Kishony and Vassily Hatzimanikatis

Available online 16th May 2011

0958-1669/$ – see front matter

DOI 10.1016/j.copbio.2011.04.017

Introduction

Classic genetic studies required that a phenotypic change was observed first and then a gene responsible for it was identified and possibly scored as a new one. As a result, genes that do not contribute or contribute little to the most often studied traits were largely neglected. After the first complete genomic sequences were released and then sequence-informed mutational studies were completed, researchers were suddenly faced with long lists of genes which were apparently active but could be deleted without causing conspicuous phenotypic effects [1]. These genes were called 'dispensable', the term is now used without quotation marks although its original meaning was that a gene is not necessary for growth under defined conditions. The present review concentrates on recent empirical work that seeks to determine how many genes are seemingly dispensable and whether their scores can be significantly reduced by measuring phenotypic differences with high accuracy and in many environments. It also briefly reports on new comparative studies of genome sequences that approach the question of essentiality and dispensability of genes by examining their history.

Prevalence of dispensable genes

A relatively easy way to obtain large numbers of mutants in bacteria is to prepare a library of clones with possibly single transposon-mediated inserts and then determine their location in the genome. It is assumed that the presence of an alive clone with a disrupted gene marks gene dispensability while absence marks essentiality. This is an obvious source of error as the distribution of transposon insertions is heterogeneous and some loci may be never hit. Furthermore, disruption of a gene need not mean its full inactivation and may influence activity of adjacent genes. Finally, clones with strong growth defects can be viable but are lost during growth in a mutagenized culture. Most of these difficulties are avoided when an alternative approach, individual targeting of every gene, is adopted. Ideally, an entire gene is replaced by a marker cassette. Only a few studies applied the last method so far (Table 1.). The number of essential genes identified in this way tends to be smaller than in random transposon mutagenesis. Indeed, an early insertional study of *Escherichia coli* reported six hundred essential genes [2] but a subsequent deletion study cancelled a half of them [3]. Generally, it appears that for many bacteria, although probably not for all, only about three hundred single gene deletions are lethal under laboratory conditions. This number is roughly similar for organisms of very different lifestyles and total genome sizes.

Yeasts are as easy to manipulate genetically as bacteria, they also offer an opportunity not accessible for the latter, that is, an efficient way of creating and maintaining a lethal knockout in a heterozygote. The first ever organism for which a systematic collection of whole gene deletions was done was *Saccharomyces cerevisiae*, the last reported so far was *Schizosaccharomyces pombe* (Table 1). In case of multicellular eukaryotes, both random and targeted mutagenesis techniques are accessible but saturated genome-wide collections of mutants are essentially absent [4]. Progress is slow because technical difficulties are much greater than in yeasts while costs can be three orders of magnitude higher. In some eukaryotes, the RNA interference can be used to silence virtually every single gene. An experiment of this type was done with *Caenorhabditis elegans* and suggested that individual silencing of about one tenth of genes leads to death or infertility [5]. However, a critical review of literature suggests that a range of 15–30% is more likely [6]. This is because RNA-mediated gene silencing must relatively often result in false findings [7]. In *Drosophila*, both old and new

Table 1

Prevalence of essential genes in unicellular organisms.

Organism	N_{PCG}[a]	N_E (%)[b]	Medium	Method	Ref.
Mycoplasma genitalium	482	382 (79)	Rich	Random transposon insertion	[39]
Mycoplasma pulmonis	782	405 (52)	Rich	Random transposon insertion	[40]
Francisella novicida	1574	396 (25)	Rich	Random transposon insertion	[41]
Haemophilus influenzae	1737	670 (38)	Rich	Random transposon insertion	[42]
Staphylococcus aureus	2632	351 (13)	Rich	Random transposon insertion	[43]
Corynebacterium glutamicum	2993	658 (22)	Rich	Random transposon insertion	[44]
Acinetobacter baylyi	3197	499 (16)	Minimal	Targeted whole gene deletion	[45]
Vibrio cholerae	3887	789 (20)	Rich	Random transposon insertion	[46]
Mycobacterium tuberculosis	3918	614 (16)	Rich	Random transposon insertion	[47]
Bacillus subtilis	4245	271 (6.4)	Rich	Targeted plasmid insertion	[48]
Escherichia coli	4288	299 (8.0)	Rich	Targeted whole gene deletion	[3]
Salmonella typhimurium	4425	257 (5.8)	Rich	Random insertion-duplication	[49]
Pseudomonas aeruginosa	5565	335 (6.0)	Rich	Random transposon insertion	[50]
Schizosaccharomyces pombe	4914	1260 (26)	Rich	Targeted whole gene deletion	[51••]
Saccharomyces cerevisiae	5797	1105 (19)	Rich	Targeted whole gene deletion	[1]

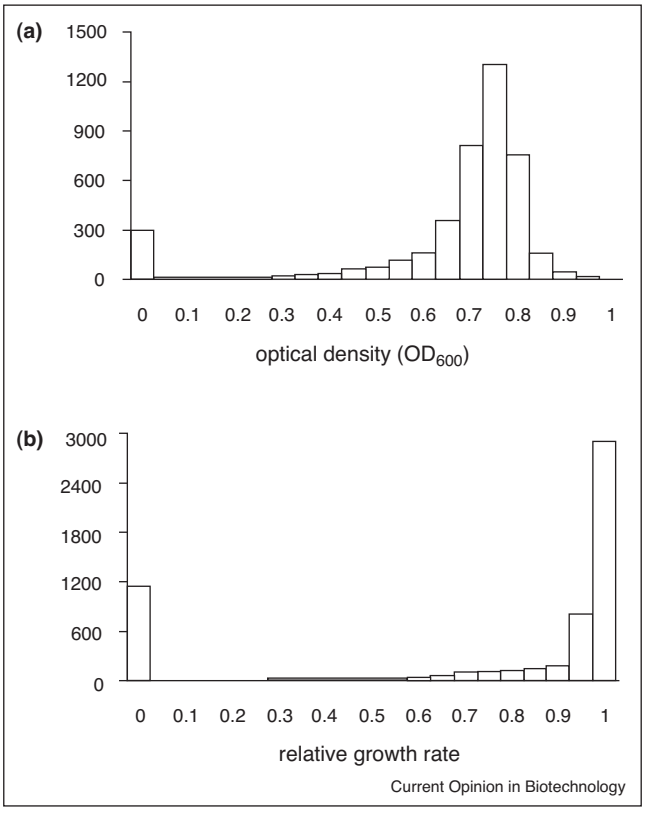
[a] Total number of protein coding genes according to the most recent publications or specialized web pages as of February 2011.
[b] Number of essential genes as stated in the cited source publications.

experiments indicate that some 25–35% of genes are essential [8,9••]. In mouse, over 40% of existing null mutations obtained through targeted insertion or deletion results in lethality or infertility but the sample covers less than one fifth of genes and is highly biased [10,11]. Chemical mutagenesis of mouse, dense but inherently random, yielded highly variable estimates with a range from 20 to 40%, or wider [12]. In sum, eukaryotes have probably several times more essential genes than prokaryotes, both in terms of numbers and proportions. Unlike in bacteria, it is rather the proportion than the number of essential genes that may be roughly similar in different eukaryotic model organisms.

Measuring gene (in)dispensability

Mutagenesis and screening for phenotypic effects of mutations on a polygenic trait, or fitness, typically yields a bimodal frequency distribution with one peak formed by lethal and the other by weak effects [13]. The abundance of small phenotypic changes appeared understandable as long as the applied mutagenesis frequently resulted in only partial gene inactivation. But, whole gene deletions also generate a bimodal distribution with an especially strong peak of small or unobservable effects (Figure 1). The bimodality is seen not only when the whole genome is considered but also within its subsets, such as genes coding for large protein complexes [14]. This leads to a basic question: what saves 'dispensable' genes from a gradual mutational decay by accumulation of minor lesions if complete loss of their activity does so little harm [15]? It has been long speculated that their metabolic role and thus the selective pressures to maintain them active can be much stronger in environments other than the standard laboratory ones. In recent studies hundreds of different chemical and physical stresses were applied and the proportion of deletions showing substantial growth defects under at least some conditions reached about 95% in *S. cerevisiae* and at least 49% in *E. coli* [16,17•]. The tested environments were in a sense unnatural but they at least helped to show that many genes become vulnerable to selection when metabolism is dis-

Figure 1



Distribution of fitness effects in genomic collection of gene deletions. **(a)** Density of cultures composed of individual *E. coli* deletion strains after a defined period of growth [3]. **(b)** Relative growth rates of homozygous *S. cerevisiae* deletion strains propagated in a common culture [52].

torted. Careful analysis of conditional growth effects can be used not only to show that the seemingly dispensable genes are useful but also what they can be used for [18].

It is also possible that genes were judged dispensable prematurely, that is, based on experiments that were not sensitive enough to detect small contributions to fitness. Indeed, even under standard laboratory conditions small but statistically significant fitness effects were frequently found when yeast strains with gene disruptions were confronted with a wild type strain in relatively long competition experiments [19]. Precise measurements of fitness were also used in the study of dispensability of paralogs originating from the whole-genome duplication in yeast. Loss of a single duplicate gene usually causes little harm but it was unsure whether this is because the duplicates genuinely compensate for each other or are simply unimportant. Careful estimations of the growth rate and the level of gene expression in strains lacking paralogs pointed to the compensation [20,21•]. Extensive competition experiments were also used to test whether loss of a dispensable gene can be actually advantageous instead of being deleterious. In yeast, loss of a gene is rarely associated with even a small fitness gain suggesting that there is little selection pressure to lose genes that are dispensable in a particular environment [22,23]. However, the assays of fitness have their limits, effects smaller than about 0.005 are currently difficult to detect. There is no guarantee that future advances in technology will help because fitness differences of this size will be always liable to the obscuring effect of small and thus difficult to control fluctuations of both the environment and genetic background.

Core genome

One question pertaining to the unequal importance of different genes is the rate at which their DNA sequence changes. In bacteria, essential genes tend to be more conserved than nonessential over both relatively short and long time scales [24]. This is only a general trend with many exceptions because the rate at which protein-coding genes evolve depends on several factors other than dispensability [25]. A more direct measure of gene significance is its distribution among different taxa. In a recent analysis of 579 sequenced eubacterial genomes, about 250 genes (gene families) per genome were identified as belonging to an 'extended core'. It means that they are present in at least 99% of all genomes. The core set is enriched in genes responsible for replication, translation, and energy homeostasis [26••]. The number and functions of the core genes are generally similar to those of essential genes. However, the genome-wide screens for essential genes are much fewer than genome sequencing projects and therefore definite comparisons of their results may be premature. The majority of an average bacterial genome is composed by other genes. They are present in only subsets of genomes and are often characteristic for different ecological niches. About a quarter of a typical bacterial genome still consists of other genes. These are found in only one or a few species; their variety is very high; their role is usually unknown but probably only accessory in most cases [26••]. The existence of core genes and niche-characteristic genes is compatible with the concept of a backbone of essential elements on which the rest of the genome is built [27]. Lateral gene transfer is omnipresent in bacteria and this could suggest that the genes accessible for ecological specialization constitute an effectively unlimited pool. This idea was tested in a study comparing 96 genome sequences derived from two closely related sympatric sister species of pathogenic bacteria (*Campylobacter coli* and *Campylobacter jejuni*). It was found that the two species have similarly sized genomes. Their core genome contains a number of species-specific genes and, importantly, demonstrates a resistance to interspecies recombination [28•]. These data not only provide support for the hypothesis of core genome in bacteria but also show that the core genes can differ even between relatively close species.

Phylogenetic analyses are much less, if at all, effective in delimiting the core genome of eukaryotes. Five yeast species that diverged before the whole genome duplication in the lineage of *S. cerevisiae* (100–150 million years ago) were recently sequenced and annotated. Their common genetic repertoire consists of approximately 3300 protein families, within a pan-proteome of approximately 5000 families for all Saccharomycetaceae. This is far from any minimal genome without redundancy especially that the five yeasts contain numerous copies of paralogous genes that, altogether, constitute a third of each genome [29•]. The quest to find a minimal gene set is effectively absent in research on multicellular eukaryotes. This is not surprising as counting of essential genes is still a problem there. But, the uncertainty about the number of protein coding genes is probably not the most important one. Still more serious are doubts what to count when new important noncoding elements are being constantly discovered and their total number cannot be yet anticipated [30]. The very notion of gene dispensability changes profoundly when we look at organisms like us. Considering the fact that genes which have paralogs of 90% sequence identity are about three times less likely to harbor known disease mutations [31] and that every human has about a hundred of such variants [32], one is left with an impression that his/her health may be actually secured by some 'redundant' or 'dispensable' genes.

Gene duplication

The above mentioned phenotypic masking of genetic damage by close relatives pertains chiefly to genes involved in metabolism. Such genes constitute a substantial portion of a genome in simple organisms but much smaller in, say, a mouse. This can be a reason why deleting one of a duplicate gene is less likely to be lethal

than deleting a singleton gene in yeasts and worms but not necessarily in mice [33]. It was actually claimed that in mouse singletons are as often essential as duplicates but definite tests are lacking as long as the collection of knockouts is not representative for the whole genome [11]. Nevertheless, the already available data suggest that duplicate genes are more often found among those coding for development than metabolism. For example, a recent study reports that new genes in *Drosophila* are equally often essential as old ones. Lethality caused by silencing of the new genes was often attributable to defects showing up at the pupal stage. This suggests an origin of essentiality through rapid neofunctionalization of a duplicate gene engaged in development [9]. Not only the expectation that new genes are less important than old ones but also a common assumption that most new genes come typically from whole genome duplications was challenged. In *Daphnia* as many as 13,000 genes — a very high number — are paralogs. These genes apparently arose through frequently occurring tandem duplications and are enriched in elements responsive to specific ecological conditions. Fine tuned adaptations to the environment can be another important factor promoting rapid evolution of new functions even if it is done mostly through differentiation of duplicate gene expression [34••]. Together these new findings promise to invigorate research on specific hypotheses explaining how duplicate, and thus normally dispensable, genes can avoid inactivation and disappearance [35,36].

Conclusions

A more appropriate name for the dispensable genes would be the genes of small or rare effects. Incorporation of such genes into systems biology will be difficult, this is already well known by students of quantitative and population genetics. One way out is to simply ignore the genes of small effects and concentrate on a possibly minimal set of essential genes. This approach is popular among researchers in the systems biology of bacteria. Its promises and challenges were recently reviewed in this journal [37]. The study of simple eukaryotic cells will possibly try to enter a similar path because the metabolic flux in the yeast cell can be modeled as reliably as in bacteria [38]. It may be much more difficult to build grand models of gene cooperation in assembling the cellular 'hardware', especially in eukaryotes, even as simple as yeast. The idea of concentrating on a set of essential genes and leaving aside dispensable ones is still less feasible for multicellular eukaryotes. However, the thorough study of well defined problems, such as the evolutionary fate and current role of duplicated genes, can be already seen as the 'subsystems biology' with good prospects for a gradual expansion.

Acknowledgements

This work was supported by the Foundation for Polish Science (The 'Mistrz' Program) and an MNiSW grant (3970/B/P01/2009/36).

References and recommended reading

Papers of particular interest, published within the period of review, have been highlighted as:

- • of special interest
- •• of outstanding interest

1. Giaever G, Chu AM, Ni L, Connelly C, Riles L, Véronneau S, Dow S, Lucau-Danila A, Anderson K, André B *et al.*: **Functional profiling of the *Saccharomyces cerevisiae* genome**. *Nature* 2002, **418**:387-391.

2. Gerdes SY, Scholle MD, Campbell JW, Balázsi G, Ravasz E, Daugherty MD, Somera AL, Kyrpides NC, Anderson I, Gelfand MS *et al.*: **Experimental determination and system level analysis of essential genes in *Escherichia coli* MG1655**. *J Bacteriol* 2003, **185**:5673-5684.

3. Baba T, Ara T, Hasegawa M, Takai Y, Okumura Y, Baba M, Datsenko KA, Tomita M, Wanner BL, Mori H: **Construction of *Escherichia coli* K-12 in-frame, single-gene knockout mutants: the Keio collection**. *Mol Syst Biol* 2006, **2**:0008.

4. Hardy S, Legagneux V, Audic Y, Paillard L: **Reverse genetics in eukaryotes**. *Biol Cell* 2010, **102**:561-580.

5. Kamath RS, Fraser AG, Dong Y, Poulin G, Durbin R, Gotta M, Kanapin A, Le Bot N, Moreno S, Sohrmann M *et al.*: **Systematic functional analysis of the *Caenorhabditis elegans* genome using RNAi**. *Nature* 2003, **421**:231-237.

6. Kemphues K: **Essential genes**. *WormBook* 2005:1-7.

7. Mohr S, Bakal C, Perrimon N: **Genomic screening with RNAi: results and challenges**. *Annu Rev Biochem* 2010, **79**:37-64.

8. Dietzl G, Chen D, Schnorrer F, Su KC, Barinova Y, Fellner M, Gasser B, Kinsey K, Oppel S, Scheiblauer S *et al.*: **A genome-wide transgenic RNAi library for conditional gene inactivation in *Drosophila***. *Nature* 2007, **448**:151-156.

9. •• Chen S, Zhang YE, Long M: **New genes in *Drosophila* quickly become essential**. *Science* 2010, **330**:1682-1685.
A sample of 195 young protein-coding genes which originated 3–35 million years ago in the *Drosophila* lineage were knocked down with RNA interference. As many as 30% of these genes were essential for viability, as much as in every other evolutionary age group that was examined. Lethality was high in the pupal stage and was attributed to diverse cellular and developmental defects. These data suggest that new genes can evolve essential functions more frequently and rapidly than previously expected.

10. Liao BY, Zhang J: **Mouse duplicate genes are as essential as singletons**. *Trends Genet* 2007, **23**:378-381.

11. Su Z, Gu X: **Predicting the proportion of essential genes in mouse duplicates based on biased mouse knockout genes**. *J Mol Evol* 2008, **67**:705-709.

12. Hentges KE, Pollock DD, Liu B, Justice MJ: **Regional variation in the density of essential genes in mice**. *PLoS Genet* 2007, **3**:e72.

13. Eyre-Walker A, Keightley PD: **The distribution of fitness effects of new mutations**. *Nat Rev Genet* 2007, **8**:610-618.

14. Fudala A, Korona R: **Low frequency of mutations with strongly deleterious but nonlethal fitness effects**. *Evolution* 2009, **63**:2164-2171.

15. Lynch M, Conery JS: **The evolutionary fate and consequences of duplicate genes**. *Science* 2000, **290**:1151-1155.

16. Hillenmeyer ME, Fung E, Wildenhain J, Pierce SE, Hoon S, Lee W, Proctor M, St Onge RP, Tyers M, Koller D *et al.*: **The chemical genomic portrait of yeast: uncovering a phenotype for all genes**. *Science* 2008, **320**:362-365.

17. • Nichols RJ, Sen S, Choo YJ, Beltrao P, Zietek M, Chaba R, Lee S, Kazmierczak KM, Lee KJ, Wong A *et al.*: **Phenotypic landscape of a bacterial cell**. *Cell* 2011, **144**:143-156.
Growth profiles of a mutant library in hundreds of stressful conditions were done in parallel. This allowed to study gene essentiality with a much greater discovery rate than in a typical laboratory environment.

18. Shuman HA, Silhavy TJ: **The art and design of genetic screens: *Escherichia coli***. *Nat Rev Genet* 2003, **4**:419-431.

19. Thatcher JW, Shaw JM, Dickinson WJ: **Marginal fitness contributions of nonessential genes in yeast**. *Proc Natl Acad Sci U S A* 1998, **95**:253-257.

20. DeLuna A, Vetsigian K, Shoresh N, Hegreness M, Colon-Gonzalez M, Chao S, Kishony R: **Exposing the fitness contribution of duplicated genes**. *Nat Genet* 2008, **40**:676-681.

21. • DeLuna A, Springer M, Kirschner MW, Kishony R: **Need-based up-regulation of protein levels in response to deletion of their duplicate genes**. *PLoS Biol* 2010, **8**:e1000347.
It has been postulated, although never tested, that the mild effect of deleting one of duplicates might be mediated by active up-regulation of the other duplicate. This study reports that such response is present in only a modest fraction of proteins but almost exclusively among those duplicates whose overlapping function is required for growth in the test environment.

22. Blanc VM, Adams J: **Evolution in *Saccharomyces cerevisiae*: identification of mutations increasing fitness in laboratory populations**. *Genetics* 2003, **165**:975-983.

23. Sliwa P, Korona R: **Loss of dispensable genes is not adaptive in yeast**. *Proc Natl Acad Sci U S A* 2005, **102**:17670-17674.

24. Jordan IK, Rogozin IB, Wolf YI, Koonin EV: **Essential genes are more evolutionarily conserved than are nonessential genes in bacteria**. *Genome Res* 2002, **12**:962-968.

25. Pal C, Papp B, Lercher MJ: **An integrated view of protein evolution**. *Nat Rev Genet* 2006, **7**:337-348.

26. •• Lapierre P, Gogarten JP: **Estimating the size of the bacterial pan-genome**. *Trends Genet* 2009, **25**:107-110.
Using 573 sequenced genomes, the authors estimate that about 250 genes per genome belong to a bacterial core genome, about one fourth of a genome consists of poorly defined accessory elements, while the rest of genes are 'characteristic' for a genome, that is, those are present in only subsets of species and often those of similar lifestyles.

27. Koonin EV: **Comparative genomics, minimal gene-sets and the last universal common ancestor**. *Nat Rev Microbiol* 2003, **1**:127-136.

28. • Lefebure T, Bitar PD, Suzuki H, Stanhope MJ: **Evolutionary dynamics of complete *Campylobacter* pan-genomes and the bacterial species concept**. *Genome Biol Evol* 2010, **2**:646-655.
An analysis of 96 genome sequences derived from two closely related sympatric sister species of pathogenic bacteria showed that their common pan-genome is finite while their core genomes demonstrate a resistance to interspecies recombination.

29. • Souciet JL, Dujon B, Gaillardin C, Johnston M, Baret PV, Cliften P, Sherman DJ, Weissenbach J, Westhof E, Wincker P *et al.*: **Comparative genomics of protoploid Saccharomycetaceae**. *Genome Res* 2009, **19**:1696-1709.
A rare example of a study searching for a 'core genome' in eukaryotes.

30. Cooper DN, Chen JM, Ball EV, Howells K, Mort M, Phillips AD, Chuzhanova N, Krawczak M, Kehrer-Sawatzki H, Stenson PD: **Genes, mutations, and human inherited disease at the dawn of the age of personalized genomics**. *Hum Mutat* 2010, **31**:631-655.

31. Hsiao TL, Vitkup D: **Role of duplicate genes in robustness against deleterious human mutations**. *PLoS Genet* 2008, **4**:e1000014.

32. Durbin RM, Abecasis GR, Altshuler DL, Auton A, Brooks LD, Gibbs RA, Hurles ME, McVean GA: **A map of human genome variation from population-scale sequencing**. *Nature* 2010, **467**:1061-1073.

33. Makino T, Hokamp K, McLysaght A: **The complex relationship of gene duplication and essentiality**. *Trends Genet* 2009, **25**:152-155.

34. •• Colbourne JK, Pfrender ME, Gilbert D, Thomas WK, Tucker A, Oakley TH, Tokishita S, Aerts A, Arnold GJ, Basu MK *et al.*: **The ecoresponsive genome of *Daphnia pulex***. *Science* 2011, **331**:555-561.
The genome of *Daphnia* gained more genes by tandem duplication than whole genome duplication. The authors postulate that duplicates quickly engage into novel combinations of interacting genes in different phases of development or under different environmental conditions. They actually provide evidence that paralogs often diverge in their expression patterns.

35. Semon M, Wolfe KH: **Consequences of genome duplication**. *Curr Opin Genet Dev* 2007, **17**:505-512.

36. Hahn MW: **Distinguishing among evolutionary models for the maintenance of gene duplicates**. *J Hered* 2009, **100**:605-617.

37. Jewett MC, Forster AC: **Update on designing and building minimal cells**. *Curr Opin Biotechnol* 2010, **21**:697-703.

38. Famili I, Forster J, Nielsen J, Palsson BO: ***Saccharomyces cerevisiae* phenotypes can be predicted by using constraint-based analysis of a genome-scale reconstructed metabolic network**. *Proc Natl Acad Sci U S A* 2003, **100**:13134-13139.

39. Glass JI, Assad-Garcia N, Alperovich N, Yooseph S, Lewis MR, Maruf M, Hutchison CA 3rd, Smith HO, Venter JC: **Essential genes of a minimal bacterium**. *Proc Natl Acad Sci U S A* 2006, **103**:425-430.

40. Dybvig K, Lao P, Jordan DS, Simmons WL: **Fewer essential genes in mycoplasmas than previous studies suggest**. *FEMS Microbiol Lett* 2010, **311**:51-55.

41. Gallagher LA, Ramage E, Jacobs MA, Kaul R, Brittnacher M, Manoil C: **A comprehensive transposon mutant library of *Francisella novicida*, a bioweapon surrogate**. *Proc Natl Acad Sci U S A* 2007, **104**:1009-1014.

42. Akerley BJ, Rubin EJ, Novick VL, Amaya K, Judson N, Mekalanos JJ: **A genome-scale analysis for identification of genes required for growth or survival of *Haemophilus influenzae***. *Proc Natl Acad Sci U S A* 2002, **99**:966-971.

43. Chaudhuri RR, Allen AG, Owen PJ, Shalom G, Stone K, Harrison M, Burgis TA, Lockyer M, Garcia-Lara J, Foster SJ *et al.*: **Comprehensive identification of essential *Staphylococcus aureus* genes using Transposon-Mediated Differential Hybridisation (TMDH)**. *BMC Genomics* 2009, **10**:291.

44. Suzuki N, Okai N, Nonaka H, Tsuge Y, Inui M, Yukawa H: **High-throughput transposon mutagenesis of *Corynebacterium glutamicum* and construction of a single-gene disruptant mutant library**. *Appl Environ Microbiol* 2006, **72**:3750-3755.

45. de Berardinis V, Vallenet D, Castelli V, Besnard M, Pinet A, Cruaud C, Samair S, Lechaplais C, Gyapay G, Richez C *et al.*: **A complete collection of single-gene deletion mutants of *Acinetobacter baylyi* ADP1**. *Mol Syst Biol* 2008, **4**:174.

46. Cameron DE, Urbach JM, Mekalanos JJ: **A defined transposon mutant library and its use in identifying motility genes in *Vibrio cholerae***. *Proc Natl Acad Sci U S A* 2008, **105**:8736-8741.

47. Sassetti CM, Boyd DH, Rubin EJ: **Genes required for mycobacterial growth defined by high density mutagenesis**. *Mol Microbiol* 2003, **48**:77-84.

48. Kobayashi K, Ehrlich SD, Albertini A, Amati G, Andersen KK, Arnaud M, Asai K, Ashikaga S, Aymerich S, Bessieres P *et al.*: **Essential *Bacillus subtilis* genes**. *Proc Natl Acad Sci U S A* 2003, **100**:4678-4683.

49. Knuth K, Niesalla H, Hueck CJ, Fuchs TM: **Large-scale identification of essential *Salmonella* genes by trapping lethal insertions**. *Mol Microbiol* 2004, **51**:1729-1744.

50. Liberati NT, Urbach JM, Miyata S, Lee DG, Drenkard E, Wu G, Villanueva J, Wei T, Ausubel FM: **An ordered, nonredundant library of *Pseudomonas aeruginosa* strain PA14 transposon insertion mutants**. *Proc Natl Acad Sci U S A* 2006, **103**:2833-2838.

51. •• Kim DU, Hayles J, Kim D, Wood V, Park HO, Won M, Yoo HS, Duhig T, Nam M, Palmer G *et al.*: **Analysis of a genome-wide set of gene deletions in the fission yeast *Schizosaccharomyces pombe***. *Nat Biotechnol* 2010, **28**:617-623.
Eight years after *S. cerevisiae*, *S. pombe* as the only other eukaryotic species was used to construct a systematic genome-wide collection of single gene deletions. As much as 83% of single-copy orthologs in the two yeasts have conserved dispensability despite the vast evolutionary distance between them.

52. Steinmetz LM, Scharfe C, Deutschbauer AM, Mokranjac D, Herman ZS, Jones T, Chu AM, Giaever G, Prokisch H, Oefner PJ *et al.*: **Systematic screen for human disease genes in yeast**. *Nat Genet* 2002, **31**:400-404.