

Bacterial gene amplification: implications for the evolution of antibiotic resistance

Linus Sandegren and Dan I. Andersson

Abstract | Recent data suggest that, in response to the presence of antibiotics, gene duplication and amplification (GDA) constitutes an important adaptive mechanism in bacteria. For example, resistance to sulphonamide, trimethoprim and β -lactams can be conferred by increased gene dosage through GDA of antibiotic hydrolytic enzymes, target enzymes or efflux pumps. Furthermore, most types of antibiotic resistance mechanism are deleterious in the absence of antibiotics, and these fitness costs can be ameliorated by increased gene dosage of limiting functions. In this Review, we highlight the dynamic properties of gene amplifications and describe how they can facilitate adaptive evolution in response to toxic drugs.

Bacterial fitness

The ability of the microorganism to survive and generate progeny. Bacterial fitness is often measured by growth rate.

Gene amplification is found in all three kingdoms of life and is important both from a fundamental evolutionary perspective^{1–3} and in medical genetics as an important contributor to many human diseases, phenotypic variability among individuals and human susceptibility to infectious diseases^{4,5}. Furthermore, gene amplification plays an important part in generating the genomic variability that is available for genetic adaptation to altered growth conditions or various stresses. During the past 30 years, studies in bacteria, as well as a range of eukaryotic pathogenic microorganisms (BOX 1), plants^{6–8}, insects^{9,10} and human tumours^{11,12}, have shown that gene duplication and amplification (GDA) is a common adaptive mechanism in response to the presence of various toxic drugs.

Bacteria can adapt to the presence of toxic levels of antibiotics using several types of response, including regulatory mechanisms that alter cellular physiology or genetic changes (such as mutation or horizontal gene transfer (HGT) of resistance determinants) that lead to the degradation or sequestration of the antibiotic, prevent its uptake, pump it out of the cell or prevent it binding to the target molecule^{13,14}. Gene amplification has been shown to confer resistance by causing the overproduction of antibiotic-modifying enzymes, target molecules and efflux pumps. Bacterial adaptation to antibiotics typically proceeds in two steps. First, the initial resistance mechanism is acquired, which in most cases confers a reduction in bacterial fitness (FIG. 1). Subsequently, bacteria will often adapt to the presence

of the costly resistance mechanism by acquiring additional compensatory mutations that reduce the fitness cost, often without loss of the resistance. As shown by recent work, such compensatory evolution can occur through gene amplification of the affected target gene or unrelated genes¹⁵.

Importantly, even though GDA involves a genetic alteration, it has properties that are different from a typical genetic adaptation by mutation or HGT. Although GDA is highly prevalent, gene duplications are far more unstable than mutations or HGT, making GDA more similar to a regulatory response (FIG. 2). The high prevalence and instability of GDAs not only provide them with properties that are different from those of typical stable mutations but also make them technically difficult to identify and study (BOX 2).

Dynamics of gene amplification

Much of our knowledge about the dynamics of GDA is based on studies of *Salmonella enterica* subsp. *enterica* serovar Typhimurium LT2, and the data provided below were obtained mainly from this strain. With the exception of the replication terminus, tandem genetic duplications have been found in all regions of the chromosome¹⁶ in sizes that range from a few kilobases to several megabases^{15,17–21}. On the basis of the frequencies and sizes of the amplified regions, it is estimated that 10% of all bacterial cells in a growing culture contain a gene duplication somewhere in their genome²².

Department of Medical Biochemistry and Microbiology, Uppsala University, BOX 582, Uppsala, S-751 23, Sweden. Correspondence to D.I.A. e-mail: Dan.Andersson@imbim.uu.se
doi:10.1038/nrmicro2174

Box 1 | Drug resistance in microorganisms other than bacteria

Similarly to the situation in bacteria, gene duplication and amplification (GDA) has also been detected in a number of eukaryotic microorganisms in response to drug exposure, including *Candida*, *Plasmodium* and *Leishmania* species. For example, *Candida glabrata* isolates that are resistant to azole antifungals (such as fluconazole) have an increased copy number of the *CYP51* gene, which encodes 14-lanosterol demethylase (the target for azoles). When fungi were grown in the absence of the drug, there was a gradual and parallel reduction in both *CYP51* gene copy number and the level of resistance⁹⁹.

Several studies of laboratory-selected and natural isolates of *Plasmodium falciparum*^{100,101} and *Plasmodium vivax*¹⁰² that are resistant to mefloquine, halofantrine and quinine show that resistance and treatment failure¹⁰³ is associated with an increased copy number and expression of multidrug resistance gene 1 (*pfmdr1* in *P. falciparum* and *pvmdr1* in *P. vivax*). The *pfmdr1* gene encodes Pgh1, a homologue to human P glycoprotein 1, which is found in the membrane of the parasite's vacuole and confers resistance to several drugs¹⁰⁴ by altering drug transport across the membrane. The human P glycoprotein is expressed by anticancer drug-resistant cells that efflux the drug from the cells^{12,105,106}. Similarly, malaria parasites that are resistant to antifolate drugs have an increased copy number (up to 11-fold) of the *gch1* gene, encoding the first enzyme, GTP cyclohydrolase 1, in the folate biosynthetic pathway. As antifolate drugs target downstream enzymes in this pathway, it was suggested that increased expression of GTP cyclohydrolase 1 provides a compensatory resistance mechanism¹⁰⁷.

Finally, antimony-resistant *Leishmania infantum* and arsenite-resistant *Leishmania amazonensis* were found to have increased copy numbers of various genes that have been implicated in resistance to these drugs^{108,109}. In the antimony-resistant *L. infantum* strain MRPA, an ATP-binding cassette gene known to be involved in antimony resistance was found to be overexpressed in the resistant strain. This locus was flanked by 1.4 kb of repeated sequences from which an extrachromosomal circular amplicon was generated in the resistant cells.

The frequency of any specific duplication is determined by the rates of formation and loss of the duplication and by any alterations in fitness associated with the altered gene dosage. Using transduction assays (BOX 2), the duplication frequency at different locations in the *S. Typhimurium* chromosome was found to vary between 3×10^{-2} and 6×10^{-5} per cell, the most common being between directly repeated *rrn* operons that provide long regions of homology for recombination^{16,19,23}. To measure the formation rate for duplication of a specific region or gene, one needs to begin with a population of cells that contain no pre-existing duplications of that region (typically <100 to 1,000 cells) and, over time, follow the initial rate of appearance of duplications. Because the rates of duplication formation and loss are generally high, the population of cells rapidly reaches equilibrium. Therefore, to measure the initial formation rate, small populations of cells must be used, which makes it technically difficult to

obtain sufficient amounts of cells or DNA to perform the measurement. Because of these difficulties, the formation rates have been indirectly calculated from measurements of loss rates and steady-state frequencies of the duplications. Loss rates are easier to measure because, by genetic placement of a selectable marker (for example, an antibiotic resistance marker) at the duplication join point, one can stably maintain the duplication in a population of cells by continuous selection. To determine the segregation rate, a large population of cells is switched from growth under selection to growth in the absence of selection, and the duplication loss rate is determined from the loss rate of the selectable marker. The rates that have been determined range from 0.01 to 0.15 per cell per generation, implying that if duplications are not continuously selected they will be rapidly lost from the population^{22,23}. Therefore, on the basis of the segregation rate (k_{loss}), the steady-state frequency of the duplication under non-selective conditions and the assumption that the duplication confers no fitness cost, one can simply estimate the formation rate of the initial duplication ($k_{\text{duplication}}$) according to the formula $k_{\text{duplication}} = \text{steady-state frequency} / k_{\text{loss}}$. These calculations suggest that the rate of formation of spontaneous duplications in *S. Typhimurium* varies between 10^{-2} to 10^{-5} per cell per generation, depending on the duplication²³.

The fitness costs (that is, reduced growth rates) that are typically associated with duplications have been rigorously measured in only a few cases. *S. Typhimurium* strains containing genetically engineered duplications (ranging in size from 72 kb to 1,246 kb) showed variable fitness costs, with phenotypes that ranged from being indistinguishable from the wild type to showing a reduction in growth rate of 20% per generation²³, without any correlation between the size of the duplicated region and the fitness cost. These results indicate that the cost is determined mainly by the genetic content of the duplicated region rather than by the metabolic cost of acquiring extra DNA. Following on from this, the gene dosage balance hypothesis suggests

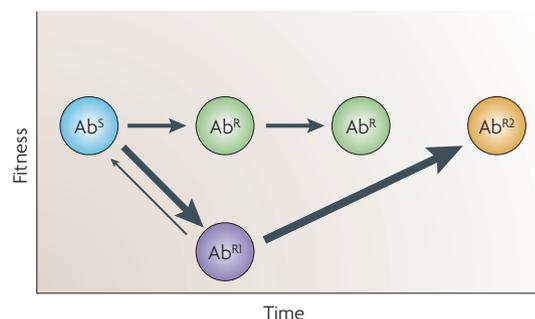


Figure 1 | **Bacterial adaptation to antibiotics.** Adaptation to antibiotics by antibiotic-susceptible (Ab^S) bacteria usually proceeds in a two-step process. First, a resistance mechanism is selected (Ab^R and Ab^{R1} ; Ab^R designates those rare cases in which no fitness cost is incurred and Ab^{R1} designates the more common cases with associated fitness costs). Subsequently, Ab^{R1} bacteria adapt to this fitness-reducing mechanism by acquiring compensatory mutations (Ab^{R2}). The thickness of the arrows indicates the relative rates^{93,94}.

Transduction assays

The study of DNA transfer from one bacterium to another through a bacteriophage in order to assess the duplication frequency of a gene. The transferred DNA contains a selectable marker in the gene of interest.

that stoichiometric imbalances in macromolecular complexes cause the reduced fitness when gene dosage is altered as a result of GDA²⁴. For example, overproduction of a specific subunit that is part of a macromolecular complex might titrate out other subunits in the same complex. Another potential reason for these costs could be that the overproduced proteins have negative effects on protein folding pathways and chaperone function.

In summary, in the absence of selection pressure for a particular GDA, its frequency in a bacterial population will depend on the forward and backward rates of duplication and amplification, as well as the relative fitness costs of all of the intermediates. Obviously, if a selective pressure is applied that favours cells containing the GDA, the steady-state frequency of amplified cells will also be influenced by the strength of this selection.

Mechanisms of GDA

Several mechanisms have been suggested that could increase gene copy number, and it is likely that different processes are involved in the formation of simple duplications and higher-level amplifications^{3,17,22,25–27}. With regard to duplication formation, at least two different mechanisms have been observed. Clearly some duplications arise by non-equal homologous recombination between long, directly orientated repeat sequences (FIG. 3a), such as ribosomal RNA operons^{16,28}, insertion sequences or transposable elements^{29–36}, and repetitive sequences such as repetitive extragenic palindromic sequences^{37–40} that are present on sister chromatids by a RecA-dependent mechanism (discussed below). However, other duplications that lack repeat sequences at the duplication join point arise frequently in a RecA-independent manner. This implies that processes capable of using short repeat sequences or random end joining in the absence of any repetitive sequence are responsible for the formation of some spontaneous genetic duplications.

Once a duplication is formed, higher-level amplification can occur by RecA-dependent recombination between the long perfect tandem repeat that is generated by the initial duplication⁴¹ or, alternatively, without any duplication intermediates by mechanisms based on rolling circle replication (RCR)^{22,27} (FIG. 3b). RCR requires the formation of a replication fork associated with RecA-dependent double-strand break repair followed by RCR to generate a long tandem array. In contrast to non-equal crossing over, RCR can generate a large tandem array in a single generation^{3,17,25}.

The number of amplified units varies between 2-fold and 40-fold, and their size varies from a few kilobases to up to several megabases, as shown by studies in *S. Typhimurium* and *Escherichia coli*, in which increased gene dosage through GDA was selected^{17,38,42,43} (TABLE 1). These numbers indicate that, under certain selective conditions, the amount of DNA in the cell can be almost doubled owing to high-level amplification of specific regions^{17,42}.

With respect to the involvement of RecA in duplication formation, homologous recombination mediated by RecA between long repetitive sequences is an important contributor, but other RecA-independent processes that use short or no repetitive sequences are responsible

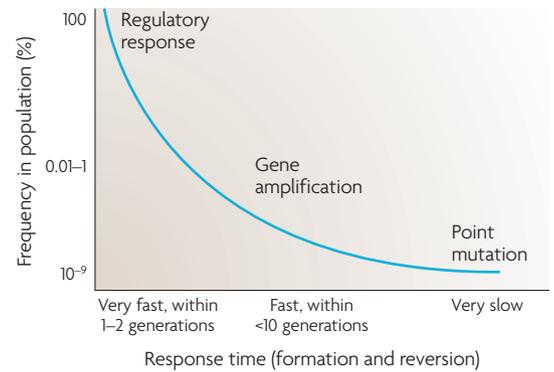


Figure 2 | Frequency and response time for point mutations, gene duplication and regulatory responses in a population. Regulatory changes in gene expression can occur rapidly in a large part of a population in response to changes in the growth environment. Gene duplications are present in a subpopulation of cells, and further gene amplification can be selected in response to changes in the growth environment. Specific point mutations are rare in a population, and it therefore takes a long time before changes in the population as a result of these mutations are seen.

for the formation of at least some duplications (TABLE 1). Evidence that RecA is not essential for the formation of certain duplications comes mainly from analysis of the join points of duplicated sequences. As RecA-mediated homologous recombination requires 20–40 nucleotides of sequence identity to form a recombinogenic complex^{44,45}, this sets a lower limit on the lengths of repeat sequences that can be used as substrates for RecA-mediated exchanges. However, short repetitive sequences (<20 bp) have been implicated in the formation of genetic duplications and deletions^{41–43,46–49}, and recombination between such short sequence homologies is often RecA independent^{46,50–54}. In some cases, the join points of tandem duplications show no repetitive sequences at all^{15,41}. Possible RecA-independent processes that could contribute to duplication formation include strand slippage during DNA replication or repair⁵⁵, pairing of single-stranded regions of sister chromosome sequences at the replication fork⁵¹ and ligation of DNA ends through the exchange of DNA gyrase subunits bound at different DNA sites⁵⁶.

Mechanisms of GDA loss

As mentioned previously, tandem duplications or amplifications are intrinsically unstable (the loss rates are as high as 0.01–0.15 per cell per generation) because of homologous recombination between the identical repeated amplified units^{22,23} (FIG. 3a). Duplications and amplifications can often be stabilized by *recA* mutations^{25,37,39,57,58}, but in some plasmid systems deletion between repeated sequences also occurs when RecA is absent^{51,59,60}. Generally, if the tandem repeated sequences are short, deletions often occur in a RecA-independent manner^{59,61,62}. Duplications and amplifications can also be stabilized by selection. Therefore, even though the repeated sequences are recurrently lost by many cells in the population, selection will continuously enrich the population for cells that maintain the growth-promoting duplication or amplification^{17,18,34,35,48,63–68}.

Non-equal homologous recombination

The exchange of genetic material between repeated regions that are similar or identical, resulting in deletion or duplication and amplification.

Rolling circle replication

Nucleic acid replication that is common in plasmids and bacteriophages and allows rapid synthesis of multiple copies of a genome or specific region.

Box 2 | How to detect gene amplifications

Many gene amplifications are overlooked or never detected, because the novel genetic changes that they cause occur only at the duplication join point, which can be many kilobase pairs away from the resistance gene of interest. The general approach of PCR amplification of candidate genes followed by sequencing for the detection of resistance mutations will not detect such amplifications of resistance genes. The general instability of amplifications also makes them easily lost if selection is not continuously maintained. The classical method for the detection of gene duplications in *Escherichia coli* and *Salmonella enterica* subsp. *enterica* serovar Typhimurium is a transduction assay¹⁶. The principle is that the insertion of a marker into the duplicated gene will produce cells with both the gene⁺ phenotype and the phenotype of the marker (for example, Tn10 and Tet^R). The frequency of gene duplications in the population can then be calculated as the number of Tet^R gene⁺ colonies divided by the total number of Tet^R colonies (gene⁺ and gene⁻). Drawbacks to this method are that it is limited to bacteria for which genetic methods are available and that it is a gene-specific assay — that is, the gene that is assayed has to be known to give a detectable phenotype if it is knocked out. It also works best for the detection of large duplications (that is, larger than a transduced fragment).

Many recently developed methods can now be used to analyse the genomes of microorganisms, such as whole-genome DNA arrays and high-throughput sequencing (reviewed in REF. 110). This has increased the possibilities of more general detection of amplifications. Pulsed field gel electrophoresis (PFGE) has long been used to 'fingerprint' bacterial genomes. Gene duplications are detected as a change in banding pattern compared with single-copy genomes, depending on the position of the gene relative to the restriction site for the restriction enzyme used. However, small duplications can be hard to detect, and many other genome rearrangements will also change the PFGE profile. Real-time PCR (RT-PCR) is the best way to measure the copy number of a specific genomic region of interest, but it will only target that region¹¹¹. Next-generation sequencing (NGS) has provided the opportunity to rapidly sequence whole bacterial genomes at a reasonable cost¹¹². For detection of gene amplification, 454 sequencing technology offers the best approach, as the longer reads (currently, 450 bp) make the detection of join points easier. However, by mapping NGS reads from any platform (Solexa, 454 or Solid) to a reference genome sequence and comparing the sequence coverage (the number of reads covering a specific region), regions with a higher copy number than the surrounding regions can be detected. The sequence coverage can therefore be used to obtain a rough estimate of the copy number of a genomic region, but the result must be verified by other methods because fragmentation and PCR amplification of the DNA to be sequenced can introduce a bias in the sequence coverage for different parts of the genome. In DNA microarray experiments that target whole-genome DNA, an increase in gene copy number will be detected as a corresponding increase in the hybridization intensity of spots that cover the amplified region^{113,114}. This will, at the same time, give information about the size of the duplication and the location of the amplification join point, as well as the copy number.

When searching for gene amplifications, global genomics methods such as whole-genome sequencing and DNA arrays have the advantage of targeting the whole genome at once, and they will therefore give additional information. However, they are still expensive and are therefore not always an appropriate choice for screening experiments. PFGE, in combination with gene duplication and amplification verification through RT-PCR or transduction assays, is probably the most cost-efficient method and has potential to screen many samples while providing information on the whole-genome scale.

Antibiotic resistance and GDA

Plasmid-borne amplifications. The first examples of increased levels of antibiotic resistance resulting from increased gene copy number were found in the early 1970s, with the discovery that different R plasmids increased in size when grown in the presence of antibiotics. The first example was plasmid NR1 (also called R100 (REF. 69) and R222 (REF. 70)), which increased in size when plasmid-containing *Proteus mirabilis* was grown in the presence of a low concentration of the antibiotics to which the plasmid provided resistance (chloramphenicol, streptomycin and sulphonamide but not tetracycline)^{35,71}. The increase in size was reversed when cells were grown in the absence of drug. The larger plasmid size was due to amplification of the resistance determinant (r-determinant) region^{35,72} and led to high-level resistance to chloramphenicol, sulphonamide and streptomycin^{73,74}. It was assumed that the genes encoding resistance to these three antibiotics resided together in the r-determinant region, whereas the genes encoding tetracycline resistance did not. This turned out to be true, and the r-determinant region of NR1 was shown to be flanked by direct repeats of insertion sequence 1 (IS1) that led to amplification through homologous recombination between sister chromatids; by contrast, tetracycline resistance genes were inserted with

a Tn10 transposon, outside the r-determinant region, in the transfer region of the plasmid^{75,76}. Importantly, functions encoded on the IS1 elements are not actively involved in the actual amplification process, but instead the sequence identity of the two IS1 copies is used as a region of homology for the initial recombination event, leading to duplication of the region between the IS1 elements⁷⁷. Peterson and Rownd⁷⁸ later showed that the frequency of the initial duplication step through IS1 elements was 150-fold lower than the following amplification to an even higher copy number, and that this was related to the smaller size of IS1 compared with the whole duplicated region.

Clewell *et al.* reported the second observation of increased plasmid size during growth in the presence of an antibiotic⁶⁴. When *Enterococcus faecalis* cells (previously known as *Streptococcus faecalis*) bearing plasmid pAMA1 were grown in the presence of tetracycline for a prolonged period of time, larger forms of the plasmid were detected with a concomitant increase in the level of tetracycline resistance^{64,79}. This increase was due to tandem amplification of up to eight copies of the r-determinant region of the plasmid. The amplified state was not stable during growth in the absence of the drug and resulted in loss of the increased resistance and the return of the plasmid to its original size^{64,79}. The

R plasmid

A plasmid that carries genes that confer antibiotic resistance.

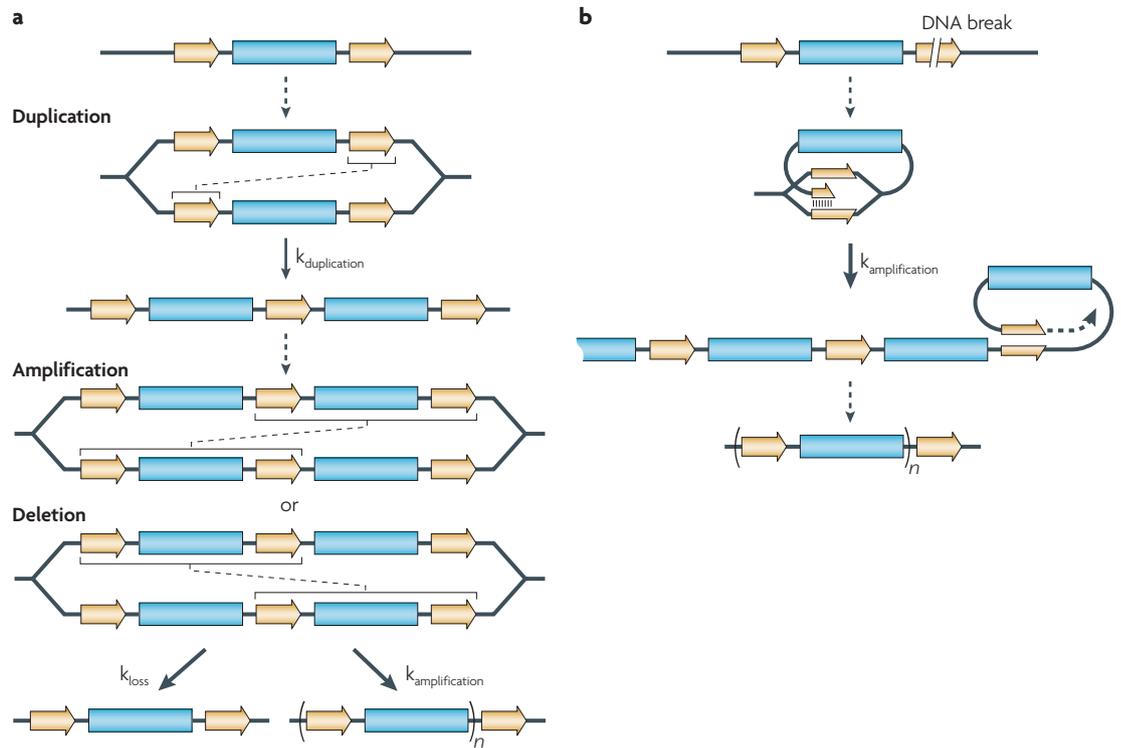


Figure 3 | Proposed mechanisms of gene duplication and amplification (GDA). **a** | Duplication occurs either through non-equal homologous recombination between directly orientated repeats or through RecA-independent mechanisms between short- or no-homology regions. **b** | GDA through rolling circle replication. A double-strand break allows single-strand invasion at a homologous or microhomologous site followed by replication. Horizontal arrows denote regions of homology (long or short) involved in recombination (dashed lines) between sister chromatids that forms the initial duplication. Horizontal square brackets indicate the size of the region of homology. Vertical arrows denote the rate of formation of the initial duplication ($k_{duplication}$), the rate of further amplification ($k_{amplification}$) and loss of amplification or duplication (k_{loss}). The thickness of vertical arrows indicates relative rates of formation and loss. The region in vertical brackets denotes the amplified unit.

amplified unit was 4.1 kb, which included the *tetL* gene and was flanked by direct repeats of 387 bp sequences that contained two versions of *oriT* (the origin for conjugal plasmid transfer), although with a low level (57% identity) of sequence similarity⁸⁰. Amplification was RecA dependent⁸¹, but the initial duplication step using the imperfect direct repeats depended on the activity of the plasmid-encoded proteins *MobB* and *MobE*, which nick and rejoin the two *oriT* regions, to form the initial duplication that was subsequently used for RecA-dependent amplification⁸⁰. Interestingly, it was later found that the pAMa1 plasmid consists of two separate replicons, and the amplified 4.1 kb region can, after intramolecular recombination in the direct repeat regions, form a self-replicating plasmid that is similar to free-standing plasmids that have been isolated from *Bacillus* species⁸².

Numerous other cases of amplification of plasmid-borne resistance genes, in a range of different bacteria, have now been described (TABLE 1), many of which have been detected in laboratory experiments by growing cells in increasing concentrations of antibiotic. In most cases, IS elements flanking the resistance determinants are involved in the initial duplication that is then followed by amplification. This

reflects the high frequency of multiple copies of identical IS elements on resistance plasmids — a feature that might be selectively stabilized by an increase in resistance through gene amplification.

Chromosomal gene amplifications. Normark *et al.* were the first to describe a chromosomal gene amplification that led to increased antibiotic resistance⁸³. Using an *E. coli* strain that contained a promoter mutation for high production of the *AmpC* β -lactamase, they selected for even higher levels of resistance by growing the cells in increasing concentrations of ampicillin. At low frequencies (10^{-8} per cell), cells with increased resistance were isolated in both *recA*⁺ and *recA*⁻ strains. However, *recA*⁻ mutations stabilized the amplifications and high-level resistance. The level of amplification was up to 30 copies, as determined by the direct correlation between β -lactam resistance and gene copy number. Later analysis of the amplifications showed that the duplication join points did not originate from any large regions of sequence homology but from different short (12–13 bp) direct repeats, and the initial duplications probably occurred through RecA-independent mechanisms^{47,48}.

Table 1 | Examples of gene amplification in antibiotic resistance development

Organism	Resistance	Gene (or genes)	Size of amplicon	Copy number	Amplification point	Refs
Plasmid-borne resistance amplification						
<i>Pseudomonas</i> sp. pB10	β -lactam	<i>bla</i> _{OXA-2}	880 bp	2	attC-integron insertion	36
<i>Enterococcus faecalis</i> pAM α 1	Tetracycline	<i>tetL</i>	4.1 kb	2–8	380 bp imperfect direct repeat	64,79–81,115
<i>Acinetobacter baumannii</i>	β -lactam	<i>bla</i> _{OXA-58}	6.8 kb	2–3	IS26 (duplicative transposition)	29
<i>Klebsiella pneumoniae</i>	β -lactam	<i>bla</i> _{SHV}	Unknown	2–8	IS26 (0.8 kb)	31
<i>Proteus mirabilis</i> NR1	Chloramphenicol, streptomycin and sulphonamide	<i>cat</i> , <i>aadA1</i> and <i>sul1</i>	2–21 kb	High	IS1 (0.77 kb)	35,71–73,116
<i>Haemophilus influenzae</i>	Tetracycline and chloramphenicol	<i>tet</i> and <i>cat</i>	12.4–13.8 kb	2–3	Presumably IS10	117
<i>Escherichia coli</i> pRSD1	Tetracycline	<i>tet</i>	5.3 kb	2–9	1.8 kb	66,118–120
<i>E. coli</i>	Chloramphenicol	<i>cat</i>	1.9 kb	2–11	IS1 (0.77 kb)	30
<i>E. coli</i>	β -lactam	Unknown	Unknown	2–30	Unknown	121
Chromosomally encoded resistance amplification						
<i>Streptococcus agalactiae</i>	Trimethoprim and sulfonamide	<i>folCEPBK</i>	13.5 kb	4	16 bp direct repeat	63
<i>E. faecalis</i>	Macrolide	<i>ermA</i>	72 bp of RBS	2	Unknown	84
<i>Staphylococcus aureus</i>	Macrolide	<i>ermA</i>	25 bp of RBS	2–3	Unknown	85
<i>Yersinia enterocolitica</i>	β -lactam	<i>blaA</i>	28 kb	2	Possible GyrA-mediated duplication	68
<i>E. coli</i>	β -lactam	<i>bla</i> _{ampC}	9–18 kb	2–40	12–13 bp direct repeat	47,48,83
<i>E. coli</i>	Sulfonamide	<i>sur</i>	18 kb	4	IS5 (1.19 kb)	32
<i>Streptococcus pneumoniae</i>	Macrolide	<i>rplV</i>	18 bp	2	No homology	86
<i>E. coli</i>	MDR efflux	<i>acrA</i> and <i>acrB</i>	149–300 kb	2–3	IS3, IS5 and IS186 (1.31 kb, 1.26 kb and 1.19 kb, respectively)	33,34
<i>S. aureus</i>	Methicillin	<i>mec</i>	~48 kb	2–3	Possibly IS257	67
<i>E. coli</i> P1 phage	Chloramphenicol and fusidic acid	<i>cat</i> and <i>fus</i>	2.6–4.8 kb	2–11	IS1 (0.77 kb)	122
<i>E. coli</i>	Mecillinam	<i>ftsQAZ</i>	80–300 kb	2	Probably IS1 mediated	123
<i>Streptomyces kanamyceticus</i>	Kanamycin	<i>kan</i> gene cluster	145 kb	2–36	CRISPR-like element homology	124
<i>Salmonella enterica</i> subsp. <i>enterica</i> serovar Typhimurium	Actinonin (GDA ameliorates fitness cost of resistance)	<i>metZ</i> and <i>metW</i>	1.9–94 kb	5–40	0 to 38 bp	15

GDA, gene duplication and amplification; IS, insertion sequence; MDR, multidrug resistance; RBS, ribosome-binding site.

Apart from the direct increase in antibiotic resistance that results from increased expression of a modifying or degrading enzyme or of efflux pumps, there are also examples in which duplications have mediated more unpredicted mechanisms to increase resistance. Duplication of part of the ribosome-binding site of the macrolide resistance gene *ermA* (erythromycin resistance methylase A) has been found in both *Staphylococcus aureus* (25 bp duplication) and *E. faecalis* (72 bp duplication). These duplications lead to constitutive expression of *ermA* by bypassing the translational attenuation system that

regulates *ermA* expression^{84,85}. An even more peculiar role of a duplication event leading to antibiotic resistance was presented by Musher *et al.*⁸⁶ During macrolide treatment of a *Streptococcus pneumoniae* infection, the bacterium developed resistance to the drug, leading to treatment failure and death of the patient. Resistance was due to an 18 bp duplication in the *rplV* gene, which encodes the ribosomal protein L22, that led to the duplication of 6 amino acids adjacent to the macrolide-binding site in the 23S ribosomal subunit. This duplication blocked binding of the drug, which resulted in resistance.

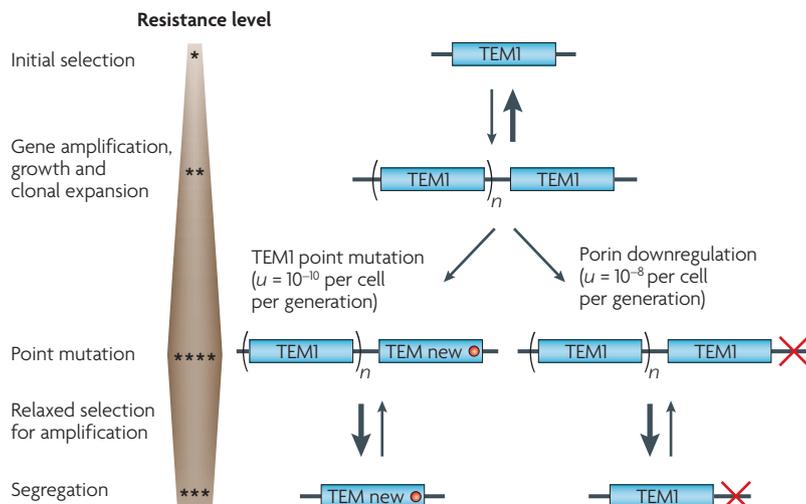


Figure 4 | Gene amplification as a facilitator of antibiotic resistance development. An initial duplication followed by further amplification of a gene that gives a small increase in tolerance against an antibiotic can be selected owing to increased gene dosage. The amplified gene array provides an increased target for mutations that increase resistance of the parental gene (left) or facilitate cell survival until another unrelated resistance mutation appears (right). If any of the later point mutations provide sufficient adaptive improvement the amplification can be lost. The thickness of the vertical arrows indicates relative rates of formation and loss. Loss-of-function mutations leading to porin downregulation are indicated by a red cross. A mutation conferring a new function in TEM1 is indicated by a red dot. u denotes mutation rates. The asterisks represent the level of resistance.

Clinical observations of resistance owing to GDA. Most characterized amplifications leading to increased resistance have been the result of laboratory set-ups that are designed to detect increased resistance, and few have been actual isolates found in clinical situations^{29,31,63,85,86}. The study by Musher *et al.*⁸⁶ that was described above is the only characterized clinical finding of a GDA that led to resistance during antibiotic treatment. Other instances of GDA in clinical isolates have all come from retrospective screens of random resistant isolates and were either found by differences in resistance levels^{29,85} or by DNA array screening⁶³. Brochet *et al.*⁶³ looked for GDA, and in 2 out of 75 screened clinical isolates of *Streptococcus agalactiae* they found higher probe ratios corresponding to a duplication of 92 kb, without any significant repeat regions at the join point, and a 13.5 kb quadruplicate repetition flanked by 16 bp imperfect direct repeats. The 13.5 kb region included the *folCEPBK* operon, which is involved in folate biosynthesis, and the amplification led to resistance against sulphonamides and trimethoprim.

Given the high resistance levels and the high frequencies of resistance through GDA found in *in vitro* studies, GDA would be expected to have a large impact on clinical resistance development. The low number of reported clinical cases probably reflects the transient nature of gene amplifications and the difficulty of detecting them both because of loss of amplification owing to growth with relaxed selection (usually the case when bacteria are isolated routinely in clinical laboratories) and the fact that gene amplification precedes more stable genetic changes, which, in some cases, also lead to higher levels

of resistance. It is likely that the clinical importance of GDA in resistance development is underestimated and that many more clinical examples would be found if studies were aimed specifically at finding them.

GDA and resistance development

In most cases in which antibiotic resistance is conferred by gene amplification, the increased resistance is linked directly to increased levels of the gene product of the amplified gene. However, owing to the high frequency with which amplifications occur in a population (FIG. 2), gene amplification can also be a first step towards increased resistance that is subsequently followed by rarer point mutations in the amplified gene (the likelihood of this occurring is increased by the increased number of target gene copies¹) or unrelated genes that further increase the resistance. The amplification is initially stabilized by the increased expression of the resistance gene but can be lost after the occurrence of a mutation that yields high-level resistance.

We have examined a system for the detection of amplification followed by mutational selection for increased β -lactam resistance⁸⁷ (FIG. 4). The TEM1 β -lactamase can hydrolyse both penicillins and early-generation cephalosporins and has trace activity towards later-generation cephalosporins⁸⁸. Evolution of resistance towards new β -lactam antibiotics has occurred mainly through mutation of existing β -lactamases, leading to an increased range of catalytic activities for these enzymes^{89,90}. As resistance to β -lactam antibiotics is linearly correlated with the level of β -lactamase⁹¹, β -lactams are a good experimental system for amplification studies. By challenging *S. Typhimurium* that carries the *bla*_{TEM1} gene on an F'-plasmid with progressively increasing levels of cephalothin (a first-generation cephalosporin) or cefaclor (a second-generation cephalosporin), mutants with increased cephalosporin resistance were rapidly selected. Most (70%) clones with increased resistance had increased *bla*_{TEM1} gene copy numbers (up to a 40-fold increase). Continued serial passage of lineages with amplifications led in several cases to a reduced copy number of *bla*_{TEM1} but maintained or increased levels of cephalosporin resistance. In several of these lineages, mutational loss of the expression of the porins *OmpC*, *OmpD* and *OmpF* caused the increased resistance by reducing antibiotic influx into the cell. This experiment shows how amplification can not only generate increased resistance as a primary response to antibiotics, but can also facilitate the development of stable increased resistance through subsequent point mutations by allowing cell growth and population expansion. Even if it was not detected in this experiment, point mutation of the *bla*_{TEM1} gene is another plausible outcome that could have yielded higher levels of cephalosporin resistance.

Other researchers have used a different approach⁹² to analyse the possibility that two gene copies of *bla*_{TEM1} on two separate plasmids can lead to the evolution of novel activity towards cefotaxime (a third-generation cephalosporin) while one gene copy with the ancestral catalytic activity towards ampicillin is retained. In this study, no control of gene copy number was performed,

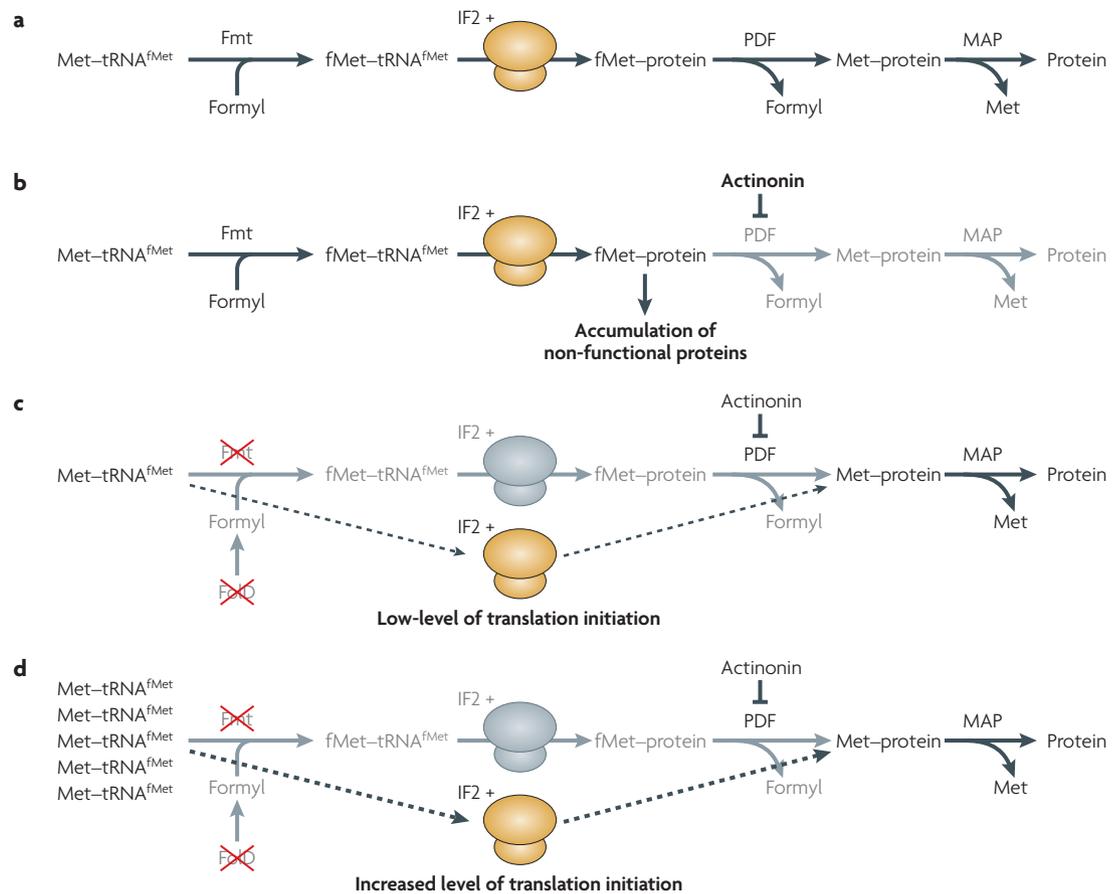


Figure 5 | Amplification of tRNA^{fMet} ameliorates the cost of actinonin resistance mutations. **a** | Pathway for translational initiation in *Salmonella enterica* subsp. *enterica* serovar Typhimurium. **b** | Inhibition of peptide deformylase activity by actinonin leads to the accumulation of non-functional proteins and cell death. **c** | Resistance mutations in *fmt* or *folD* bypass the deformylation deficiency by abolishing formylation. **d** | Amplification of *metZ* and *metW*, the genes that encode initiation tRNA^{fMet}, leads to an increased cellular concentration of Met-tRNA^{fMet} and increases initiation without formylation. fMet, formyl-methionine; Fmt, formyl transferase; IF, initiation factor; MAP, methionine aminopeptidase; PDF, peptide deformylase.

other than maintenance of both plasmids. They found that mutants with activity against cefotaxime readily evolved, but the two plasmids could not coexist even in the presence of both antibiotics. For unknown reasons, the mutated *bla*_{TEM1} plasmid together with the other *bla*_{TEM1} plasmid posed a high fitness cost to the bacterium.

Reducing fitness costs of resistance

Resistance mutations frequently show a general fitness cost that manifests as a reduced growth rate and/or virulence in the absence of the drug. During growth in the absence of the drug, mutations often arise that compensate for the fitness cost of the resistance mutation, leading to strains with restored fitness and the maintenance of drug resistance⁹³⁻⁹⁵. In one recently published case, compensation occurred through the amplification of a wild-type gene that was unrelated to the original resistance mutation (FIG. 5). The compensatory amplification was observed in *fmt* (which encodes formyl transferase) mutants that are resistant to deformylase inhibitors¹⁵. This

class of drugs inhibits bacterial growth by preventing the removal of the formyl group from formyl-methionine (fMet), which is incorporated as the first amino acid in most proteins. As removal of the formyl group is essential to the function of many proteins, this results in cell death (FIG. 5b). Cells become resistant to the inhibitor through *fmt* mutations that eliminate formylation of methionyl-tRNA^{fMet} (Met-tRNA^{fMet}; charged initiator tRNA), thereby bypassing the need for removal of the formyl group (FIG. 5c). This resistance mutation confers a large fitness cost owing to the impairment of the initiation of protein synthesis in the absence of the formyl group. This fitness cost can be partly compensated, without loss of drug resistance, when tRNA^{fMet} is overproduced and allows initiation of protein synthesis to occur using Met-tRNA^{fMet} rather than formylated Met-tRNA^{fMet} (FIG. 5d). Overproduction of tRNA^{fMet} is achieved by amplification of two tandem genes, *metZ* and *metW*, both of which encode normal tRNA^{fMet}. This amplification restores nearly normal fitness to strains with the resistance mutation.

Evolutionary and clinical consequences

With regard to evolution in general, and the evolution of antibiotic resistance in particular, the high prevalence and intrinsic instability of most GDAs have several important implications. First, the high steady-state frequency of GDA of most genes in a bacterial population suggests that selection for cells with gene amplification rather than mutation is the initial response to selective conditions that are resolvable by increased gene activity⁹⁶. Therefore, if growth is limited by an external factor (for example, an antibiotic) or an internal factor (for example, a deleterious resistance mutation), and the problem can be mitigated by increasing the level of a gene product, common GDAs are more likely to be selected than rare point mutations. Often it is the gene that confers the limitation that is amplified (for example, *bla* genes in the examples presented above), but sometimes increased dosage of an unrelated non-cognate gene can resolve the problem (for example, *metZ* and *metW* amplification in the *fmt* mutant in the example discussed above). As shown by a recent study in *E. coli*, non-cognate suppression is frequently observed, suggesting that promiscuity and multifunctionality are common properties of bacterial genes⁹⁷. In this paper, the authors demonstrated that, out of 104 single-gene knockout strains, 20% of these auxotrophs could be rescued by the overexpression of at least 1 non-cognate gene. As the deleted gene and the suppressor were generally genetically unrelated, this finding suggests that promiscuity was a consequence of contingency.

Second, GDA could provide an initial solution to a selective problem and thereby allow a population to expand to a sufficiently large size for it to generate a subsequent stable mutational solution. Therefore, GDA could serve as an intermediate and facilitator in the development of a stable genetic adaptation (as exemplified by the *bla*_{TEM1} case described above). Such stable changes could occur outside the amplified region and be allowed by population expansion alone, or they could affect the amplified genes, in which case the target number is increased by both cell growth and the additional gene copies per cell.

Finally, as the parent amplification is often unstable and may be dispensable after a stable adaptation occurs, it is conceivable that the intermediate amplifications leave no sequence trace in the genome. Therefore, the intermediate role of amplification can escape detection when stable intermediates have been established, because the amplification is rapidly lost after selection is relaxed. This could be relevant in clinical settings, and it is possible that the contribution of GDA to antibiotic resistance, antibiotic treatment failures and host adaptation is missed because the GDA has already been lost when the clinical isolates are characterized genetically. Therefore, when searching for GDA in clinical isolates, it might be worthwhile to examine the strains immediately after isolation from the patients without any preceding non-selective growth in the laboratory. Here, whole-genome sequencing methods that rely on small amounts of DNA are of interest and should allow 'in situ-like' analysis of GDA in a bacterial population directly recovered from a patient⁹⁸.

1. Bergthorsson, U., Andersson, D. I. & Roth, J. R. Ohno's dilemma: evolution of new genes under continuous selection. *Proc. Natl Acad. Sci. USA* **104**, 17004–17009 (2007).
Discusses the development of a new model (innovation, amplification and divergence) that can explain the evolution of new genes under continuous selection.
2. Ohno, S. *Evolution by Gene Duplication* (Springer, Berlin, 1970).
3. Romero, D. & Palacios, R. Gene amplification and genomic plasticity in prokaryotes. *Annu. Rev. Genet.* **31**, 91–111 (1997).
4. Beckmann, J. S., Estivill, X. & Antonarakis, S. E. Copy number variants and genetic traits: closer to the resolution of phenotypic to genotypic variability. *Nature Rev. Genet.* **8**, 639–646 (2007).
5. Conrad, B. & Antonarakis, S. E. Gene duplication: a drive for phenotypic diversity and cause of human disease. *Annu. Rev. Genomics Hum. Genet.* **8**, 17–35 (2007).
6. Harms, C. T. *et al.* Herbicide resistance due to amplification of a mutant acetohydroxyacid synthase gene. *Mol. Gen. Genet.* **233**, 427–435 (1992).
7. Papanikou, E., Brotherton, J. E. & Widholm, J. M. Length of time in tissue culture can affect the selected glyphosate resistance mechanism. *Planta* **218**, 589–598 (2004).
8. Suh, H., Hepburn, A. G., Kriz, A. L. & Widholm, J. M. Structure of the amplified 5-enolpyruvylshikimate-3-phosphate synthase gene in glyphosate-resistant carrot cells. *Plant Mol. Biol.* **22**, 195–205 (1993).
9. Hemingway, J. The molecular basis of two contrasting metabolic mechanisms of insecticide resistance. *Insect Biochem. Mol. Biol.* **30**, 1009–1015 (2000).
10. Raymond, M., Chevillon, C., Guillemaud, T., Lenormand, T. & Pasteur, N. An overview of the evolution of overproduced esterases in the mosquito *Culex pipiens*. *Philos. Trans. R. Soc. Lond. B* **353**, 1707–1711 (1998).
11. Albertson, D. G. Gene amplification in cancer. *Trends Genet.* **22**, 447–455 (2006).
12. Fojo, T. Multiple paths to a drug resistance phenotype: mutations, translocations, deletions and amplification of coding genes or promoter regions, epigenetic changes and microRNAs. *Drug Resist. Updat.* **10**, 59–67 (2007).
13. Harbottle, H., Thakur, S., Zhao, S. & White, D. G. Genetics of antimicrobial resistance. *Anim. Biotechnol.* **17**, 111–124 (2006).
14. Normark, B. H. & Normark, S. Evolution and spread of antibiotic resistance. *J. Intern. Med.* **252**, 91–106 (2002).
15. Nilsson, A. I. *et al.* Reducing the fitness cost of antibiotic resistance by amplification of initiator tRNA genes. *Proc. Natl Acad. Sci. USA* **103**, 6976–6981 (2006).
The first demonstration that the fitness costs associated with antibiotic resistance can be ameliorated by gene amplification.
16. Anderson, P. & Roth, J. Spontaneous tandem genetic duplications in *Salmonella typhimurium* arise by unequal recombination between rRNA (*rrn*) cistrons. *Proc. Natl Acad. Sci. USA* **78**, 3113–3117 (1981).
The first genome-wide analysis of duplication frequencies in a bacterial chromosome.
17. Andersson, D. I., Slechta, E. S. & Roth, J. R. Evidence that gene amplification underlies adaptive mutability of the bacterial *lac* operon. *Science* **282**, 1133–1135 (1998).
The first demonstration that adaptive mutation in the lac operon can be explained by gene amplification of the selected lac target gene.
18. Kugelberg, E., Kofoid, E., Reams, A. B., Andersson, D. I. & Roth, J. R. Multiple pathways of selected gene amplification during adaptive mutation. *Proc. Natl Acad. Sci. USA* **103**, 17319–17324 (2006).
19. Sonti, R. V. & Roth, J. R. Role of gene duplications in the adaptation of *Salmonella typhimurium* to growth on limiting carbon sources. *Genetics* **123**, 19–28 (1989).
20. Straus, D. S. Selection for a large genetic duplication in *Salmonella typhimurium*. *Genetics* **80**, 227–237 (1975).
21. Straus, D. S. & D'Ari Straus, L. Large overlapping tandem genetic duplications in *Salmonella typhimurium*. *J. Mol. Biol.* **103**, 143–153 (1976).
22. Roth, J. R. *et al.* In *Escherichia coli and Salmonella: Cellular and Molecular Biology* 2nd edn (eds Neidhardt, F. C. *et al.*) 2256–2276 (American Society for Microbiology, Washington DC, 1996).
23. Pettersson, M. E., Sun, S., Andersson, D. I. & Berg, O. G. Evolution of new gene functions: simulation and analysis of the amplification model. *Genetica* **135**, 309–324 (2008).
24. Veitia, R. A. Gene dosage balance in cellular pathways: implications for dominance and gene duplicability. *Genetics* **168**, 569–574 (2004).
25. Galitski, T. & Roth, J. R. Pathways for homologous recombination between chromosomal direct repeats in *Salmonella typhimurium*. *Genetics* **146**, 751–767 (1997).
26. Hastings, P. J., Ira, G. & Lupski, J. R. A microhomology-mediated break-induced replication model for the origin of human copy number variation. *PLoS Genet.* **5**, e1000327 (2009).
27. Petit, M. A., Mesas, J. M., Noirot, P., Morel-Deville, F. & Ehrlich, S. D. Induction of DNA amplification in the *Bacillus subtilis* chromosome. *EMBO J.* **11**, 1317–1326 (1992).
28. Lehner, A. F. & Hill, C. W. Involvement of ribosomal ribonucleic acid operons in *Salmonella typhimurium* chromosomal rearrangements. *J. Bacteriol.* **143**, 492–498 (1980).

29. Bertini, A. *et al.* Multicopy *bla*_{OXA58} gene as a source of high-level resistance to carbapenems in *Acinetobacter baumannii*. *Antimicrob. Agents Chemother.* **51**, 2324–2328 (2007).
30. Chandler, M., de la Tour, E. B., Willems, D. & Caro, L. Some properties of the chloramphenicol resistance transposon Tn9. *Mol. Gen. Genet.* **176**, 221–231 (1979).
31. Hammond, D. S., Schooneveldt, J. M., Nimmo, G. R., Huygens, F. & Giffard, P. M. *bla*_{SHV} genes in *Klebsiella pneumoniae*: different allele distributions are associated with different promoters within individual isolates. *Antimicrob. Agents Chemother.* **49**, 256–263 (2005).
32. Nichols, B. P. & Guay, G. G. Gene amplification contributes to sulfonamide resistance in *Escherichia coli*. *Antimicrob. Agents Chemother.* **33**, 2042–2048 (1989).
33. Nicoloff, H., Perreten, V. & Levy, S. B. Increased genome instability in *Escherichia coli* lon mutants: relation to emergence of multiple-antibiotic-resistant (*Mar*) mutants caused by insertion sequence elements and large tandem genomic amplifications. *Antimicrob. Agents Chemother.* **51**, 1293–1303 (2007).
34. Nicoloff, H., Perreten, V., McMurry, L. M. & Levy, S. B. Role for tandem duplication and Lon protease in AcrAB–TolC-dependent multiple antibiotic resistance (*Mar*) in an *Escherichia coli* mutant without mutations in *marRAB* or *acrRAB*. *J. Bacteriol.* **188**, 4413–4423 (2006).
35. Rownd, R. & Mickel, S. Dissociation and reassociation of RTF and *r*-determinants of the R-factor NR1 in *Proteus mirabilis*. *Nature New Biol.* **234**, 40–43 (1971).
36. Szczepanowski, R., Krahn, I., Puhler, A. & Schluter, A. Different molecular rearrangements in the integron of the IncP-1β resistance plasmid pB10 isolated from a wastewater treatment plant result in elevated β-lactam resistance levels. *Arch. Microbiol.* **182**, 429–435 (2004).
37. Haack, K. R. & Roth, J. R. Recombination between chromosomal IS200 elements supports frequent duplication formation in *Salmonella typhimurium*. *Genetics* **141**, 1245–1252 (1995).
38. Jessop, A. P. & Clugston, C. Amplification of the ArgF region in strain HfrP4X of *E. coli* K-12. *Mol. Gen. Genet.* **201**, 347–350 (1985).
39. Lin, R. J., Capage, M. & Hill, C. W. A repetitive DNA sequence, *rhs*, responsible for duplications within the *Escherichia coli* K-12 chromosome. *J. Mol. Biol.* **177**, 1–18 (1984).
40. Shyamala, V., Schneider, E. & Ames, G. F. Tandem chromosomal duplications: role of REP sequences in the recombination event at the join-point. *EMBO J.* **9**, 939–946 (1990).
41. Reams, A. B. & Neidle, E. L. Gene amplification involves site-specific short homology-independent illegitimate recombination in *Acinetobacter* sp. strain ADP1. *J. Mol. Biol.* **338**, 643–656 (2004).
- A thorough analysis of the sequence requirements for gene amplification.**
42. Tlsty, T. D., Albertini, A. M. & Miller, J. H. Gene amplification in the *lac* region of *E. coli*. *Cell* **37**, 217–224 (1984).
- A key paper showing that *lac* gene amplification is the predominantly selected response when cells are starved for lactose.**
43. Whoriskey, S. K., Nghiem, V. H., Leong, P. M., Masson, J. M. & Miller, J. H. Genetic rearrangements and gene amplification in *Escherichia coli*: DNA sequences at the junctures of amplified gene fusions. *Genes Dev.* **1**, 227–237 (1987).
44. Shen, P. & Huang, H. V. Homologous recombination in *Escherichia coli*: dependence on substrate length and homology. *Genetics* **112**, 441–457 (1986).
45. Watt, V. M., Ingles, C. J., Urdea, M. S. & Rutter, W. J. Homology requirements for recombination in *Escherichia coli*. *Proc. Natl Acad. Sci. USA* **82**, 4768–4772 (1985).
46. Albertini, A. M., Hofer, M., Calos, M. P. & Miller, J. H. On the formation of spontaneous deletions: the importance of short sequence homologies in the generation of large deletions. *Cell* **29**, 319–328 (1982).
47. Edlund, T., Grundstrom, T. & Normark, S. Isolation and characterization of DNA repetitions carrying the chromosomal β-lactamase gene of *Escherichia coli* K-12. *Mol. Gen. Genet.* **173**, 115–125 (1979).
48. Edlund, T. & Normark, S. Recombination between short DNA homologies causes tandem duplication.
- The first description of a chromosomal gene amplification that led to increased antibiotic resistance.**
49. Albertini, A. M., Hofer, M., Calos, M. P., Tlsty, T. D. & Miller, J. H. Analysis of spontaneous deletions and gene amplification in the *lac* region of *Escherichia coli*. *Cold Spring Harb. Symp. Quant. Biol.* **47**, 841–850 (1983).
50. Jones, I. M., Primrose, S. B. & Ehrlich, S. D. Recombination between short direct repeats in a *recA* host. *Mol. Gen. Genet.* **188**, 486–489 (1982).
51. Lovett, S. T., Drapkin, P. T., Suter, V. A. Jr & Gluckman-Peskind, T. J. A sister-strand exchange mechanism for *recA*-independent deletion of repeated DNA sequences in *Escherichia coli*. *Genetics* **135**, 631–642 (1993).
52. Lovett, S. T., Gluckman, T. J., Simon, P. J., Suter, V. A. Jr & Drapkin, P. T. Recombination between repeats in *Escherichia coli* by a *recA*-independent, proximity-sensitive mechanism. *Mol. Gen. Genet.* **245**, 294–300 (1994).
53. Marvo, S. L., King, S. R. & Jaskunas, S. R. Role of short regions of homology in intermolecular illegitimate recombination events. *Proc. Natl Acad. Sci. USA* **80**, 2452–2456 (1983).
54. Mazin, A. V., Kuzminov, A. V., Dianov, G. L. & Salganik, R. I. Mechanisms of deletion formation in *Escherichia coli* plasmids. II. Deletions mediated by short direct repeats. *Mol. Gen. Genet.* **228**, 209–214 (1991).
55. Trinh, T. Q. & Sinden, R. R. The influence of primary and secondary DNA structure in deletion and duplication between direct repeats in *Escherichia coli*. *Genetics* **134**, 409–422 (1993).
56. Ikeda, H., Shiraishi, K. & Ogata, Y. Illegitimate recombination mediated by double-strand break and end-joining in *Escherichia coli*. *Adv. Biophys.* **38**, 3–20 (2004).
57. Anderson, R. P. & Roth, J. R. Tandem chromosomal duplications in *Salmonella typhimurium*: fusion of histidine genes to novel promoters. *J. Mol. Biol.* **119**, 147–166 (1978).
58. Hill, C. W., Foulds, J., Soll, L. & Berg, P. Instability of a missense suppressor resulting from a duplication of genetic material. *J. Mol. Biol.* **39**, 563–581 (1969).
59. Bi, X. & Liu, L. F. *recA*-independent and *recA*-dependent intramolecular plasmid recombination. Differential homology requirement and distance effect. *J. Mol. Biol.* **235**, 414–423 (1994).
60. Matfield, M., Badawi, R. & Brammar, W. J. *Rec*-dependent and *Rec*-independent recombination of plasmid-borne duplications in *Escherichia coli* K12. *Mol. Gen. Genet.* **199**, 518–523 (1985).
61. Bzymek, M. & Lovett, S. T. Instability of repetitive DNA sequences: the role of replication in multiple mechanisms. *Proc. Natl Acad. Sci. USA* **98**, 8319–8325 (2001).
62. Puopolo, K. M., Hollingshead, S. K., Carey, V. J. & Madoff, L. C. Tandem repeat deletion in the alpha C protein of group B *Streptococcus* is *recA* independent. *Infect. Immun.* **69**, 5037–5045 (2001).
63. Brochet, M., Couve, E., Zouine, M., Poyart, C. & Glaser, P. A naturally occurring gene amplification leading to sulfonamide and trimethoprim resistance in *Streptococcus agalactiae*. *J. Bacteriol.* **190**, 672–680 (2008).
64. Clewell, D. B., Yagi, Y. & Bauer, B. Plasmid-determined tetracycline resistance in *Streptococcus faecalis*: evidence for gene amplification during growth in presence of tetracycline. *Proc. Natl Acad. Sci. USA* **72**, 1720–1724 (1975).
65. Kroll, J. S., Loynds, B. M. & Moxon, E. R. The *Haemophilus influenzae* capsulation gene cluster: a compound transposon. *Mol. Microbiol.* **5**, 1549–1560 (1991).
66. Mattes, R., Burkardt, H. J. & Schmitt, R. Repetition of tetracycline resistance determinant genes on R plasmid pRSD1 in *Escherichia coli*. *Mol. Gen. Genet.* **168**, 173–184 (1979).
67. Matthews, P. R. & Stewart, P. R. Amplification of a section of chromosomal DNA in methicillin-resistant *Staphylococcus aureus* following growth in high concentrations of methicillin. *J. Gen. Microbiol.* **134**, 1455–1464 (1988).
68. Seoane, A., Sanchez, E. & Garcia-Lobo, J. M. Tandem amplification of a 28-kilobase region from the *Yersinia enterocolitica* chromosome containing the *blaA* gene. *Antimicrob. Agents Chemother.* **47**, 682–688 (2003).
69. Sugino, Y. & Hirota, Y. Conjugal fertility associated with resistance factor R in *Escherichia coli*. *J. Bacteriol.* **84**, 902–910 (1962).
70. Falkow, S., Citarella, R. V. & Wohlhieter, J. A. The molecular nature of R-factors. *J. Mol. Biol.* **17**, 102–116 (1966).
71. Rownd, R., Kasamatsu, H. & Mickel, S. The molecular nature and replication of drug resistance factors of the Enterobacteriaceae. *Ann. NY Acad. Sci.* **182**, 188–206 (1971).
72. Perlman, D. & Rownd, R. H. Transition of R factor NR1 in *Proteus mirabilis*: molecular structure and replication of NR1 deoxyribonucleic acid. *J. Bacteriol.* **123**, 1013–1034 (1975).
73. Hashimoto, H. & Rownd, R. H. Transition of the R factor NR1 and *Proteus mirabilis*: level of drug resistance of nontransitioned and transitioned cells. *J. Bacteriol.* **123**, 56–68 (1975).
74. Huffman, G. A. & Rownd, R. H. Transition of deletion mutants of the composite resistance plasmid NR1 in *Escherichia coli* and *Salmonella typhimurium*. *J. Bacteriol.* **159**, 488–498 (1984).
75. Hu, S., Ohtsubo, E. & Davidson, N. Electron microscopic heteroduplex studies of sequence relations among plasmids of *Escherichia coli*: structure of F13 and related F-primes. *J. Bacteriol.* **122**, 749–763 (1975).
76. Ptashne, K. & Cohen, S. N. Occurrence of insertion sequence (IS) regions on plasmid deoxyribonucleic acid as direct and inverted nucleotide sequence duplications. *J. Bacteriol.* **122**, 776–781 (1975).
77. Peterson, B. C. & Rownd, R. H. Homologous sequences other than insertion elements can serve as recombination sites in plasmid drug resistance gene amplification. *J. Bacteriol.* **156**, 177–185 (1983).
78. Peterson, B. C. & Rownd, R. H. Drug resistance gene amplification of plasmid NR1 derivatives with various amounts of resistance determinant DNA. *J. Bacteriol.* **161**, 1042–1048 (1985).
79. Yagi, Y. & Clewell, D. B. Plasmid-determined tetracycline resistance in *Streptococcus faecalis*: tandemly repeated resistance determinants in amplified forms of pAMα1 DNA. *J. Mol. Biol.* **102**, 583–600 (1976).
80. Francia, M. V. & Clewell, D. B. Amplification of the tetracycline resistance determinant of pAMα1 in *Enterococcus faecalis* requires a site-specific recombination event involving relaxase. *J. Bacteriol.* **184**, 5187–5193 (2002).
81. Yagi, Y. & Clewell, D. B. Amplification of the tetracycline resistance determinant of plasmid pAMα1 in *Streptococcus faecalis*: dependence on host recombination machinery. *J. Bacteriol.* **143**, 1070–1072 (1980).
82. Perkins, J. B. & Youngman, P. *Streptococcus* plasmid pAMα1 is a composite of two separable replicons, one of which is closely related to *Bacillus* plasmid pBC16. *J. Bacteriol.* **155**, 607–615 (1983).
83. Normark, S., Edlund, T., Grundstrom, T., Bergstrom, S. & Wolf-Watz, H. *Escherichia coli* K-12 mutants hyperproducing chromosomal β-lactamase by gene repetitions. *J. Bacteriol.* **132**, 912–922 (1977).
84. Min, Y. H. *et al.* Molecular analysis of constitutive mutations in *ermB* and *ermA* selected *in vitro* from inducibly MLSB-resistant enterococci. *Arch. Pharm. Res.* **31**, 377–380 (2008).
85. Yoon, E. J., Kwon, A. R., Min, Y. H. & Choi, E. C. Foggy D-shaped zone of inhibition in *Staphylococcus aureus* owing to a dual character of both inducible and constitutive resistance to macrolide-lincosamide-streptogramin B. *J. Antimicrob. Chemother.* **61**, 533–540 (2008).
86. Musher, D. M. *et al.* Emergence of macrolide resistance during treatment of pneumococcal pneumonia. *N. Engl. J. Med.* **346**, 630–631 (2002).
- A clinically relevant observation of how gene amplification of a ribosome-binding site in the *ermA* gene can cause increased expression of a resistance gene and result in treatment failure.**
87. Sun, S., Berg, O. G., Roth, J. R. & Andersson, D. I. Contribution of gene amplification to evolution of increased antibiotic resistance in *Salmonella typhimurium*. *Genetics* 1 Jun 2009 (doi:10.1534/genetics.109.103028).
88. Matagne, A. *et al.* The diversity of the catalytic properties of class A β-lactamases. *Biochem. J.* **265**, 131–146 (1990).
89. Livermore, D. M. β-lactamases in laboratory and clinical resistance. *Clin. Microbiol. Rev.* **8**, 557–584 (1995).
90. Petrosino, J., Cantu, C. 3rd & Palzkill, T. β-lactamases: protein evolution in real time. *Trends Microbiol.* **6**, 323–327 (1998).

91. Nordstrom, K., Ingram, L. C. & Lundback, A. Mutations in R factors of *Escherichia coli* causing an increased number of R-factor copies per chromosome. *J. Bacteriol.* **110**, 562–569 (1972).
92. Holloway, A. K., Palzkill, T. & Bull, J. J. Experimental evolution of gene duplicates in a bacterial plasmid model. *J. Mol. Evol.* **64**, 215–222 (2007).
93. Andersson, D. I. Persistence of antibiotic resistant bacteria. *Curr. Opin. Microbiol.* **6**, 452–456 (2003).
94. Andersson, D. I. The biological cost of mutational antibiotic resistance: any practical conclusions? *Curr. Opin. Microbiol.* **9**, 461–465 (2006).
95. Andersson, D. I. & Levin, B. R. The biological cost of antibiotic resistance. *Curr. Opin. Microbiol.* **2**, 489–493 (1999).
96. Roth, J. R., Kugelberg, E., Reams, A. B., Kofoid, E. & Andersson, D. I. Origin of mutations under selection: the adaptive mutation controversy. *Annu. Rev. Microbiol.* **60**, 477–501 (2006).
97. Patrick, W. M., Quandt, E. M., Swartzlander, D. B. & Matsumura, I. Multicopy suppression underpins metabolic evolvability. *Mol. Biol. Evol.* **24**, 2716–2722 (2007).
- A landmark paper showing that promiscuity and multifunctionality are common properties of bacterial genes.**
98. Ishoey, T., Woyke, T., Stepanauskas, R., Novotny, M. & Lasken, R. S. Genomic sequencing of single microbial cells from environmental samples. *Curr. Opin. Microbiol.* **11**, 198–204 (2008).
99. Marichal, P. *et al.* Molecular biological characterization of an azole-resistant *Candida glabrata* isolate. *Antimicrob. Agents Chemother.* **41**, 2229–2237 (1997).
100. Cowman, A. F., Galatis, D. & Thompson, J. K. Selection for mefloquine resistance in *Plasmodium falciparum* is linked to amplification of the *pfmdr1* gene and cross-resistance to halofantrine and quinine. *Proc. Natl Acad. Sci. USA* **91**, 1143–1147 (1994).
101. Preechapornkul, P. *et al.* *Plasmodium falciparum* *pfmdr1* amplification, mefloquine resistance, and parasite fitness. *Antimicrob. Agents Chemother.* **53**, 1509–1515 (2009).
102. Suwanarusk, R. *et al.* Amplification of *pvmr1* associated with multidrug-resistant *Plasmodium vivax*. *J. Infect. Dis.* **198**, 1558–1564 (2008).
103. Gonzalez-Pons, M., Szeto, A. C., Gonzalez-Mendez, R. & Serrano, A. E. Identification and bioinformatic characterization of a multidrug resistance associated protein (ABCC) gene in *Plasmodium berghei*. *Malaria J.* **8**, 1 (2009).
104. Reed, M. B., Saliba, K. J., Caruana, S. R., Kirk, K. & Cowman, A. F. Pgh1 modulates sensitivity and resistance to multiple antimalarials in *Plasmodium falciparum*. *Nature* **403**, 906–909 (2000).
105. Sharom, F. J. ABC multidrug transporters: structure, function and role in chemoresistance. *Pharmacogenomics* **9**, 105–127 (2008).
106. Yuan, H. *et al.* Strategies to overcome or circumvent P-glycoprotein mediated multidrug resistance. *Curr. Med. Chem.* **15**, 470–476 (2008).
107. Nair, S. *et al.* Adaptive copy number evolution in malaria parasites. *PLoS Genet.* **4**, e1000243 (2008).
108. Leprohon, P. *et al.* Gene expression modulation is associated with gene amplification, supernumerary chromosomes and chromosome loss in antimony-resistant *Leishmania infantum*. *Nucleic Acids Res.* **37**, 1387–1399 (2009).
109. Lin, Y. C. *et al.* Two distinct arsenite-resistant variants of *Leishmania amazonensis* take different routes to achieve resistance as revealed by comparative transcriptomics. *Mol. Biochem. Parasitol.* **162**, 16–31 (2008).
110. Medini, D. *et al.* Microbiology in the post-genomic era. *Nature Rev. Microbiol.* **6**, 419–430 (2008).
111. Lee, C., Lee, S., Shin, S. G. & Hwang, S. Real-time PCR determination of rRNA gene copy number: absolute and relative quantification assays with *Escherichia coli*. *Appl. Microbiol. Biotechnol.* **78**, 371–376 (2008).
112. Mardis, E. R. The impact of next-generation sequencing technology on genetics. *Trends Genet.* **24**, 133–141 (2008).
113. Porwollik, S. *et al.* DNA amplification and rearrangements in archival *Salmonella enterica* serovar Typhimurium LT2 cultures. *J. Bacteriol.* **186**, 1678–1682 (2004).
114. Shiu, S. H. & Borevitz, J. O. The next generation of microarray research: applications in evolutionary and ecological genomics. *Heredity* **100**, 141–149 (2008).
115. Yagi, Y. & Clewell, D. B. Identification and characterization of a small sequence located at two sites on the amplifiable tetracycline resistance plasmid pAMα1 in *Streptococcus faecalis*. *J. Bacteriol.* **129**, 400–406 (1977).
116. Tanak, N., Cramer, J. H. & Rownd, R. H. EcoRI restriction endonuclease map of the composite R plasmid NR1. *J. Bacteriol.* **127**, 619–636 (1976).
117. Spies, T., Laufs, R. & Riess, F. C. Amplification of resistance genes in *Haemophilus influenzae* plasmids. *J. Bacteriol.* **155**, 839–846 (1983).
118. Schmitt, R., Bernhard, E. & Mattes, R. Characterisation of Tn1721, a new transposon containing tetracycline resistance genes capable of amplification. *Mol. Gen. Genet.* **172**, 53–65 (1979).
119. Schoffl, F., Arnold, W., Puhler, A., Altenbuchner, J. & Schmitt, R. The tetracycline resistance transposons Tn1721 and Tn1771 have three 38-base-pair repeats and generate five-base-pair direct repeats. *Mol. Gen. Genet.* **181**, 87–94 (1981).
120. Wiebauer, K., Schraml, S., Shales, S. W. & Schmitt, R. Tetracycline resistance transposon Tn1721: *recA*-dependent gene amplification and expression of tetracycline resistance. *J. Bacteriol.* **147**, 851–859 (1981).
121. Odakura, Y., Tanaka, T., Hashimoto, H. & Mitsuhashi, S. Mutation of R factors capable of specifying hypersynthesis of penicillinase. *Antimicrob. Agents Chemother.* **3**, 315–324 (1973).
122. Meyer, J. & Iida, S. Amplification of chloramphenicol resistance transposons carried by phage P1Cm in *Escherichia coli*. *Mol. Gen. Genet.* **176**, 209–219 (1979).
123. Vinella, D., Cashel, M. & D'Ari, R. Selected amplification of the cell division genes *ftsQ*–*ftsA*–*ftsZ* in *Escherichia coli*. *Genetics* **156**, 1483–1492 (2000).
124. Yanai, K., Murakami, T. & Bibb, M. Amplification of the entire kanamycin biosynthetic gene cluster during empirical strain improvement of *Streptomyces kanamyceticus*. *Proc. Natl Acad. Sci. USA* **103**, 9661–9666 (2006).

An interesting observation of how high-level gene amplification of the kanamycin biosynthetic gene cluster occurred in *Streptomyces kanamyceticus* as a consequence of selection for strains with increased antibiotic production.

Acknowledgements

This work was supported by grants to D.I.A. from the Swedish Research Council and the European Commission 6th Framework Program, and to L.S. from Uppsala University.

DATABASES

Entrez Gene: <http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=gene>
[ermA](http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=gene) | [rplV](http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=gene) | [tetL](http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=gene)

Entrez Genome Project: <http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=genomeprj>
[Enterococcus faecalis](http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=genomeprj) | [Escherichia coli](http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=genomeprj) | [Proteus mirabilis](http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=genomeprj) | [Salmonella enterica subsp. enterica serovar Typhimurium LT2](http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=genomeprj) | [Staphylococcus aureus](http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=genomeprj) | [Streptococcus agalactiae](http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=genomeprj) | [Streptococcus pneumoniae](http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=genomeprj)

UniProtKB: <http://ca.expasy.org/sprot>
[AmpC](http://ca.expasy.org/sprot) | [L22](http://ca.expasy.org/sprot) | [MobB](http://ca.expasy.org/sprot) | [MobE](http://ca.expasy.org/sprot) | [OmpC](http://ca.expasy.org/sprot) | [OmpD](http://ca.expasy.org/sprot) | [OmpE](http://ca.expasy.org/sprot) | [TEM1](http://ca.expasy.org/sprot)

FURTHER INFORMATION

Dan I. Andersson's homepage: <http://www.imbim.uu.se/forskning/anderssonresearch.html>

ALL LINKS ARE ACTIVE IN THE ONLINE PDF