Order and disorder in bacterial genomes
Eduardo PC Rocha

The availability of sequenced bacterial genomes allows a deeper understanding of their organizational features that are related with fundamental cellular processes such as coordinated gene expression, chromosome replication and cell division. Nevertheless, recent genome comparisons and experimental work highlighted the fluidity of bacterial chromosomes, including genome rearrangements that imperil the selective features of chromosome order. As a result, the clash between elements generating rearrangements and chromosome organization is a classic case of evolutionary conflict.

Introduction
Creativity and organization are in conflict, and bacterial genomes do not escape this general rule; for example, when creativity is related to the generation of genetic novelty, and organization is associated with the stability of selective arrangements of genetic elements in the chromosome. The availability of genome data has allowed the understanding of many features concerning both chromosome organizational structure and the generation of genetic variability in bacterial populations. Chromosome organization involves selective features such as the distribution of genes relative to replication, segregation or expression, which can be biased by expression levels, essentiality or function. It also involves the biased distribution of nucleotides and oligonucleotides. Many of these features have been recently surveyed [1–3].

Generation of genetic variability depends on mutation rates, horizontal gene transfer or intra-genomic recombination mechanisms. The latter have a particularly important role in adaptation via gene dosage effects or fast local sequence variation [4,5]. Yet, they also tend to produce rearrangements in the chromosome, disrupting chromosome organization [6]. This creates an evolutionary conflict between the two processes. The resulting trade-offs between organization and genotypic creativity depend on bacterial ecological, genetical and physiological characteristics such as lifestyles, population sizes, recombination mechanisms and growth rates.

This review focuses on recent advances in the understanding of the balance between the stability of genome organization and the generation of variability through recombination processes.

The constraints of chromosome organization
Bacterial chromosomes are adaptively organized in function of processes such as cellular factories and chromosome replication and segregation (Figure 1). A classic example of the former is the organization of functionally related genes in polycistronic units, allowing the development of sophisticated strategies for the regulation of gene expression [2]. Organization beyond the level of operons resulting from the dynamic interaction of the chromosome and gene expression with the cell factory involve the association between gene expression and cell compartmentalization, chromosome segregation or cell differentiation [1,7]. In Escherichia coli, genes involved in the sulphur metabolism cluster together, possibly to allow compartmentalization of the metabolism itself and its toxic metabolic intermediates [8]. Translation related genes also cluster at a supra-operonic level in many bacterial genomes [9]. For example, highly expressed genes and stress related genes of E. coli and Bacillus subtilis are clustered on the two opposite sides of the origin of replication [10]. Genomic islands are also frequently associated with genes involved in pathogenicity and antibiotic resistance [5,11], but also catabolic pathways [12] and symbiosis [13]. The clustering of closely related functions results in their easier spread by horizontal gene transfer [2]. Physical proximity also facilitates co-expression, since even divergently oriented close operons tend to be co-expressed [14**]. This implies that a proper description of prokaryotic genome organization must include supra-operonic organization [15*], and is consistent with the neighborhood conservation of genes in different operons between distantly related genomes [9].
Well-studied bacteria have one single replication origin (ori), meaning that each replication round doubles the frequency in the cell of genes near the origin relative to genes near the terminus. The gene dosage effect is amplified in fast-growing bacteria containing multiple rounds of replication under exponential growth [16]. This trend creates a gene dosage gradient from the origin towards the terminus of replication. This allows over-expression of highly expressed genes, but also serves as regulatory purposes. In the B. subtilis genome, the positioning of spoIIR near ori allows it to be expressed early in the forespore in the asymmetric segregation during sporulation [1]. The differential expression of the genes distributed along the chromosome might also drive chromosome segregation. Shortly after replication initiation, the new ori regions migrate quickly to opposite poles of the cell. This could result from the motor driving force of the RNA polymerase (coupled with the high frequency of leading strand and highly expressed genes at the origin) [17**] or as a result of ribosome crowding near highly expressed genes that would pull the two origins apart (eventually linked with the simultaneous insertion in the membrane of translated proteins) [10,18]. Recent work suggests that chromosome segregation, in the last replication stages, depends on asymmetrically distributed oligonucleotides between the leading and the lagging strands, and polarized towards the terminus [19]. These motifs could direct the pumping of the chromosome into each of the daughter cells, thus allowing the alignment of the dif site with the septum. Although the sequence and importance of such motifs is still unclear [20], this suggests a direct link between replication-associated biases and chromosome segregation both in early and in late replication stages.

Genes are also more often coded in the leading than in the lagging replicating strand. The level of this gene strand bias varies widely between species, and is probably associated with genome stability and the composition of DNA polymerase [21]. The early observation that rDNA and ribosomal protein genes are systematically coded in the leading strand suggested that avoidance of frequent head-on collisions between the replication fork and the RNA polymerase was responsible for this pattern [22]. This model was recently modified to account for the fact that essential genes, not highly expressed genes, are preferentially positioned in the leading strand [23**,24]. Head-on collisions are more likely to lead to truncated transcripts, which if translated into incomplete peptides result in non-functional proteins and complexes. This is expected to be particularly important for essential genes, and might explain why 94% and 76% of essential genes from B. subtilis and of E. coli, respectively, are coded in the leading strand [23**].

The creative role of repeats in genomes
Recent overviews show that most genomes possess hundreds or thousands of repeated elements capable of recombining by homologous or illegitimate recombination [25**,26]. These repeats are part of duplicated genes, regulatory elements or insertion sequences (IS). They are constantly created by recombination, horizontal transfer, or transposition, and also deleted by further recombination or by the accumulation of point mutations [4,6,25*].

The compactness of bacterial genomes suggests that only selection pressure(s) or self-replication of repeats allow their stable maintenance. Positive selection can result, for example, from gene dosage effects or from the generation of genetic variability. Stable gene amplification results in the multiple copies of rDNA operons in many bacteria that allow faster growth through efficient access to resources [27]. In Pseudomonas aeruginosa, inversions between IS elements allow enhanced adaptation to the host by silencing antigens and lead to hypermutability [28**]. Transient gene amplification is also found when...
gene dosage is important to face unusual concentrations of a toxic substance or a nutrient [29,30], or motivated by selfish elements, such as in the expansion of restriction and modification systems facing replacement by an homologous sequence [31*]. Disappearance of selection pressure usually leads to deletion of the extra copies of genetic material. However, this is not without long-term consequences. It was recently proposed that transient selection for amplification of the lac operon involves co-amplification of a neighbor error-prone polymerase and is responsible for ‘directed’ evolution of E. coli cells under starvation in presence of lactose [32**].

Intra-chromosomal recombination results in local high-frequency sequence changes. It leaves the remaining chromosome sequence unaffected, contrary to the accumulation of point mutations, and only requires elements present in the chromosome, contrary to horizontal transfer. As a result, intra-chromosomal recombination is associated with adaptation strategies to frequently encountered, but intrinsically heterogeneous, stresses [4,5,33]. Illegitimate recombination between closely spaced repeats results in local genetic diversity and typically local rearrangements leading to the variation in the expression of genes with important phenotypical consequences [4,5,33]. Homologous recombination between distant repeats creates genetic variation, for example, by generating new chimerical genes or new genome architectures.

How order balances disorder
Intra-chromosomal recombination, through chromosome rearrangements, antagonizes the selective features of the organization of bacterial genomes (Table 1). The resulting trade-off is a function of the selective advantage of each element in the bacterium evolutionary and ecological context. The negative association between repeats and genome stability was frequently suggested [6,34,35*], and was statistically confirmed for γ-proteobacteria, where the genomes with the most repeats are also the ones showing accelerated rearrangement rates [36**]. This association has prompted experimental work aimed at understanding the dynamics of bacterial chromosomes by testing rearrangement scenarios. Early work focused on the outcomes of homologous recombination between two copies of a repeat in E. coli [37] and Salmonella enterica serovar Typhimurium [38]. In both cases some chromosomal segments were found refractory to inversion, but the respective authors gave different interpretations to their results. Indeed, unobserved inversions can result from recombination impairment by mechanistic problems [38], or from the deleterious effects of the rearrangements [37]. Given the low frequencies of homologous recombination in these experimental setups, work to evaluate the relative importance of each factor has recently used phage related site-specific recombination systems to induce inversions [39**,40*].

In the Firmicute Lactococcus lactis, no non-permissive recombination points were identified in the genome, although some inversions led to lower growth rates or even lethality [39**]. Several of the large inversions leading to lower growth can be attributed to the resulting chromosome asymmetry and/or to loss of advantageous gene dosage effects (Figure 2). Interestingly, all inversions leading to lethality involved switching many genes from the leading to the lagging strand. This is likely to be highly deleterious, if not lethal, in this genome that has an estimated ~90% of essential and overall ~80% of genes in the leading strand [24]. A reappraisal of the homologous recombination experiments in Salmonella [38] and E. coli [37] shows that the majority of the observed non-permissive inversions also involve shifts of leading strand genes to the lagging strand (Figure 2). Some of the exceptions result from inversions placing the origin and terminus of replication close together in the chromosome, which is likely to result in slower growth and complicate chromosome segregation. Among the 20 inversion intervals tested in Salmonella through recombinase-mediated rearrangements, some were more infrequent than others [40*]. Two recombination sites were inaccessible to plasmids, but oddly not to other chromosome regions, suggesting that some DNA regions could be hidden or protected from site-specific recombination relative to other chromosome and extra-chromosome regions [40*]. Constrains on chromosome organization in the nucleoid could lead to regionalization of promiscuous regions

<table>
<thead>
<tr>
<th>Table 1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Elements capable of generating genetic diversity by recombination mechanisms, their putative sources, the mechanisms by which variability can be generated and their effect in the chromosome organization.</td>
</tr>
<tr>
<td><strong>Element</strong></td>
</tr>
<tr>
<td>---</td>
</tr>
<tr>
<td>Long repeats</td>
</tr>
<tr>
<td>Shortly spaced repeats</td>
</tr>
<tr>
<td>Insertion sequences</td>
</tr>
<tr>
<td>Restriction systems</td>
</tr>
</tbody>
</table>
within the genome, and data from *L. lactis* and *E. coli* also identified one origin and one terminus domain refractory to inversions [38,39**].

A recent work used the genome information of *Sinorhizobium meliloti* to identify large repeats, mostly IS, present in more than one of its three replicons (one chromosome and two mega-plasmids) [41**]. These 133 repeats were then found to provide for integration hotspots by homologous recombination in 91 cases. A variant where all replicons were merged into one single chromosome showed similar growth rate and symbiosis efficiency in

---

**Figure 2**

(a) Different consequences of different rearrangements (inversions) scenarios resulting from intra-chromosomal recombination between repeats (arrows indicate inversion breakpoints). The text labels are as in Figure 1. (b) A summary of some of the experiments performed in *L. lactis* [39**] and *Salmonella* [38], in relation to the distribution of essential and highly expressed genes. Inversions take place between one fixed point and another element in the chromosome linked by the semi-circle. Black lines indicate permissive intervals, red lines indicate inversions that are not possible or viable, and blue lines indicate inversions leading to lower growth rate (not indicated for *Salmonella*). In the center circles are distributed the highly expressed (black) and essential (red) genes in the two genomes (as determined elsewhere [24]). The leading strand is green and the lagging strand is red.
rich medium, but lower growth rate in minimal medium, relative to the wild-type. Since the original genomic structure is usually found in nature, the selective disadvantages of these kinds of events will merit further experimental analysis in an evolutionary context.

Systematic competition experiments between inverted and wild-type strains, allowing the assessment of the relative fitness cost for chromosome rearrangements, are unfortunately still unavailable. Because the rearrangements observed in wild-type genomes have been subject to natural selection, they might enlighten the effects of selection on the rearrangements that did occur in nature. Although rearrangements in the genomes of enterobacteria are frequent in laboratory conditions they are rarely observed in nature [6]. For example, very few inversions are frequent in laboratory conditions and wild-type strains, allowing the assessment of the relative fitness cost for chromosome rearrangements, are unfortunately still unavailable. Because the rearrangements observed in wild-type genomes have been subject to natural selection, they might enlighten the effects of selection on the rearrangements that did occur in nature. Although rearrangements in the genomes of enterobacteria are frequent in laboratory conditions they are rarely observed in nature [6]. For example, very few inversions are observed between the sequenced genomes of E. coli and S. enterica, which diverged about 100 million years ago. This probably reflects the selective constraints associated with chromosome organization (Figure 1). By disrupting selective features such as gene dosage, chromosome symmetry or gene strand bias, inversions often lead to lower growth rates [37,39**,42]. Unsurprisingly, the most frequently observed inversions between rDNA or IS are symmetrical relative to the origin and terminus of replication in bacteria [6,43] and in archaea [44]. These rearrangements minimize the disruption of the chromosome organization relative to replication and segregation [42,45]. Symmetrical rearrangements could also result from the biased repair of stalled replication forks by illegitimate recombination with the ssDNA present in the other replication fork [43]. The importance of this mechanism relative to the counter-selection of highly asymmetric rearrangements remains to be determined.

Since different positioning of repeats leads to different rearrangements, one might expect the distribution of repeats to be non-random. Although both direct and inverse repeats may generate sequence diversity by recombination, only the latter lead to chromosome inversions. Therefore, chromosome organization can be partly reconciled with the existence of repeats, if these are mostly in the direct conformation. This is indeed observed in most genomes [25*], and is particularly remarkable in Mycoplasma genitalium and Mycoplasma pneumoniae. These genomes rely on homologous recombination to produce genetic variability in their major adhesin. Because these genomes code 80% of their genes in the leading strand, inversions are expected to be strongly counter selected. As a result, direct repeats outnumber inverse repeats by a factor larger than nine in M. pneumoniae and larger than 57 in M. genitalium [46*], and only translocations, but not inversions, are observed [47]. Repeats also tend to be placed in the chromosomes around the origin of replication more symmetrically than expected [25*]. Whether such repeats are the cause and/or the consequence of frequent symmetric recombination remains to be determined. In any case, and as described above, they are expected to be the ones that produce the less disorganizing chromosome inversions (Figure 2). All these results suggest that trade-offs between positive selection associated with some repeats and purifying selection for genome organization can lead to the confinement of instability in certain regions of the chromosome.

**Taming disorder**

Genetic markers in the linear chromosome of Streptomyces are lost with frequencies of $10^{-4}$ to $10^{-2}$ per spore, and deletions often attain more than 2 Mb (i.e. around a fourth of the genome) [48]. Other instabilities in these genomes involve large chromosome arms deletions, replacements and switches, as well as chromosome circularization after the deletion of telomeres and fusion resulting in partial genome duplication [48–50]. The genome sequence of S. coelicolor nicely illustrates why these genomes are so tolerant to these changes [51**]. The central half of the chromosome, around the replication origin, includes most of the conserved genes and nearly all the essential genes. Inversely, the regions near the telomeres include many genes coding for antibiotic production/resistance and transposases. These regions are thus intrinsically unstable. This is because of lower purifying selection for chromosome organization in regions lacking housekeeping genes and higher positive selection for diversification of antibiotics-associated genes. Such a confinement of genomic instability can be found in many other genomes. Borrelia burgdorferi generates genetic variability by recombination between repeats present in plasmids [52], which allows the stabilization of the linear chromosome while maintaining the generation of genetic diversity. In Pirunella, 65% of transposases and 85% of integrases/recombinases are in the chromosome half surrounding the terminus, where they mediated a large inversion [53]. Finally, several genomes show evidence of plasticity zones, where deletions, inversions and insertions of genetic material are much more frequent than in the rest of the chromosome [54–56]. Order and disorder can be partly reconciled by the stabilization of some regions of the chromosome, where positive selection for organization is stronger, while leaving others free to evolve and generate variability.

The number of repeats and the rate of genome rearrangements are associated with different ecological constraints. Many bacteria are under periodic stresses; for example, imposed by the immune system for human pathogens. As a result of the process of adaptation to these stresses, they have evolved sophisticated strategies for generating variability at the relevant loci by recombination. However, the chromosomes of obligatory intracellular pathogens, such as Chlamydia, Spirochetes, or Rickettsia, exhibit...
few repeated elements, and this might be related to their relatively protected intracellular environment. Conversely, the chromosomes of mycoplasmas [46] and facultative pathogens such as Helicobacter pylori [57], Streptococcus [58], Tropheryma whippleti [59], or Neisseria [60] show many repeated elements allowing the generation of antigen and tropism variation, and resulting in frequent genome rearrangements. Several of these pathogens, because they need to contain the recombination elements allowing fast adaptation, have probably relaxed the selection for some organizational features of their chromosome and this is revealed in their frequent rearrangements. Within strains of the same species, the differences in ecology can also lead to different rearrangement rates. Within S. enterica, the host-specific serovars exhibit more frequent rearrangements than the generalists, even though their intra-chromosomal recombination rate is similar [61]. Specialists tend to suffer from more frequent bottlenecks than generalists, which facilitates the fixation of mildly deleterious rearrangements. This underlines the difference between rearrangement events in nature and in the laboratory: purifying selection acts on many features of chromosomal organization and its intensity depends on population parameters such as size and clonality.

Conclusions and perspectives

The use of E. coli as a model organism has contributed much to our understanding of bacterial genetics. However, if comparative genomics is to allow us to understand organization and evolution of the genome it must now be complemented by experimental studies in distantly related species, where many variables remain unknown. First, many of the organizational elements of the bacterial chromosome are poorly understood. These include the reasons why Firmicutes have much higher frequencies of leading strand genes [21], the higher-than-operonic level of chromosome organization of gene expression [9,14,15], and the role and importance of signals directing chromosome segregation [20]. Until these elements have been properly characterized it is difficult to assign them a quantitative fitness effect and to understand why they show such a large variability among bacteria. Second, the experimental study of intra-chromosomal recombination mechanisms has been historically confined to enterobacteria. However, the genes related with recombination vary significantly even among the genomes of γ-proteobacteria [62,63]. Hence, chromosomes are likely to engage into intra-chromosomal recombination with different frequencies in different species. It has been pointed out that the lack of key enzymes in recombination pathways (e.g. RecBCD and RecA) is responsible for the genome stability of intracellular bacteria such as Rickettsia and Buchnera [62,63]. As a case in point, these genomes also have very few large repeats and no IS. However, this association between recombination-deficient genomes and stability is not trivial, since the partial inactivation of RecA in Streptomyces lividans has the exact opposite effect of destabilizing the chromosome [64]. This is not completely surprising as RecA is also involved in the repair of stalled replication forks and in the regulation of the SOS response. Under these conditions, the stability of Buchnera and Rickettsia would be the result of recombination deficiency and very stable environments, where replication fork arrests are rare, selection for genetic variation weak and repeats nearly nonexistent.

Experimental approaches to study chromosome evolution and tolerance to changes include induction of chromosome inversions [37,38,39,40], subdivision [65], circularization [59], and fusion [41,42]. Artificial manipulations of genome structure raise questions about the viability of the constructs in nature when submitted to natural selection. This will have to be tackled by experimental evolutionary studies. The fitness effects of chromosomal organisational features must also be more precisely evaluated and placed into the ecological context of the different species. For example, gene dosage effects are expected to be important only in fast-growing bacteria [3], whereas gene strand bias seems to be independent of the growth rate [21], and horizontal transfer is more important in larger genomes [2]. Comparative genomics will then have a major role in the understanding of the trade-off between order and disorder in bacterial genomes. Such studies could provide important clues on how the selective features of chromosome organization can or cannot be harmonized with the creative features of genetic change that allows the fast adaptation of bacteria.

Acknowledgements

I'm most indebted to the many people with whom I've worked and discussed on the issues presented in this review, in particular to Antoine Danchin, Guillaume Achaz and Alain Blanchard. Gilles Fischer, Pascal Le Bourgeois and Guillaume Achaz provided most valuable comments and criticisms to the manuscript. I apologize for all the works that could not be cited due to space limitations.

References and recommended reading

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

- of special interest
- ** of outstanding interest


This work shows that co-expression in bacteria is frequent among proximal genes even if they do not belong to the same operon (e.g. divergently oriented genes). The localisation of such co-expressed genes tends to be conserved in evolution.


The analysis of gene position and orientation in bacterial genomes using signal-processing techniques indicates that gene clustering and orientation cannot be associated with a single characteristic number of units. This suggests that genes tend to assemble and co-orient over any scale of observation greater than a few kilobases. Therefore, gene organization is important well beyond the operon level.


RNA polymerase is a powerful and abundant motor in replicating bacterial cells. If the RNA polymerase movements are restricted in the cell, its action results in DNA translocation. Because highly expressed genes are more frequent near the origin in chromosomes of fast growing bacteria, RNA polymerase can have an important role in separating the newly replicating chromosomes (see also [10,18]).


The essential character of a gene, not its expression rate, is a major determinant of its position in the leading strand of genomes. This suggests that collisions between polymerases may result mostly in transcription abortions, with consequences for the performance of essential functions. It also highlights an important constraint to genome rearrangements (see also [24]).


If the presence of repeats is positively selected or caused by self-replication but disrupts chromosomal organization, then one would expect repeats to be distributed as to minimize their negative impact. This paper is the first contribution to indicate that the distribution of repeats in bacterial genomes is indeed constrained in such a way, but at a very different degree in different genomes.


This paper shows that rearrangements by homologous recombination between insertion sequences provide an adaptive advantage to opportunistic Pseudomonas aeruginosa strains. Adaptation is provided by silencing antigens and by inducing mutator phenotypes. The latter may allow a wider exploration of the fitness landscape by the population of bacteria.


Attempts to replace a restriction and modification (RM) system from the B. subtilis chromosome resulted in the amplification of the corresponding genes. This is mediated by the RM functions themselves.


DinB-dependent mutagenesis in stationary phase may result in adaptive evolution at the lac locus because of the co-amplification of dinB with lac. Amplification of the error-prone polymerase dinB would lead to general mutagenesis and increasing the frequency of Lac" mutants. Subsequent selection for amplified leaky lac mutants decreases and deletions lead to a single Lac" locus. Although it is controversial if this implies that general mutagenesis is not a stress response, it nicely illustrates the important roles of transient amplifications in genome evolution.


The analysis of three closely related Bordetella genomes, suggests that B. pertussis and B. parapertussis are derivatives of B. bronchiseptica-like ancestors. Evolution of these two host-restricted pathogens resulted in a large number of gene deletions and some rearrangements. The latter are very abundant in B. pertussis where they are flanked by insertion sequences.

www.sciencedirect.com

Current Opinion in Microbiology 2004, 7:519-527


This manuscript describes the propensity for genome inversions in a non-γ-proteobacterial genome. Inversions in Lactococcus lactis confirmed features observed in enterobacteria (i.e. ori and ter domains, deleterious effects of replicating strand switches), but the study also tests recombination intermediates and analyses the effects of inversions on growth rates.

40. Garcia-Russell N, Harmon TG, Le TO, Amaladas NH, Mathewson RD, Segall AM: Unequal access of chromosomal regions to each other in Salmonella: probing chromosome structure with phage lambda integrase-mediated long-range rearrangements. Mol Microbiol 2004, 52:299-344. Among the 16 inversions shown in this work that do not contain the two ‘hidden’ loci, seven are frequent and nine are rare or very rare. This illustrates the heterogeneity of recombination encounters between different places in the genome. Yet, among the frequent inversions, only one leads to switching genes from the leading to the lagging strand, whereas fitness effects account for part of the measured lower recombination frequencies.

41. Guo X, Flores M, Mavingui P, Fuentes SI, Hernandez G, Davila G, Palacios R: Natural genomic design in Sinorhizobium meliloti: novel genomic architectures. Genome Res 2003, 13:1810-1817. Departing from the identification of large repeats common to at least two of the three replicons of Sinorhizobium meliloti, the authors predict and find genome architectures resulting from multiple integration events. It is unclear at this stage how replication and segregation proceed in the integrated genome, and how inversions in co-integrates affect the recovery of the original plasmids. Yet, this clearly demonstrates the plasticity of the Sinorhizobium genome.


45. Mackiewicz P, Mackiewicz, D, Kowalczuk M, Cebrat S: Pyrococcus furiosus/C.15/C15: aspects of the rare or very rare inversions. This suggests that fitness effects account for part of the measured lower recombination frequencies.

46. Rocha EPC, Blanchard A: Genomic repeats, genome plasticity and the dynamics of Mycoplasma evolution. Nucleic Acids Res 2002, 30:2031-2042. Contrary to other small genomes, mycoplasmas, which are chronic pathogens, possess many repeats that allow the generation of adaptive genotypic diversity. However, possibly because the most variable genes correspond to very different proteins in different genomes, the types and roles of these repeats also differ.


51. Bentley SD, Chater KF, Cerdeno-Taraga AM, Challis GL, Thomson JR, James KD, Harris DE, Quail MA, Kieser H, Harper D et al.: Complete genome sequence of the model actinomycete Streptomyces coelicolor A3(2). Nature 2002, 417:141-147. The central half of the Streptomyces coelicolor chromosome is conserved and contains most of the housekeeping functions and essential genes. The rest includes many horizontally transferred genes (e.g. for antibiotic production) and repeated elements, resulting in a remarkable confinement of chromosomal instability.


