Polarisation of prokaryotic chromosomes

Jean R Lobry* and Jean-Michel Louarn†

In many prokaryotes, asymmetrical mutational or selective pressures have caused compositional skews between complementary strands of replication arms, especially sensitive in the distribution of guanine and cytosine. In *Escherichia coli*, most of the guanine/cytosine skew is caused by mutation rates differing on leading and lagging strands, but contribution of skewed functionally important guanine-rich motifs (Chi and Rag sites), which control chromosome repair or positioning, is noticeable. Interference between replication and gene expression plays a minor role. The situation may be different in other bacteria. Studies of chromosome processing and bacterial taxonomy might profit from consideration of chromosome polarisation.

Addresses
*Laboratoire de Biométrie et Biologie évolutive, CNRS UMR 5558, Université Claude Bernard, 43 Boulevard du 11 Novembre 1918, F-69622 Villeurbanne cedex, France
e-mail: lobry@biomserv.univ-lyon1.fr
†Laboratoire de Microbiologie et de Génétique moléculaire, CNRS UMR 5100, 118 route de Narbonne, F-31062 Toulouse cedex, France
e-mail: louarn@ibcg.biotoul.fr

Current Opinion in Microbiology 2003, **6**:101–108

This review comes from a themed issue on
Cell regulation
Edited by Andrée Lazdunski and Carol Gross

1369-5274/03/$ – see front matter

DOI 10.1016/S1369-5274(03)00024-9

Abbreviations

CDR	chromosome dimer resolution
DAZ	*dif* activity zone
PR1	parity rule 1
PR2	parity rule 2

Introduction

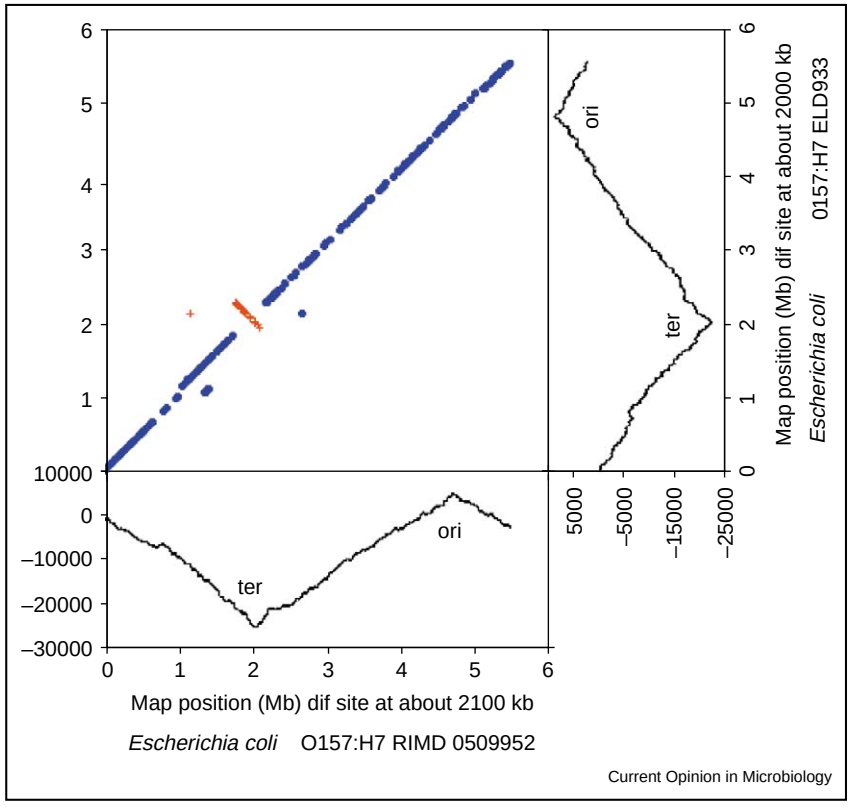
In bacteria and their phages, guanine (G)/cytosine (C) and adenine (A)/thymine (T) ratios of complementary strands often differ. This property has long been utilised for the purification of complementary strands of phage DNAs [1–3] and even of megabase-sized genomes, such as that of *Streptococcus pneumoniae* [4]. The information from large-scale sequencing confirms that most bacterial genomes display base composition skews [5••,6]. The term 'chirochore' was coined to describe a DNA sequence displaying such a compositional skew [7]. As a general rule, each replication arm of a bacterial chromosome appears to be a chirochore, and the term 'replichore' expresses this fact [8]. Thus, this compositional skew generates a dual polarisation on a bi-directionally replicating chromosome. Genomicists now tend to consider a shift of the G/C skew to mark the origin and terminus of replication. This correlation has proved to be useful as the starting point for mapping replication origins in bacteria [9–11], but the biological significance of such G/C skews requires thorough discussion. In this review, we present a discussion, centred on *Escherichia coli*, of the phenomenon.

When mutation and selection are assumed to be symmetric with respect to both DNA strands, and then the rates of reciprocal substitutions in a specific strand to be equal according to the parity rule 1 (PR1) of Sueoka [12], the intrastrand base composition at equilibrium is expected to reach the parity rule 2 (PR2) state, in which the frequencies of adenine and thymine, and guanine and cytosine, are equal within each strand [12,13]. This PR2 state is more than an equilibrium point because convergence towards PR2 continues even if substitution rates are modified. There is no way to escape from the PR2 state under the PR1 hypothesis [14]. Therefore, G/C and A/T skews, violations of the PR2 state, are strong indications that some selective or mutational pressure affected asymmetrically complementary strands during chromosome evolution.

Chromosome polarisation conservation

Chromosomal gene order is not generally conserved during bacterial evolution [15]. Even the organisation of operons is plastic [16]. The few operons that are conserved in prokaryotes code for physically interacting proteins [17•]. In this chaotic situation, the comparison of prokaryotic chromosomes has, nevertheless, revealed a conservation of the orientation and the relative distance of certain genes (or groups of genes) from the origin and terminus of replication, at least during recent evolution. This was found by applying the matrix dot plot [18] at the genome scale to observe characteristic X-shaped dot plots (as exemplified in Figure 1) crossing on the replication origin-to-terminus axis [19–21,22••,23••,24,25••, see also update]. Thus, the chromosome polarisation revealed by gene orientation in the replichores tends to be conserved during evolution [26••]. This observation is intriguing. The majority of bacterial genes are encoded on the leading strand [27], so it is legitimate to wonder whether A/T and G/C skews result primarily from constraints imposed to gene expression by replication, turning into constraints to base composition imposed by the genetic code, or whether they result primarily from a mutation pressure that is different on leading and lagging strands, turning into usage (therefore selection) of the skews, for instance to control operations of chromosome processing. Evidence supporting the second model in the case of

Figure 1



Example of a recent chromosomal inversion between the genome sequences of two different isolates of *Escherichia coli* 0157:H7, and its absence of effect on the A/T and G/C skews. The central panel is a dot plot comparison at the DNA level, at which each point represents 100% exact identity over 5 kb in the direct versus direct (blue dots) or direct versus complementary (red crosses) strands between the two chromosomes. The genomes of the two isolates are almost identical (although one is from Japan and the other from the USA), with most differences corresponding to sequencing indeterminations. The inversion is located near the terminus of replication, and the *dif* site is located within the inverted region. External panels represent the cumulated A/T and G/C composite skew in third codon positions of coding sequences, as computed with the *oriloc* software [68]. The maximum value of the composite skew index is located at *oriC*, the origin of replication, and the minimum value at *terC*, where termination occurs preferentially. Note that the general pattern is the same between the two isolates. This is because the inversion includes the terminus region, so that asymmetric features are conserved. This is a general property of polarised features, not restricted to A/T and G/C skews. (See the enlightening figures in [25••] for an explanation.) Sequence accession numbers in the DDBJ/EMBL/GenBank database are BA000007 and AE005174 for the x axis and y axis, respectively.

E. coli is presented. The situation can be different in other bacteria, as discussed later.

Preferential gene orientation contributes little to the G/C skew in *E. coli*

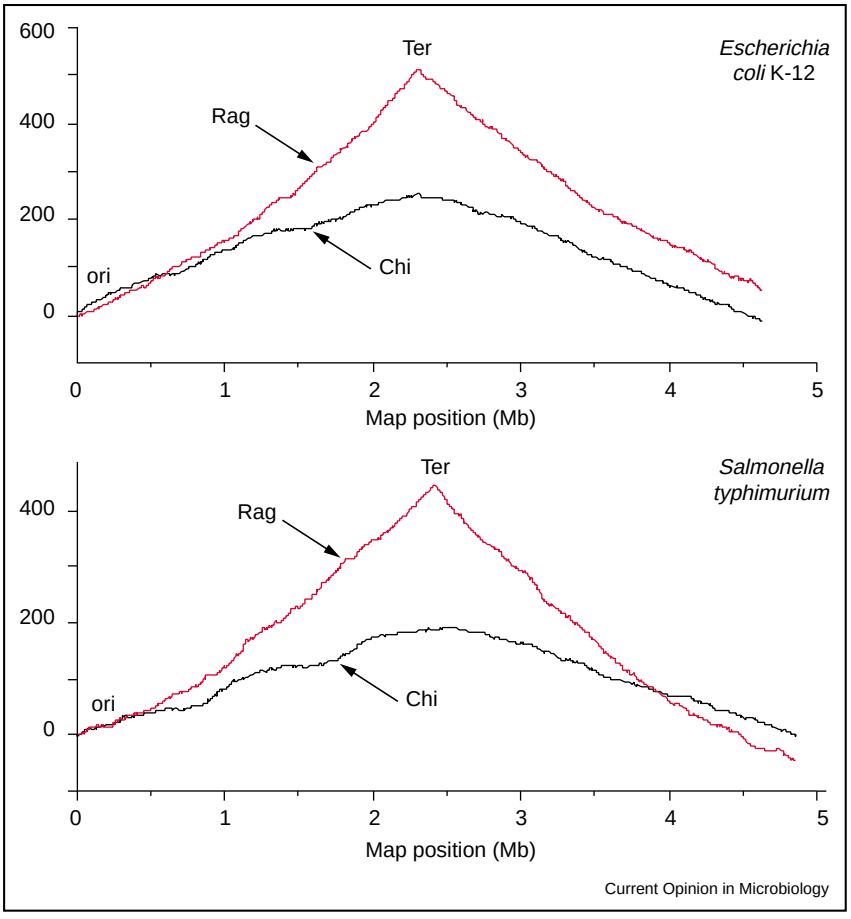
In *E. coli*, 54% of coding sequences are found on the leading strand, but highly expressed genes are more often transcribed in the direction of replication. Thus, constraints imposed to gene expression by replication may exist. For instance, transcription may oppose unidirectional obstacles to replication [27]. This idea has received some experimental support [28]. The constraints involved are weak, as many inversions within a replication arm result in normal viability at the laboratory scale [29].

McLean *et al.* [6] noticed that, in *E. coli* and other bacterial genomes, the G/C skew is stronger for the third letter of codons than for the two first letters. Tillier and Collins [20,30] confirmed the observation, and showed that the sequence of a gene may be influenced by whether it is encoded on the leading or lagging strand. In general, biases are higher at weakly selected positions, third codon positions or intergenic spaces, and are highly correlated between these two classes (see [5••] for a recent update). These observations do not support the contention that, in *E. coli*, constraints owing to codon usage have shaped the base composition skews.

G-rich skewed elements of *E. coli*: the Chi and Rag motifs

Salzberg *et al.* [31] found two families of skewed G-rich octamers in *E. coli*. The skew of these octamers shifts sign near *oriC* and in the terminus region, and is stronger than that expected from the global G/C skew, as exemplified

Figure 2



Simple DNA walks for Rag motifs (RGNAGGGS) and Chi sites (GCTGGTGG) along the *E. coli* K-12 chromosome (top) and the closely related *Salmonella typhimurium* LT2 chromosome (bottom). The walker is gliding along the published strand of the chromosome and moves up by one unit when encountering the motif under analysis, and moves down by one unit when the complement of the motif is encountered. For instance, in the *E. coli* chromosome, starting from the origin (ori) up to the terminus, there are about 500 more Rag motifs in the direct orientation than in the reverse orientation, and this bias is inverted after the terminus. In both organisms, the shift of strand preference occurs precisely at *dif* for this motif (see also Figure 3), whereas the shift for Chi-site strand preference is much less clear-cut. The walk would yield a linear trace if the bias were constant along the chromosome, but note that for Rag motifs, the curve is lambda-shaped, signifying that the bias increases towards the terminus.

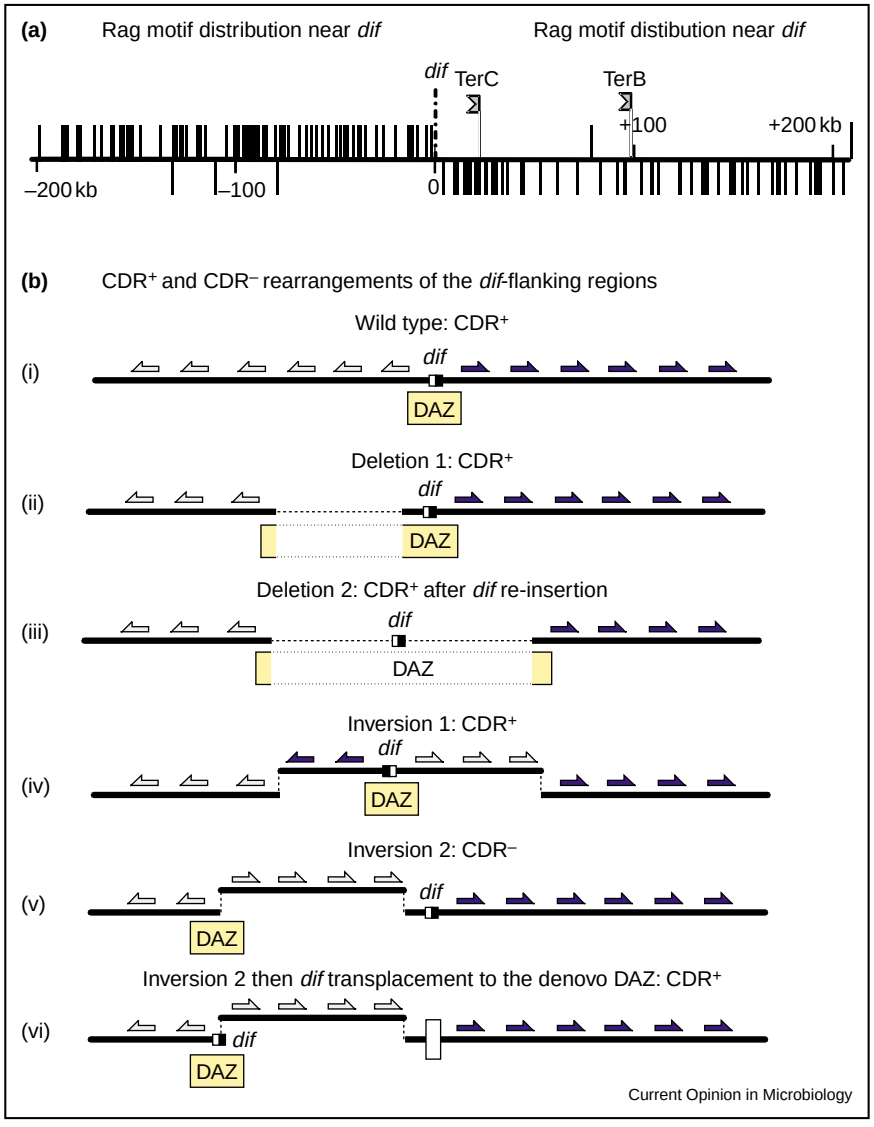
below. The skew of some motifs of each family is functionally important. Chi sites (GCTGGTGG) belong to the TGGT family, and 75% of them are found on the leading strand (Figure 2). The predicted orientation bias, if it were solely dictated by the excess of G on the leading strand, is estimated to 56.5% for this 5G-1C-containing octamer (see Box 1). Thus, the bias of Chi sites strongly exceeds the predicted value. When a RecBCD complex moving on DNA meets the 3′ end of a Chi site, this complex shifts from a purely nucleolytic activity to become a producer of a 3′-ended single-strand, the substrate for homologous recombination. In its preferential orientation, a Chi site may facilitate the re-starting of aborted replication forks [32,33•]. The selective value of this orientation may have favored the generation of the skew, but other explanations can be considered [34,35].

The skew of Rag motifs (RGNAGGGS) is more remarkable. Their actual skew of 82% is much stronger than the theoretical value, predicted to 56.7% from the skew of their components (see Box 1), and increases towards the terminus (Figure 2). The skew shift occurs abruptly in the terminus and, *dif*, a 28 bp site devoted to chromosome

Box 1 Expected skews.

Oligomer expected skew values were calculated for *E. coli* replichore 1. Probabilities of occurrence of a given oligomer and of its complement were deduced from base composition of the leading strand, assuming no other constraint: A, 0.2444; T, 0.2460; C, 0.2469; G, 0.2627. For example, the Chi-site expected frequency on this strand is $F1=[(0.2627)^5\ (0.2460)^2\ (0.2469)$. The frequency of its complement is $F2=[(0.2469)^5\ (0.2444)^2\ (0.2627)]$. Chi expected skew is $(F1/F1+F2)\times 100$.

Figure 3



The activity of *dif* in chromosome dimer resolution (CDR) is controlled by repeated and oriented elements of the flanking regions. **(a)** *dif* is located in the junction between two regions strongly polarised by Rag elements (indicated by vertical lines; their nearly 100% skew shifts sign at *dif*). **(b)** Genetic analysis of the DAZ phenomenon. (i) Experiments of *dif* transplacement indicate that *dif* allows CDR only if located in a small zone (10–20 kb) around its natural map position. (ii) Deletions near *dif*, whatever their importance, conserve a CDR^+ phenotype, provided *dif* is conserved. (iii) Deletions, including *dif*, whatever their importance, are of course CDR^-, but regain a CDR^+ phenotype, provided *dif* is re-inserted at the junction between the severed chromosomes. (iv) Inversions, including the *dif* site, remain CDR^+, whatever their importance. (v) Inversions affecting the regions near *dif*, but not *dif* itself, are CDR^-, provided the *dif*-proximal endpoint maps within 20–30 kb from *dif*. (vi) Inversion of *dif*-flanking regions may generate a *de novo* DAZ at the *dif*-distal endpoint, demonstrated by a CDR^+ phenotype if *dif* is transplaced to this inversion endpoint. Altogether, these results support the model that the DAZ is the junction between series of multiple and interchangeable inversely oriented elements (DAZ-organising polar elements [DOPES], represented here by arrows), for which Rag sequences are excellent candidates (see [a]).

dimer resolution (CDR), maps at the point where polarity switches [36]. The same phenomenon is conserved in many chromosome termini (not shown). An unusual feature of *dif* activity in dimer resolution is that the *dif* site must be located within a small region, the *dif* activity zone (DAZ). Genetic analysis [37–39] revealed that the DAZ occurs at the junction between regions displaying opposite polarisation (Figure 3). The polarizing motifs, not yet rigorously identified, are probably the Rag motifs, and the polarity-reading factor is most probably FtsK. This septum-anchored protein [40•] is a DNA translocase. It apparently mobilises DNA in a direction dictated by its polarity, and stops where polarity shifts, that is, near *dif*. It ensues that *dif* sites of a chromosome dimer are placed at the septum, favoring resolution by site-specific recombination [41••–43••]. FtsK may be the clearing factor for

ridding the port of the septum of DNA loops ([43••]; J Corre and J-M Louarn, unpublished data). Because chromosome positioning makes the terminus stand in the cell close to the septum or where it will form [44], the extreme skew of Rag motifs of the terminus (97%) has perhaps evolved primarily to help FtsK prevent the capture of the region by the septum, and was subsequently used for dimer resolution.

Altogether, the strong skew of Chi and Rag elements accounts for about 7% of the global G/C skew of *E. coli* (but for about 14%, if the extended Rag family RRNAGGG is counted). These specific G-rich motifs contribute importantly to the global skew, but other G-rich oligomers less so. The actual orientation bias of all 5G-containing octamers found on the *E. coli* chromosome (60% [45]) is not very different from the predicted one, 56.2% (see Box 1). It must be emphasised here that, given that the G/C skew remains detectable after removal of all skewed octamer sequences [30], functions involving skewed motifs have not been the sole mechanism responsible for the chirochore structure of this chromosome.

Other examples of polarised repeated elements

In *E. coli*, the first repeated polar elements identified were the replication pause (Ter) sites. They all map in the same orientation on a given replichore (5 on each replichore). These 23 bp sequences are rather AT-rich [46]. The existence of polarised elements, of unknown nature, in the regions peripheral to the terminus may be inferred from data on the deleterious effects of certain inversions [29,47]. The other repeated elements so far identified, such as the BIMEs, are not polarised [48].

Origin of G/C and A/T skews: the differential mutation pressure model

Given that, in *E. coli*, the chirochore structure cannot be explained solely by preferential gene orientation and G-rich skewed motifs, contribution of asymmetric mutation pressures, which could be caused by differences in replication modes of leading and lagging strands, must have been important [49••]. Unequal fidelity of leading and lagging strand synthesis has been detected, by looking at the rates of chromosomal spontaneous mutations, in *mutL* mutants (a genotype favoring rapid evolution [50]). For instance, T/G mispairing (the most frequent mismatch) occurs preferentially during leading strand synthesis [51]. Applied to a genome under initial PR2 state, this should make this genome gain Gs and Ts on the leading strand. This fits with observed skews [5••]. It was also observed that +1G frameshifts occur fivefold more frequently during synthesis of series of Gs on the leading strand, whereas −1G frameshifts occur at comparable frequencies on both strands [52••]. This may have favored series of Gs on the leading strand, for example, 6-G series display an orientation bias of 67% significantly greater than the expected one, 60% (see Box 1). Preferential cytosine deamination (C → T), which happens *in vitro* more frequently on single-stranded DNA, could also affect preferentially the leading strand [53]. Other mechanisms might coexist, as indicated by the observation that, in *Streptomyces coelicolour*, the biases observed are in complete contradiction with the predictions from the cytosine deamination model [54••].

When evolution meets polarisation

The conservation of *dif* and of polarisation of its flanking regions occurs in a context of great variability within the terminus regions, which are among the most divergent segments in closely related bacteria [22••]. For instance, the regions flanking *dif* are different in *Salmonella typhimurium* and *E. coli*, although *dif* and Rag-determined opposite polarities are conserved (Figure 2). The rapid evolution of terminus correlates with its recombination [55–57,58••]. RecD inactivation is known to trigger hyperrecombination in the terminus [55], and this is perhaps the key to understanding terminus evolution. Linear DNA entering into the bacterium is degraded up to the first correctly oriented Chi site by RecBCD, which then loses its RecD-associated nuclease activity. The bacterium consequently becomes a $RecD^-$ phenocopy [59,60], which is manifested by an increased recombination activity in the terminus. The region would become a preferential target for integration of exogenous DNA. But the new material will be tolerated only if adequately polarised. Mutation selection will tend to eliminate inappropriately oriented polar motifs and to gain correctly oriented ones. The pressure for polarisation may prevail over that for gene functionality, explaining why prophages present in the *E. coli* terminus are extremely polarised, and cryptic, and why the region harbors no essential genes. The terminus has often been considered as junk, but this disparaging view has to be revised; the region is important, but only its polarisation matters, not its gene content.

Conclusions

The present review has centered on *E. coli*, the only organism in which polarity-mediated functions have been identified. Combining the functions implicated in chromosome processing with strand-specific mutability due to replication asymmetry provides a rationale for its chromosome polarisation. Aside from fork repair and septum clearing, other functions — condensation and separation of nascent chromosomes, perhaps replication itself —could use polarisation to monitor their progress through the chromosome. Like the Chi-directed recombination and FtsK-dependent mobilisation, they might use motifs that arose from the mutation-determined G/C skew. Skewed sequences clearly require more attention, as do studies on the functions of polarisation in chromosomal processing.

Differing situations may be found in other bacteria. The noticeable G/C skew of *Haemophilus influenzae* [7] is not

associated with highly skewed short motifs [31,45], although the dimer chromosome resolution system is close to that of *E. coli* (F-X Barre, personal communication). Strongly asymmetric mutation pressures may have predominantly contributed to preferential gene orientation in genomes displaying very strong chirochore structure, such as that of *Borrelia burgdorferi* (see [61••] for a detailed analysis of the surprising genome of this bacterium). Conversely, in *Mycoplasma genitalium*, a weak asymmetric mutation pressure is associated with strong chirochore structure. In this organism, about 80% of the coding sequences are found on the leading strand. The shaping factor is probably a strong interference between replication and transcription [5••]. New taxonomy guidelines will likely emerge as analysis of bacterial sequences allows us to trace the conservation and evolution of processes generating chromosome polarisation, or making use of polarisation.

Update

Recent studies have extended the taxonomic range of X-shaped dot plots with the genomic comparison of *Oceanobacillus iheyensis* versus *Bacillus subtilis* and *B. halodurans* (Figure 4 in [62]), *Brucella suis* versus *Mesorhizobium loti* (Figure 3 in [63]), *Agrobacterium tumefaciens* and *Sinorhizobium meliloti* (Figure 3 in [64], Figure 3 in [65]) *Streptococcus agalactiae* versus *S. pyogenes* and *S. pneumoniae* (Figure 3 in [66]).

Acknowledgements

We gratefully thank Stanislav Cebrat, Henry Krisch and Noboru Sueoka for useful comments, careful lecture and correction of the manuscript.

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

1. Cordes S, Epstein HT, Marmur J: **Some properties of the deoxyribonucleic acid of phage alpha**. *Nature* 1961, **191**:1097-1098.

2. Hradecna Z, Szybalski W: **Fractionation of the complementary strands of coliphage λ DNA based on the asymmetric distribution of the poly I,G-binding sites**. *Virology* 1967, **32**:633-643.

3. Rudner R, Karkas JD, Chargaff E: **Separation of microbial deoxyribonucleic acids into complementary strands**. *Proc Natl Acad Sci USA* 1969, **63**:152-159.

4. Roger M: **Evidence for conversion of heteroduplex transforming DNAs to homoduplexes by recipient pneumococcal cells (DNA strand resolution-bacterial transformation-genetic recombination)**. *Proc Natl Acad Sci USA* 1972, **69**:466-470.

5. •• Lobry JR, Sueoka N: **Asymmetric directional mutation pressures in bacteria**. *Genome Biology* 2002, **3**:RESEARCH0058.
This paper presents a recent analysis of the G/C and A/T skews in 43 sequenced bacterial chromosomes. It shows that skews are highly correlated with weakly selected positions, and that they are not correlated with the overall G+C content.

6. McLean MJ, Wolfe KH, Devine KM: **Base composition skews, replication orientation, and gene orientation in 12 prokaryote genomes**. *J Mol Evol* 1998, **47**:691-696.

7. Lobry JR: **Asymmetric substitution patterns in the two DNA strands of bacteria**. *Mol Biol Evol* 1996, **13**:660-665.

8. Blattner FR, Plunkett G III, Bloch CA, Perna NT, Burland V, Riley M, Collado-Vides J, Glasner JD, Rode CK, Mayhew GF *et al.*: **The complete genome sequence of *Escherichia coli* K-12**. *Science* 1997, **277**:1453-1474.

9. Picardeau M, Lobry JR, Hinnebusch BJ: **Physical mapping of an origin of bidirectional replication at the center of the *Borrelia burgdoferi* linear chromosome**. *Mol Microbiol* 1999, **32**:437-445.

10. Myllykallio H, Lopez P, Lopez-garcia P, Heilig R, Saurin W, Zivanovic Y, Philippe H, Forterre P: **Bacterial mode of replication with eukaryotic-like machinery in a hyperthermophilic archaeon**. *Science* 2000, **288**:2212-2215.

11. Zawilak A, Cebrat S, Mackiewicz P, Krol-Hulewicz A, Jakimowicz D, Messer W, Gosciniak G, Zakrewska-Czerwinska J: **Identification of a putative chromosomal replication origin from *Helicobacter pylori* and its interaction with the initiator protein DnaA**. *Nucleic Acids Res* 2001, **29**:2251-2259.

12. Sueoka N: **Intrastrand parity rules of DNA base composition and usage biases of synonymous codons**. *J Mol Evol* 1995, **40**:318-325. [published erratum appears in *J Mol Evol* 1995, **42**:323.]

13. Lobry JR: **Properties of a general model of DNA evolution under no-strand-bias conditions**. *J Mol Evol* 1995, **40**:326-330. [Published erratum appears in *J Mol Evol* 1995, **41**:680.]

14. Lobry JR, Lobry C: **Evolution of DNA base composition under no-strand-bias conditions when the substitution rates are not constant**. *Mol Biol Evol* 1999, **16**:719-723.

15. Mushegian AR, Koonin EV: **Gene order is not conserved in bacterial evolution**. *Trends Genet* 1996, **12**:289-290.

16. Itho H, Takemoto K, Mori H, Gojobori T: **Evolutionary instability of operon structures disclosed by sequence comparisons of complete microbial genomes**. *Mol Biol Evol* 1999, **16**:332-346.

17. • Wolf YI, Rogozin IB, Kondrashov AS, Koonin EV: **Genome alignment, evolution of prokaryotic genome organization, and prediction of gene function using genomic context**. *Genome Res* 2001, **11**:356-372.
Because functionally related genes are often clustered in operons, one would expect such structures to be conserved on an evolutionary scale. This is clearly not the case, as demonstrated by the systematic search performed by the authors of this paper on 25 complete prokaryotic genomes. This is an up-to-date demonstration of this surprising negative result, with many references to previous and consistent findings.

18. Gibbs AJ, McIntyre GA: **The diagram, a method for comparing sequences. Its use with amino-acid and nucleotide sequences**. *Eur J Bioch* 1970, **16**:1-13.

19. Read TD, Brunham RC, Shen C, Gill SR, Heidelberg JF, White O, Hickey EK, Peterson J, Utterback T, Berry K *et al.*: **Genome sequence of *Chlamydia trachomatis* MoPn and *Chlamydia pneumoniae* AR39**. *Nucleic Acids Res* 2000, **28**:1397-1406.

20. Tillier ER, Collins RA: **Genome rearrangement by replication-directed translocation**. *Nat Genet* 2000, **26**:195-197.

21. Eisen JA, Heidelberg JF, White O, Salzberg SL: **Evidence for symmetric chromosomal inversions around the replication origin in bacteria**. *Genome Biology* 2000, **1**:RESEARCH0011.1-RESEARCH0011.9.

22. •• Suyama M, Bork P: **Evolution of prokaryotic gene order: genome rearrangements in closely related species**. *Trends Genet* 2001, **17**:10-13.
In Figure 1, the dot plots are ordered according to increasing divergence at the amino acid level. The progressive X-fication from Figure 1a–g is striking. The last two panels, corresponding to *Mycoplasma* and *Chlamydia*, are outliers: their rate of genome rearrangement is abnormally low (neglecting the effect due to substitution rate acceleration in these lineages). The authors of this paper suggest that this could be a consequence of the deficiency in RecG helicase in these species.

23. •• Moran NA, Mira A: **The process of genome shrinkage in the obligate symbiont *Buchnera aphidicola*. *Genome Biology* 2001, **2**:RESEARCH0054.1-RESEARCH0054.12.
Buchnera aphidicola is an obligate endosymbiont of aphids. Compared with *E. coli*, a closely related free-living bacteria, its genome size has been

drastically reduced from 4.6 Mb to 0.6 Mb. Despite this sevenfold difference in genome size, an X-shaped dot plot is visible in Figure 6 of this paper.

24. Takami H, Nakasone K, Takaki Y, Maeno G, Sasaki R, Masui N, Fuji F, Hirama C, Nakamura Y, Ogasawara N *et al.*: **Complete genome sequence of the alkaliphilic bacterium *Bacillus halodurans* and genomic sequence comparison with *Bacillus subtilis***. *Nucleic Acids Res* 2000, **28**:4317-4331.

25. •• Zivanovic Y, Lopez P, Philippe H, Forterre P: ***Pyrococcus* genome comparison evidences chromosome shuffling-driven evolution**. *Nucleic Acids Res* 2002, **30**:1902-1910.
There are also X-shaped dot plots in archae (Figure 1), confirming their bacteria-like mode of chromosome replication.

26. •• Mackiewicz P, Mackiewicz D, Kowalczuk M, Cebrat S: **Flip-flop around the origin and terminus of replication in prokaryotic genomes**. *Genome Biology* 2001, **2**:INTERACTIONS1004.1-INTERACTIONS1004.4.
This is a very interesting discussion about the potential underlying reasons of chromosome polarisation conservation. The impact of various chromosomal inversion scenarios on chromosome polarisation is well illustrated by Figure 3. Only a symmetrical inversion encompassing the origin or terminus maintains both chromosome polarisation and replichore length (scenario 3a). In complement to the authors' arguments for their third point, note that a chromosome-length polymorphism study for *Escherichia coli* is in favor of a selective force to maintain replichores of similar size [67].

27. Brewer BJ: **When polymerases collide: replication and the transcriptional organization of the *E. coli* chromosome**. *Cell* 1988, **53**:679-686.

28. French S: **Consequences of replication fork movement through transcription units *in vivo***. *Science* 1992, **258**:1362-1365.

29. Rebollo JE, François V, Louarn JM: **Detection and possible role of two large nondivisible zones on the *Escherichia coli* chromosome**. *Proc Natl Acad Sci USA* 1988, **85**:9391-9395.

30. Tillier ERM, Collins RA: **The contributions of replication orientation, gene direction, and signal sequences to base-composition asymmetries in bacterial genomes**. *J Mol Evol* 2000, **50**:249-257.

31. Salzberg SL, Salzberg AJ, Kerlavage AR, Tomb JF: **Skewed oligomers and origins of replication**. *Gene* 1998, **217**:57-67.

32. Kuzminov A: **Recombinational repair of DNA damage in *Escherichia coli* and bacteriophage lambda**. *Microbiol Mol Biol Rev* 1999, **63**:751-813.

33. • Gruss A, Michel B: **The replication-recombination connection: insights from genomics**. *Curr Opin Microbiol* 2001, **4**:595-601.
This review updates the work by Kuzminov (1999) [32] and shows, in particular, that the skew of Chi sites may vary with the organism.

34. El Karoui M, Biaudet V, Schbath S, Gruss A: **Characteristics of Chi distribution on different bacterial genomes**. *Res Microbiol* 1999, **150**:579-587.

35. Colbert T, Taylor AF, Smith GR: **Genomics, Chi sites and codons: ''islands of preferred DNA pairings'' are oceans of ORFs**. *Trends Genet* 1998, **14**:485-488.

36. Capiaux H, Cornet F, Corre J, Guijo M, Perals K, Rebollo JE, Louarn JM: **Polarization of the *Escherichia coli* chromosome: a view from the terminus**. *Biochimie* 2001, **83**:161-170.

37. Cornet F, Louarn J, Patte J, Louarn JM: **Restriction of the activity of the recombination site *dif* to a small zone of the *Escherichia coli* chromosome**. *Genes Dev* 1996, **10**:1152-1161.

38. Kuempel P, Hogaard A, Nielsen M, Nagappan O, Tecklenburg M: **Use of a transposon (Tndif) to obtain suppressing and nonsuppressing insertions of the *dif* resolvase site of *Escherichia coli***. *Genes Dev* 1996, **10**:1162-1171.

39. Perals K, Cornet F, Merlet Y, Delon I, Louarn JM: **Functional polarization of the *Escherichia coli* chromosome terminus: the dif site acts in chromosome dimer resolution only when located between long stretches of opposite polarity**. *Mol Microbiol* 2000, **36**:33-43.

40. • Donachie WD: **FtsK: Maxwell's demon?** *Mol Cell* 2002, **9**:206-207.
This is a recent review on the multiple activities of the FtsK protein as a keystone of the replication-division coupling.

41. •• Aussel L, Barre FX, Aroyo M, Stasiak A, Stasiak AZ, Sherratt D: **FtsK is a DNA motor protein that activates chromosome dimer resolution by switching the catalytic state of the XerC and XerD recombinases**. *Cell* 2002, **108**:195-205.
See annotation Corre and Louarn (2002) [43••].

42. •• Capiaux H, Lesterlin C, Perals K, Louarn JM, Cornet F: **Functional replacement of the *E. coli dif* recombination site reveals a dual role for the FtsK protein in chromosome segregation**. *EMBO Reports* 2002, **3**:523-526.
See annotation Corre and Louarn (2002) [43••].

43. •• Corre J, Louarn JM: **Evidence from terminal recombination gradients that FtsK uses replichore polarity to control terminus positioning at division**. *J Bacteriol* 2002, **184**:3801-3807.
Papers [41••–43••] present the available evidence to show that FtsK is a DNA translocase [41••], which guides the sites involved in dimer resolution to the septum [42••], and does this by mobilizing terminus DNA in a polarity-dependent manner [43••].

44. Niki H, Yamaichi Y, Hiraga S: **Dynamic organization of chromosomal DNA in *Escherichia coli***. *Genes Dev* 2000, **14**:212-223.

45. Uno R, Nakayama Y, Arakawa K, Tomita M: **The orientation bias of Chi sequences is a general tendency of G-rich oligomers**. *Gene* 2000, **259**:207-215.

46. Coskun-Ari FF, Hill TM: **Sequence-specific interactions in the Tus-Ter complex and the effect of base pair substitutions on arrest of DNA replication in *Escherichia coli***. *J Biol Chem* 1997, **272**:26448-26456.

47. Guijo MI, Patte J, del Mar Campos M, Louarn JM, Rebollo JE: **Localized remodeling of the *Escherichia coli* chromosome: the patchwork of segments refractory and tolerant to inversion near the replication terminus**. *Genetics* 2001, **157**:1413-1423.

48. Bachellier S, Clement JM, Hofnung M: **Short palindromic repetitive DNA elements in enterobacteria: a survey**. *Res Microbiol* 1999, **150**:627-639.

49. •• Kowalczuk M, Mackiewicz P, Mackiewicz D, Nowicka A, Dudkiewicz M, Dudek MR, Cebrat S: **DNA asymmetry and the replicational mutational pressure**. *J Appl Genet* 2001, **42**:553-577.
An up-to-date review about DNA asymmetries and potential underlying mechanisms, the best current entry to previous literature on the subject.

50. Taddei F, Radman M, Maynard-Smith J, Toupance B, Gouyon PH, Godelle B: **Role of mutator alleles in adaptive evolution**. *Nature* 1997, **387**:700-702.

51. Fijalkowska IJ, Jonczyk P, Tkaczyk MM, Bialoskorska M, Schaaper RM: **Unequal fidelity of leading strand and lagging strand DNA replication on the *Escherichia coli* chromosome**. *Proc Natl Acad Sci USA* 1998, **95**:10020-10025.

52. •• Gawel D, Maliszewska-Tkaczyk M, Jonczyk P, Schaaper RM, Fijalkowska IJ: **Asymmetry of frameshift mutagenesis during leading and lagging-strand replication in *Escherichia coli***. *Mutat Res* 2002, **501**:129-136.
Papers [51,52••] report measurements of mutation frequencies within the *lac* operon inserted in either orientations at the *attλ* position on the chromosome. The authors have taken advantage of well-identified *lacZ* mutants to determine frequencies and strand preferences of various mutational events. They clearly establish the higher mutability of the leading strand during normal replication (the use of *mutL* strains, eliminating post replicative repair of mismatches, facilitated the detection of effects due to replication errors).

53. Frank AC, Lobry JR: **Asymmetric substitution patterns: a review of possible underlying mutational or selective mechanisms**. *Gene* 1999, **238**:65-77.

54. •• Bentley SD, Chater KF, Cerdeño-Tárraga AM, Challis GL, Thomson NR, 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.

This paper presents the first counter-example to the universality of compositional biases orientations, implying the rejection of the cytosine deamination theory as a universal mechanism at the origin of asymmetric directional pressure in bacteria.

55. Corre J, Patte J, Louarn JM: **Prophage lambda induces terminal recombination in *Escherichia coli* by inhibiting chromosome dimer resolution: an orientation-dependent *cis*-effect lending support to bipolarization of the terminus**. *Genetics* 2000, **154**:39-48.

56. Horiuchi T, Fujimura Y, Nishitani H, Kobayashi T, Hidaka M: **The DNA replication fork blocked at the Ter site may be an entrance for the RecBCD enzyme into duplex DNA**. *J Bacteriol* 1994, **176**:4656-4663.

57. Peters JE, Craig NL: **Tn7 transposes proximal to DNA double-strand breaks and into regions where chromosomal DNA replication terminates**. *Mol Cell* 2000, **6**:573-582.

58. •• Huber KE, Waldor MK: **Filamentous phage integration requires the host recombinases XerC and XerD**. *Nature* 2002, **417**:656-659.

Papers [55–57,58••] present all recombination events disclosed in the terminus, in addition to dimer resolution. [55] shows that the *dif* region undergoes recombinogenic lesions at very high frequency when chromosome dimer resolution is inhibited, and when RecD is inactivated. [56] provided the first indication that fork arrest at Ter sites is recombinogenic. [57] shows an example of preferred transposition in the terminus (generalised Tn7 transposition). [58••] describes how the *dif* site of *Vibrio cholerae* chromosome 1 has been used for integration of a bacteriophage carrying important virulence determinants.

59. Myers RS, Kuzminov A, Stahl FW: **The recombination hot spot chi activates RecBCD recombination by converting *Escherichia coli* to a recD mutant phenocopy**. *Proc Natl Acad Sci USA* 1995, **92**:6244-6248.

60. Koppen A, Krobitsch S, Thoms B, Wackernagel W: **Interaction with the recombination hot spot chi *in vivo* converts the RecBCD enzyme of *Escherichia coli* into a chi-independent recombinase by inactivation of the RecD subunit**. *Proc Natl Acad Sci USA* 1995, **92**:6249-6253.

61. •• Kowalczuk M: **Mutational and selection pressures in the *Borrelia burgdorferi* genome [PhD Thesis]**. Wroclaw: Wroclaw University; 2002.

This polish PhD dissertation (core text in English) is available online in Smorfland's web site (http://smorfland.microb.uni.wroc.pl/). It describes a valuable study of an intriguing situation: *Borrelia* genome is exceptional in that the asymmetric mutation pressure is so high that the amino acid content of proteins is greatly affected. This is because the genome has undergone the most extreme, to our current knowledge, asymmetric mutation pressure.

62. Takami H, Takaki Y, Uchiyama I: **Genome sequence of *Oceanobacillus iheyensis* isolated from the Iheya Ridge and its unexpected adaptive capabilities to extreme environments**. *Nucleic Acids Res* 2002, **30**:3927-3935.

63. Paulsen IT, Seshadri R, Nelson KE, Eisen JA, Heidelberg JF, Read TD, Dodson RJ, Umayam L, Brinkac LM, Beanan MJ *et al.*: **The *Brucella suis* genome reveals fundamental similarities between animal and plant pathogens and symbionts**. *Proc Natl Acad Sci USA* 2002, **99**:13148-13153.

64. Goodner B, Hinkle G, Gattung S, Miller N, Blanchard M, Qurollo B, Goldman BS, Cao Y, Askenazi M, Halling C *et al.*: **Genome sequence of the plant pathogen and biotechnology agent *Agrobacterium tumefaciens* C58**. *Science* 2001, **294**:2323-2328.

65. Wood DW, Setubal JC, Kaul R, Monks DE, Kitajima JP, Okura VK, Zhou Y, Chen L, Wood GE, Almeida NF Jr *et al.*: **The genome of the natural genetic engineer *Agrobacterium tumefaciens* C58**. *Science* 2001, **294**:2317-2323.

66. Glaser P, Rusniok C, Buchrieser C, Chevalier F, Frangeul L, Msadek T, Zouine M, Couve E, Lalioui L, Poyart C *et al.*: **Genome sequence of *Streptococcus agalactiae*, a pathogen causing invasive neonatal disease**. *Mol Microbiol* 2002, **45**:1499-1513.

67. Bergthorsson U, Ochman H: **Distribution of chromosome length variation in natural isolates of *Escherichia coli***. *Mol Biol Evol* 1998, **15**:6-16.

68. Frank AC, Lobry JR: **Oriloc: prediction of replication boundaries in unannotated bacterial chromosomes**. *Bioinformatics* 2000, **16**:560-561.