American Journal of Virology 2 (1): 1-7, 2013
ISSN: 1949-0097
©2013 Science Publication
doi:10.3844/ajvsp.2013.1.7 Published Online 2 (1) 2013 (http://www.thescipub.com/ajv.toc)

The Bacterial Chromosomal Sequence and Related Issues

Thi Huyen Phan and Duc Luong Nguyen

Department of Biotechnology, Faculty of Chemical Engineering,
Ho Chi Minh City University of Technology, VNU-HCM,
268 Ly Thuong Kiet St., Ward 14, District 10, Ho Chi Minh City, Vietnam

Received 2012-08-31, Revised 2013-01-01; Accepted 2013-07-04

ABSTRACT

There has been a rapid increasing in the number of bacterial genomes that are completely sequenced recently. However, nucleotide arrangement in bacterial chromosome is still not clear, though it is known that the chromosomal strand is symmetric as a whole but asymmetric locally. The unequal distribution of adenine and thymine as well as cytosine and guanine in the local genomic sequences has been suggested to be associated with transcription and replication. In the bacterial chromosome, distribution of nucleotides in the leading and lagging strands, which interact with the structurally asymmetric DNA polymerase during replication process, has also been reported to be biased. Recent studies on changing the nucleotide sequence of bacterial chromosomes have revealed the importance of nucleotide arrangement in chromosome, which exerts influences on success of the chromosome engineering experiments. This review highlights common properties of bacterial chromosomal sequences with related issues involved in the nucleotide arrangement in chromosome, emphasizing that arrangement of nucleotides in chromosome regarding local strand asymmetry and chromosomal strand symmetry needs to be made clear so as to help in experimental design for effectiveness in changing chromosomal sequences or creating artificial chromosomes.

Keywords: Bacterial Chromosome, Strand Symmetry, Strand Asymmetry, Biased Nucleotide Distribution, Chromosomal Change

1. INTRODUCTION

Little is known about the arrangement of nucleotides in bacterial chromosomes while the number of bacterial genomes that are completely sequenced has been increasing. With advances in bioinformatics as well as in molecular technology and genetic manipulation, these completely sequenced genomes have become a valuable resource for understanding the chromosomal structure and essentially biological processes in bacterial cells. However, the technical advances are not enough for one to be always successful in changing the chromosomes, as recent efforts on changing the bacterial chromosomes have revealed the involvement of chromosomal nucleotide arrangement in biological processes which exert influences on success of the chromosomal engineering experiments. This article reviews some common properties of bacterial chromosomal sequences and related issues that would help in figuring out the chromosomal nucleotide arrangement.

1.1. The Chromosomal Sequence and Different Replicating Strands

Most bacterial chromosomes are circular, double-stranded DNA molecules, which follow the Chargaff's first parity rule, i.e., in each molecule, numbers of adenines and thymines are similar as are those of cytosines and guanines (Chargaff, 1950). It has been found that the Chargaff's first parity rule applied to single-strand DNA sequences, which is called the

Corresponding Author: Thi Huyen Phan, Department of Biotechnology, Faculty of Chemical Engineering, Ho Chi Minh City University of Technology, VNU-HCM, 268 Ly Thuong Kiet St., Ward 14, District 10, Ho Chi Minh City, Vietnam

Chargaff's second parity rule (Rudner *et al*., 1968; Forsdyke and Mortimer, 2000), is also true for the chromosomes. Even when applied to oligomers, the Chargaff's second parity rule still holds true, i.e., in the whole single-stranded sequence of chromosome, number of an oligomer is almost equal to that of its reverse complement. The chromosomal DNA strand is thus said to be symmetric (Qi and Cuticchia, 2001; Baisnee *et al*., 2002; Powdel *et al*., 2009). Shown in **Fig. 1** is the strand symmetry in *Bacillus subtilis* and *Escherichia coli* chromosomes. In contrast, in local chromosomal sequences, the Chargaff's second parity rule is violated, i.e., local sequences of the chromosome are asymmetric (Freeman *et al*., 1998; Mradzek and Karlin, 1998; Lobry and Sueoka, 2002). **Figure 2** shows GC skew as a measure of strand asymmetry in every 1-kb segment of the *B.subtilis* and *E. coli* chromosomes. These strand symmetric and asymmetric results pose a question of how the nucleotides are arranged in chromosome so that the local genomic strand is asymmetric but the whole chromosomal strand becomes symmetric.

In relation to biological processes associated with the chromosomal DNA, the nucleotide skew has been recognized to correlate with direction of transcription (Bell and Forsdyke, 1999). Green *et al*. (2003); Touchon *et al*. (2005) and Polak and Arndt (2008) further suggested the local genomic DNA skews to be involved in replication and transcription. The bacterial circular chromosome usually possesses a replication origin and a replication terminus, with its two replichores are approximately equal in length. During the course of DNA synthesis, the replication process occurs bi-directionally from the replication origin (**Fig. 3**).

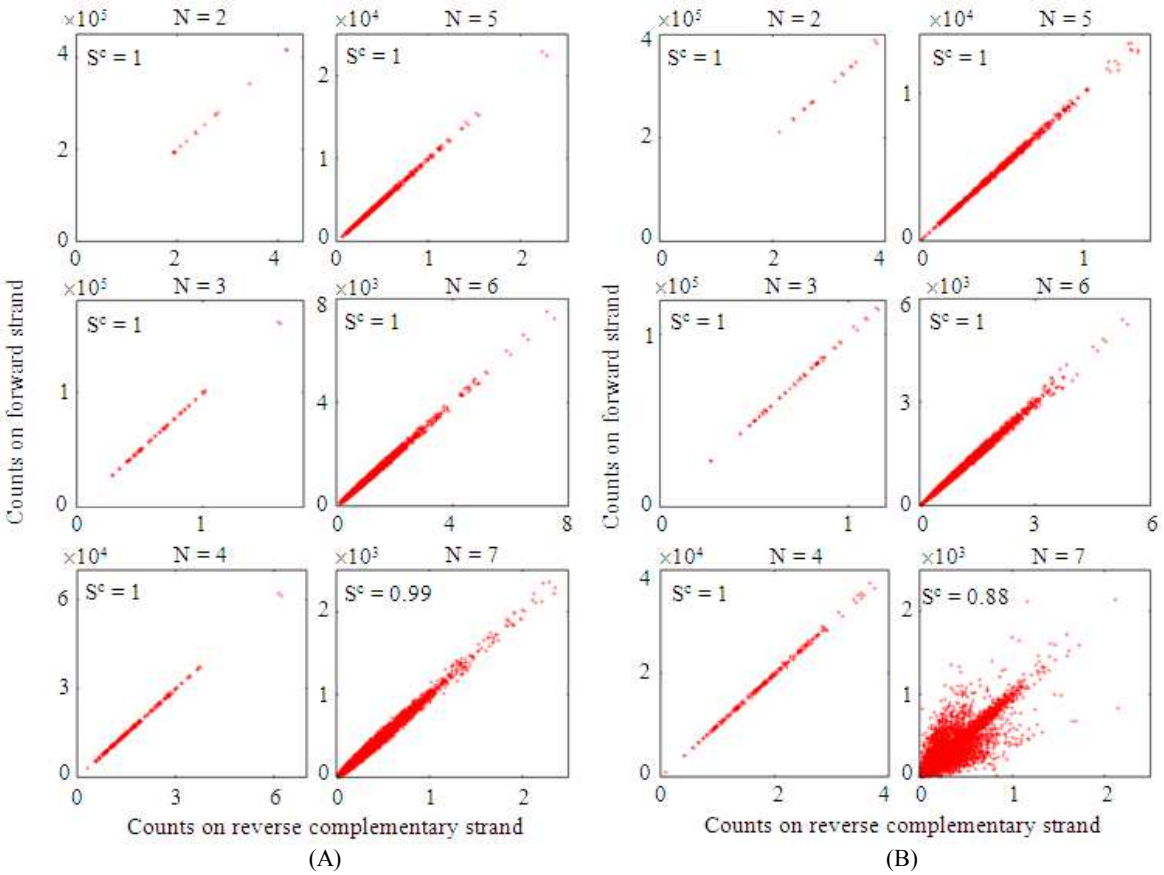

Fig. 1. Strand symmetry in bacterial chromosomes. Plotted are oligomer counts, with the oligomer size N = 2, 3, 4, 5, 6 or 7 in the forward strand and the reverse complementary strand of *B.subtilis* strain 168 chromosome (A) and *E. coli* strain K12 substrain MG1655 chromosome (B), according to the method described by Baisnee *et al*. (2002). S^c shows the degree of chromosomal strand symmetry

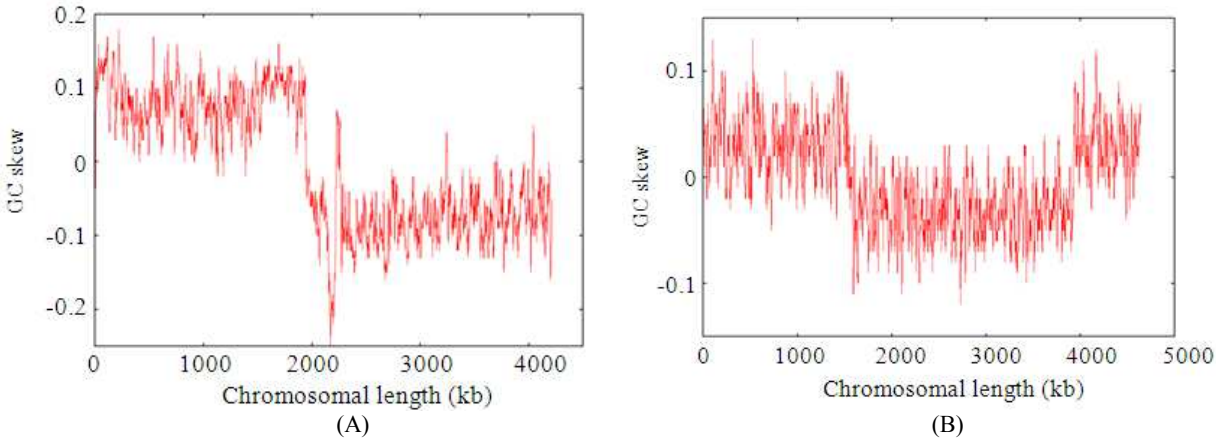

Fig. 2. GC skews in the local sequences of bacterial chromosomes. GC skew was calculated and plotted for every 1kb-segment along the *B. subtilis* strain 168 chromosome (A) and the *E. coli* strain K12 substrain MG1655 chromosome (B). Note that replication origin and terminus positions are at 0 and 1947 kb, respectively, in (A) and at 3925 and 1550 kb, respectively, in (B), according to Worning *et al*. (2006)

Lobry and Sueoka (2002) reported that in the bacterial chromosomes, third codon positions in the leading strand were relatively rich in guanine than in cytosine and in thymine than in adenine in comparison with the lagging strand. Frequencies of some DNA octamers in bacterial chromosomes were also reported to be remarkably different in the two complementary strands and also in both sides of the replication origin (Salzberg *et al*., 1998). Obviously, DNA arrangement in the leading strand is different from that in the lagging strand. However, detail about this difference needs to be more investigated. It has been reasoned that the difference in frequencies of oligomers in the leading and lagging strands are probably due to different causes such as the distribution of nucleotides in the chromosome, the biased distribution of genes and corresponding regulation signals in the two replicating strands (Rocha, 2004). Another study transformed the gene order and strand position in a chromosome *in silico* turned the chromosome out to be symmetric (Poptsova *et al*., 2009).

1.2. The Asymmetric DNA Polymerase, the Leading and Lagging Strands

In the replication process, the DNA is synthesized in the 5' to 3' orientation, as illustrated in **Fig. 3**. While the replication fork is moved in either direction from the origin to the terminus, one of the two chromosomal strands must be replicated as leading and the other as lagging. Prior to the process of DNA synthesis, the DNA polymerase binds to the complementary DNA strands separated by the single-stranded binding proteins. Using each strand as a template, it initiates the DNA synthesis by elongating the RNA primer, which is already built on the template by the RNA primase, in the 5' to 3' orientation. Structurally, the DNA polymerase is an asymmetrical protein dimer (Maki *et al*., 1988; Maki and Kornberg, 1988; Glover and McHenry, 2001; McHenry, 2003). The first unit of the dimer composes of protein subunits including α, β, θ, ε and τ, while the second unit composes of subunits α, β, θ, ε, γ, δ, δ', χ and ψ. The α, ε and θ components in each unit together form the polymerase core which is capable of catalyzing DNA synthesis. During replication, the core in the first unit is responsible for synthesizing the leading strand while that in the second unit is responsible for synthesizing the lagging strand. The β subunit, or the sliding clamp, is responsible for promoting and helping the polymerase core to bind to the DNA template, while δ and δ' subunits are considered to make unique contact with τ and γ, respectively. τ is presumed to act as a tether holding the lagging polymerase core, enabling this lagging strand polymerase to cycle when a blocking oligonucleotide is encountered. χ and ψ subunits are suggested to participate specifically in lagging strand replication. It is obvious that the replication machineries for the leading and the lagging strands are not identical. The question remains is how the nucleotides are organized in the chromosome to be compatible with the replication machineries.

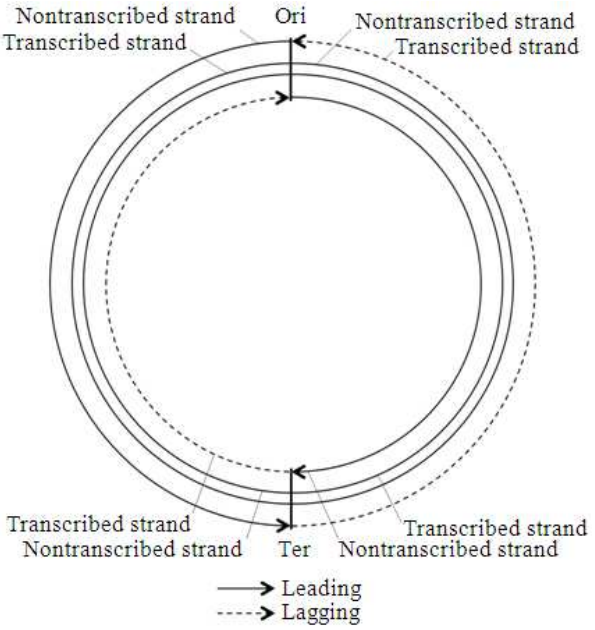

Fig. 3. Bidirectional replication in the bacterial chromosome. Ori and Ter denote the origin and terminus of replication, respectively. Two middle circles denote the double-stranded chromosomal sequence with its two strands serving as templates for synthesizing the strands denoted by the smallest and the largest circles. Transcribed strand and Nontranscribed strand denote the strand that serves as the template for transcription and the strand that is complementary to the template for transcription, respectively

As the degree of nucleotide skews in different bacterial chromosomes are different and the DNA arrangement in the leading strand is different from that in the lagging strand of each bacterial chromosome, with related issues presented below, nucleotide arrangement in the leading and lagging strands as well as aspects in interactions between the replication machineries and the complementary chromosomal DNA strands need to be discovered.

1.3. The Oligonucleotide-Mediated recombination

Using short single-stranded oligonucleotides, studies on changing nucleotide sequence of chromosome during replication process have been shown to be successful (Ellis *et al*., 2001; Li *et al*., 2003). To change the nucleotide sequence of chromosome, one must design an oligonucleotide so that it can be recombined to either the Transcribed (T) strand or the Non-Transcribed (NT) strand, as shown in **Fig. 3**, during replication by the λ-Red Beta protein. However, the recombination efficiency of the single-stranded oligonucleotides to the T strand is not the same as that to the NT strand. In addition, the recombinants may not be obtained. One the one hand, Ellis *et al*. (2001) reported that the recombination efficiency of single-stranded oligonucleotides to the NT strand was higher or lower than that to the T strand depending upon the replication direction. On the other hand, Li *et al*. (2003) showed that this was also dependent on the sequence of single-stranded oligonucleotide. Changing the chromosomal sequence using this way has been suggested to be involved in the replication process (Huen *et al*., 2006). The efficiency of recombination using short single-stranded oligonucleotides was reported to be also influenced by the transcriptional activity, of which the presence of RNA polymerase on the T strand was proved to be the cause (Liu *et al*., 2002). Nevertheless, bias in recombination of the single-stranded oligonucleotides to the NT and T strands obtained by Liu *et al*. (2002) is not the same as that obtained by Li *et al*. (2003). In reality, the single-stranded oligonucleotide sequences and the targeted sequences used in these two studies were different. Apparently, chromosomal sequence changes via recombination process mediated by short oligonucleotides are dependent upon the nucleotide sequence of oligonucleotides and targeted DNA which has been suggested to be related and associated to the processes of transcription and replication. While the nucleotide arrangement in chromosome is still not clear, choosing targeting chromosomal strands for recombination and designing oligonucleotide sequences are still challenged jobs for the oligonucleotide-mediated recombination.

1.4. The Chromosomal Sequence Changes and Negative Consequences

The *E. coli* genome has been known to organize into four macrodomains and two non-structured regions (Niki *et al*., 1999). The macrodomains are also known not to interact with each other but with the neighboring non-structured DNA (Valens *et al*., 2004). Thiel *et al*. (2012) found that one of the macrodomains, the Ter macrodomain, was constrained for reducing DNA mobility and delaying loci segregation when it was associated with the cell division machinery. It is shown that intrareplichore inversions to some extent perturbed nucleoid distribution and early steps of cytokinesis, leading to wrong distribution of nucleoid and thus to abnormal cell morphology (Esnault *et al*., 2007).

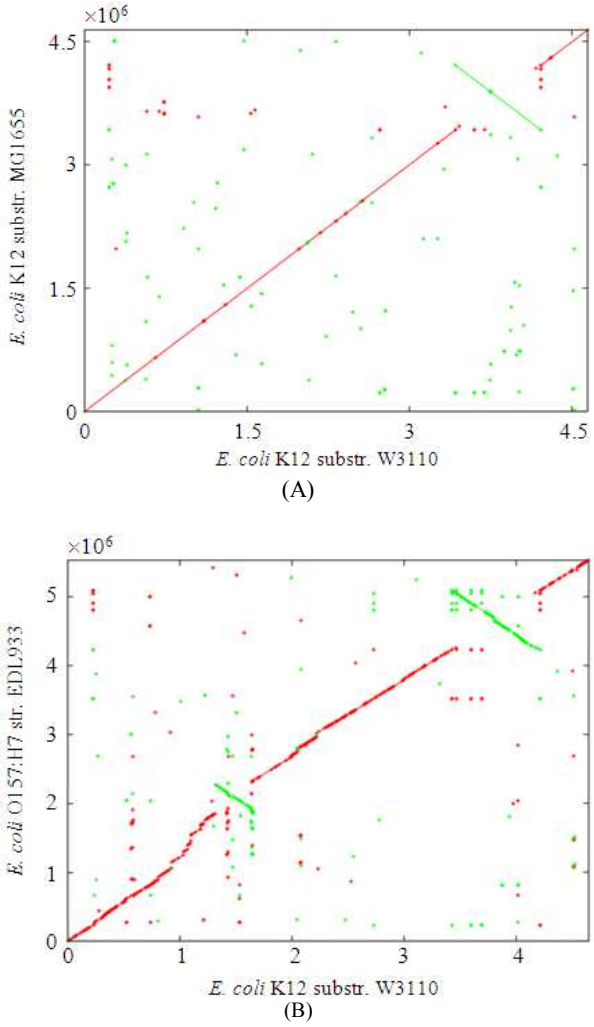

Fig. 4. Common interreplichore inversions in closely related bacterial chromosomes. The chromosomal sequences were compared using MUM mer plot (Kurtz *et al*., 2004). Each axis shows the length of chromosome ($\times 10^6$bp) of each bacterium indicated by its name. Red plots denote the homology between the forward sequences in the first chromosome and the forward sequences in the second chromosome while blue plots denote that between the forward sequences in the first and the reverse complementary sequences in the second. Replication origin and terminus positions in *E. coli* K12 substrain MG1655 chromosome are at 3925 and 1550 kb, respectively, in *E. coli* K12 substrain W3110 chromosome at 3714 and 1554 kb, respectively and in *E. coli* O157:H7 strain EDL933 chromosome at 4783 kb and 2010 kb, respectively (Worning *et al*., 2006)

Interreplichore inversions also led to wrong distribution of nucleoid in the cell, which seriously affected processes prior to cytokinesis, making the cell unable to divide or able to divide for abnormal nucleoid containing cells (Niki *et al*., 1999; Esnault *et al*., 2007). In *B. subtilis*, changing the position of replication origin or making the intrareplichore inversions also reduced the replication rate (Wang *et al*., 2007; Srivatsan *et al*., 2010). Nucleotide arrangement in the chromosome therefore appears to play important roles in the replication and the processes of cell division. However, not every inversion in the bacterial chromosomes negatively affects the cells. Interreplichore inversions, which remain the approximate equality of replichore lengths, are in fact a common characteristic of the bacterial chromosomes (**Fig. 4**). It has been proposed that inversions are a major factor that makes the whole strand of chromosome symmetric (Albrecht-Buehler, 2006; Okamura *et al*., 2007). Nevertheless, biological significance of the natural inversions in bacterial chromosomes is not clear.

During evolution, the bacterial chromosomes can undergo gain or loss of DNA and subsequent rearrangement. As the two replichores are different in DNA composition and the whole chromosomal strand is symmetric, the nucleotide rearrangement in chromosome may be directed. In fact, attempts in addition of DNA into a chromosome are not always successful. When trying to insert the whole *Synechocystis* PCC6803 chromosome into the *B. subtilis* chromosome, Itaya *et al*. (2005) were successful only when the combined chromosome was symmetric with respect to the replication origin and terminus. Similarly, chromosomes of *E. coli* and *Haemophilus influenzae* were also successfully combined. However, difficulties existed in gene transfer, DNA replication, interactions in gene regulation and compatibility of gene products of the two bacteria. In addition, a number of *H. influenzae* genes were completely incompatible to *E. coli* (Holt *et al*., 2007). In a different study, an artificial bacterial chromosome was also successfully created but its genes were the same in sequence as well as in order as in the natural chromosome (Gibson *et al*., 2008). Strand symmetric and asymmetric properties of the bacterial chromosome, which are consequences of the chromosomal nucleotide arrangement, appear to play an important role in biological processes regarding the replication rate and the stability of bacterial chromosome. Understanding the relationship between these strand symmetry and asymmetry would provide a

basic background for making an artificial chromosome successfully without being dependent on nucleotide sequence of the naturally existing chromosome.

2. CONCLUSION

We have presented some general emerging properties of the bacterial chromosomal sequences with related issues involved in the nucleotide arrangement in chromosome. The nucleotide arrangement in the bacterial chromosomes has been shown to be involved in the replication, transcription and cell division, which in turn influence the effectiveness of the chromosomal genetic engineering processes. We emphasize that a good understanding of nucleotide arrangement in the bacterial chromosomes would help in experimental design in order to make the chromosomal engineering experiments predictable and effective. Therefore, there is a need to figure out the nucleotide arrangement in the bacterial chromosomes.

3. REFERENCES

Albrecht-Buehler, G., 2006. Asymptotically increasing compliance of genomes with Chargaff's second parity rules through inversions and inverted transpositions. Proc. Natl. Acad. Sci. USA., 103: 17828-17833. DOI: 10.1073/pnas.0605553103

Baisnee, P.F., S. Hampson and P. Baldi, 2002. Why are complementary DNA strands symmetric? Bioinformatics, 18: 1021-1033. PMID: 12176825

Bell, S.J. and D.R. Forsdyke, 1999. Deviations from Chargaff's second parity rule correlate with direction of transcription. J. Theor. Biol., 197: 63-76. PMID: 10036208

Chargaff, E., 1950. Chemical specificity of nucleic acids and mechanism of their enzymatic degradation. Experientia, 6: 201-209. DOI: 10.1007/BF02173653

Ellis, H.M., D. Yu, T. DiTizio and D.L. Court, 2001. High efficiency mutagenesis, repair and engineering of chromosomal DNA using single- stranded oligonucleotides. Proc. Natl. Acad. Sci. U.S.A., 98: 6742-6746. DOI: 10.1073/pnas.121164898

Esnault, E., M. Valens, O. Espeli and F. Boccard, 2007. Chromosome structuring limits genome plasticity in *Escherichia coli*. PLoS Genet., 3: e226-e226. DOI: 10.1371/journal.pgen.0030226

Forsdyke, D.R. and J.R. Mortimer, 2000. Chargaff's legacy. Gene, 261: 127-137.

Freeman, J.M., T.N. Plasterer, T.F. Smith and S.C. Mohr, 1998. Patterns of genome organization in bacteria. Science, 279: 1827-1827. DOI: 10.1126/science.279.5358.1827a

Gibson, D.G., G.A. Benders, C. Andrews-Pfannkoch, E.A. Denisova and H. Baden-Tillson *et al*., 2008. Complete chemical synthesis, assembly and cloning of a *Mycoplasma genitalium* genome. Science, 319: 1215-1220. DOI: 10.1126/science.1151721

Glover, B.P. and C.S. McHenry, 2001. The DNA polymerase III holoenzyme: An asymmetric dimeric replicative complex with leading and lagging strand polymerases. Cell, 105: 925-934.

Green, P., B. Ewing, W. Miller, P.J. Thomas and E.D. Green, 2003. Transcription-associated mutational asymmetry in mammalian evolution. Nat. Genet., 33: 514-517. DOI: 10.1038/ng1103

Holt, R.A., R. Warren, S. Flibotte, P.I. Missirlis and D.E. Smailus, 2007. Rebuilding microbial genomes. Bioessays, 29: 580-590. DOI: 10.1002/bies.20585

Huen, M.S.Y., X.T. Li, L.Y. Lu, R.M. Watt and D.P. Liu *et al*., 2006. The involvement of replication in single stranded oligonucleotide-mediated gene repair. Nucl. Acids Res., 34: 6183-6194. DOI: 10.1093/nar/gkl852

Itaya, M., K. Tsuge, M. Koizumi and K. Fujita, 2005. Combining two genomes in one cell: Stable cloning of the *Synechocystis* PCC6803 genome in the *Bacillus subtilis* 168 genome. Proc. Natl. Acad. Sci. U.S.A., 44: 15971-15976. DOI: 10.1073/pnas.0503868102

Kurtz, S., A. Phillippy, A.L. Delcher, M. Smoot and M. Shumway *et al*., 2004. Versatile and open software for comparing large genomes. Genome Biol., 5: R12-R12. DOI: 10.1186/gb-2004-5-2-r12

Li, X., N. Costantino, L. Lu, D. Liu and R.M. Watt *et al*., 2003. Identification of factors influencing strand bias in oligonucleotide-mediated recombination in *Escherichia coli*. Nucl. Acids Res., 31: 6674-6687. DOI: 10.1093/nar/gkg844

Liu, L., M.C. Rice, M. Drury, S. Cheng and H. Gamper *et al*., 2002. Strand bias in targeted gene repair is influenced by transcriptional activity. Mol. Cell Biol., 22: 3852-3863. PMID: 11997519

Lobry, J.R and N. Sueoka, 2002. Asymmetric directional mutation pressures in bacteria. Genome Biol., 3: 0058.1-0058.14. DOI: 10.1186/gb-2002-3-10-research0058

Maki, H., S. Maki and A. Kornberg, 1988. DNA Polymerase III holoenzyme of *Escherichia coli*. IV. The holoenzyme is an asymmetric dimer with twin active sites. J. Biol. Chem., 263: 6570-6578. PMID: 3283128

Maki, S. and A. Kornberg, 1988. DNA polymerase III holoenzyme of *Escherichia coli*. III. Distinctive processive polymerases reconstituted from purified subunits. J. Biol. Chem., 263: 6561-6569. PMID: 3283127

McHenry, C.S., 2003. Chromosomal replicases as asymmetric dimers: Studies of subunit arrangement and functional consequences. Mol. Microbiol., 49: 1157-1165. DOI: 10.1046/j.1365-2958.2003.03645.x

Mradzek, J. and S. Karlin, 1998. Strand compositional asymmetry in bacterial and large viral genomes. Proc. Natl. Acad. Sci. U.S.A., 95: 3720-3725.

Niki, H., Y. Yamaichi and S. Hiraga, 1999. Dynamic organization of chromosomal DNA in *Escherichia coli*. Gen. Dev., 14: 212-223. DOI: 10.1101/gad.14.2.212

Okamura, K., J. Wei and S.W. Scherer, 2007. Evolutionary implications of inversions that have caused intra-strand parity in DNA. BMC Genomics, 8: 160-160. DOI: 10.1186/1471-2164-8-160

Polak, P. and P.F. Arndt, 2008. Transcription induces strand-specific mutations at the 5' end of human genes. Genome Res., 18: 1216-1223. DOI: 10.1101/gr.076570.108

Poptsova, M.S., S.A. Larionov, E.V. Ryadchenko, S.D. Rybalko, I.A. Zakharov *et al*., 2009. Hidden chromosome symmetry: in silico transformation reveals symmetry in 2D DNA walk trajectories of 671 chromosomes. PLoS ONE, 4: e6396-e6396. DOI: 10.1371/journal.pone.0006396

Powdel, B.R., S.S. Satapathy, A. Kumar, P.K. Jha and A.K. Buragohain *et al*., 2009. A study in entire chromosomes of violations of the intra-strand parity of complementary nucleotides (Chargaff's second parity rule). DNA Res., 16: 325-343. DOI: 10.1093/dnares/dsp021

Qi, D. and A.J. Cuticchia, 2001. Compositional symmetries in complete genomes. Bioinformatics, 17: 557-559. PMID: 11395434

Rocha, E.P.C., 2004. The replication-related organization of bacterial genomes. Microbiology, 150: 1609-1627. DOI: 10.1099/mic.0.26974-0

Rudner, R., J.D. Karkas and E. Chargaff, 1968. Separation of *B. subtilis* DNA into complementary strands. III. Direct analysis. Proc. Natl. Acad. Sci. U.S.A., 60: 921-922. PMID: 4970114

Salzberg, S.L., A.J. Salzberg, A.R. Kerlavage and J.F. Tomb, 1998. Skewed oligomers and origins of replication. Gene, 217: 57-67.

Srivatsan, A., A. Tehranchi, D.M. MacAlpine and J.D. Wang, 2010. Co-orientation of replication and transcription preserves genome integrity. PLoS Genet., 6: e1000810-e1000810. DOI: 10.1371/journal.pgen.1000810

Thiel, A., M. Valens, I. Vallet-Gely, O. Espeli and F. Boccard, 2012. Long-range chromosome organization in *E. coli*: A site-specific system isolates the Termacrodomain. PLoS Genet., 8: e1002672-e1002672. DOI: 10.1371/journal.pgen.1002672

Touchon, M., S. Nicolay, B. Audit, E.B. Brodie and Y. d'Aubenton-Carafa *et al*., 2005. Replication-associated strand asymmetries in mammalian genomes: Toward detection of replication origins. Proc. Natl. Acad. Sci. U.S.A., 102: 9836-9841. DOI: 10.1073/pnas.0500577102

Valens, M., S. Penaud, M. Rossignol, F. Cornet and F. Boccard, 2004. Macrodomain organization of the *Escherichia coli* chromosome. EMBO J., 23: 4330-4341. DOI: 10.1038/sj.emboj.7600434

Wang, J.D., M.B. Berkmen and A.D. Grossman, 2007. Genome-wide coorientation of replication and transcription reduces adverse effects on replication in *Bacillus subtilis*. Proc. Natl. Acad. Sci. U.S.A., 104: 5608-5613. DOI: 10.1073/pnas.0608999104

Worning, P., L.J. Jensen, P.F. Hallin, H.H. Staerfeldt and D.W. Ussery, 2006. Origin of replication in circular prokaryotic chromosomes. Environ. Microbiol., 8: 353-361. DOI: 10.1111/j.1462-2920.2005.00917x