

# Genome DNA Sequence Variation, Evolution, and Function in Bacteria and Archaea

Hiromi Nishida\*

Agricultural Bioinformatics Research Unit,  
Graduate School of Agricultural and Life Sciences,  
University of Tokyo, Japan

## Abstract

Comparative genomics has revealed that variations in bacterial and archaeal genome DNA sequences cannot be explained by only neutral mutations. Virus resistance and plasmid distribution systems have resulted in changes in bacterial and archaeal genome sequences during evolution. The restriction-modification system, a virus resistance system, leads to avoidance of palindromic DNA sequences in genomes. Clustered, regularly interspaced, short palindromic repeats (CRISPRs) found in genomes represent yet another virus resistance system. Comparative genomics has shown that bacteria and archaea have failed to gain any DNA with GC content higher than the GC content of their chromosomes. Thus, horizontally transferred DNA regions have lower GC content than the host chromosomal DNA does. Some nucleoid-associated proteins bind DNA regions with low GC content and inhibit the expression of genes contained in those regions. This

form of gene repression is another type of virus resistance system. On the other hand, bacteria and archaea have used plasmids to gain additional genes. Virus resistance systems influence plasmid distribution. Interestingly, the restriction-modification system and nucleoid-associated protein genes have been distributed via plasmids. Thus, GC content and genomic signatures do not reflect bacterial and archaeal evolutionary relationships.

## Distribution of genome base compositions and mutational biases

Among all published genome sequences, *Candidatus Zinderia insecticola* has a genome with the lowest guanine-cytosine (GC) content (13.5%) (McCutcheon and Moran, 2010), and *Anaeromyxobacter dehalogenans* 2CP-C has a genome with the highest GC content (74.9%) (Thomas et al., 2008). The distribution of GC content of bacterial genomes is rather distinct from a normal (Gaussian) distribution (Figure 1). On the other hand, within each bacterium, the distribution of the GC content of the genes is similar to a normal distribution (Figure 2), suggesting that each bacterium has maintained its genomic GC content. Bacteria have been thought to possess directionality, driven by neutral forces, toward higher or lower levels of GC

\*Corresponding author: [ahn@mail.ecc.u-tokyo.ac.jp](mailto:ahn@mail.ecc.u-tokyo.ac.jp)

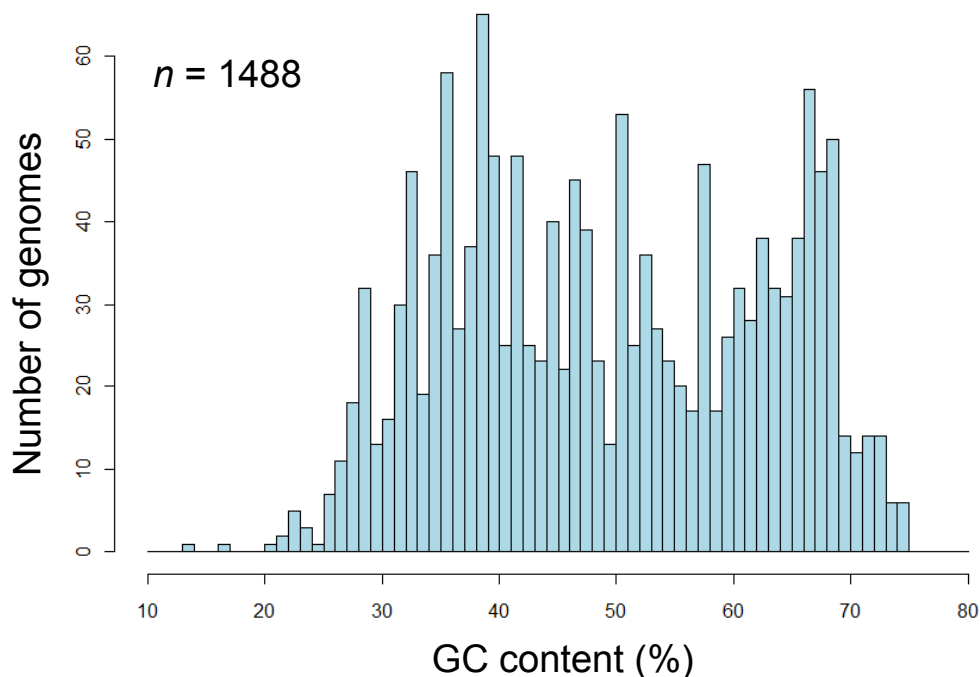


Figure 1. Distribution of GC content of bacterial genomes.

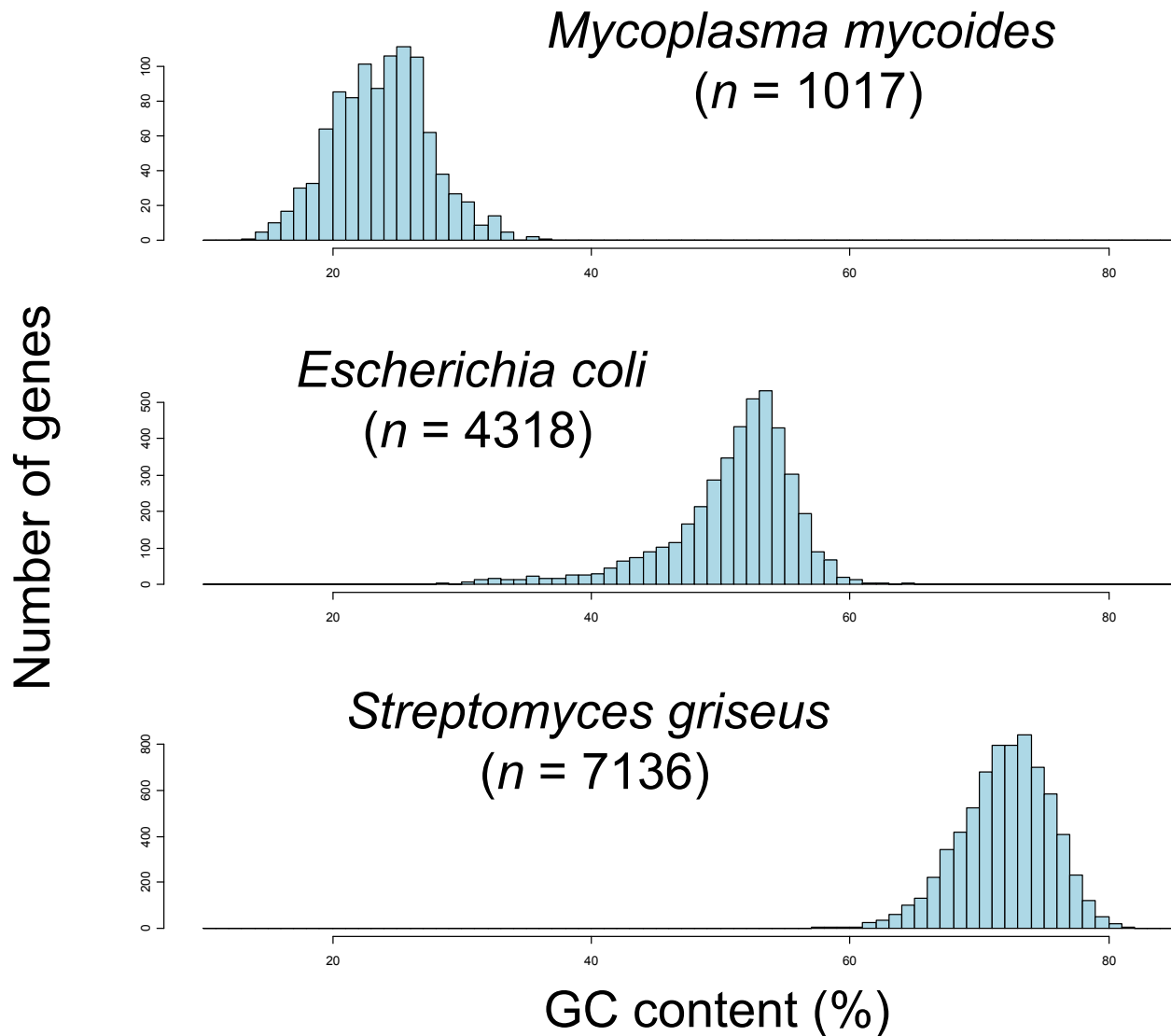


Figure 2. Distribution of GC content of genes of three bacteria species, *Mycoplasma mycoides* (genome with low GC content), *Escherichia coli* (genome with middle GC content), and *Streptomyces griseus* (genome with high GC content).

content in their DNA (Sueoka, 1961; Freese, 1962; Sueoka, 1962; Sueoka, 1988).

However, recent studies have shown that the variation of the GC content among bacteria is driven by selection, in which mutations from GC to adenosine-thymine (AT) are more common than mutations from AT to GC (Hershberg and Petrov, 2010; Hildebrand et al., 2010; Rocha and Feil, 2010). Lind and Andersson (2008) compared the genomes of 2 *Salmonella typhimurium* mutants and showed a bias toward mutations from GC to AT. Rocha and Danchin (2002) suggested that GC content variation may be related to the higher energy cost and limited availability of G and C over A and T. However, many bacterial species such as Actinobacteria have a high GC content genome. How have these bacteria maintained a high GC content? DNA polymerase components involved in DNA replication have been reported to directly influence the GC contents of genomes (Zhao et al., 2007; Wu et al., 2012).

#### GC content and genomic signature

Oligonucleotide frequencies (genomic signatures) within a genome can be observed and compared with other genomes (Campbell et al., 1999; Deschavanne et al., 1999). Undoubtedly, genomes with similar genomic signatures have similar GC contents. Interestingly, genomes with similar GC contents also have similar genomic signatures (Albrecht-Buehler, 2007a; Albrecht-Buehler, 2007b; Zhang and Wang, 2011), with the exception of *Deinococcus radiodurans* and *Thermus thermophilus* (Nishida et al., 2012a). Phylogenetic relationships based on genomic signature comparison of 89 bacteria (Nishida et al., 2012a) were found to be completely different from those based on gene content or orthologous protein sequence comparison (Nishida et al., 2011). This indicates that organisms with genomic signature similarity do not represent closely related organisms in evolutionary terms (Albrecht-Buehler, 2007a; Bohlin, 2011).

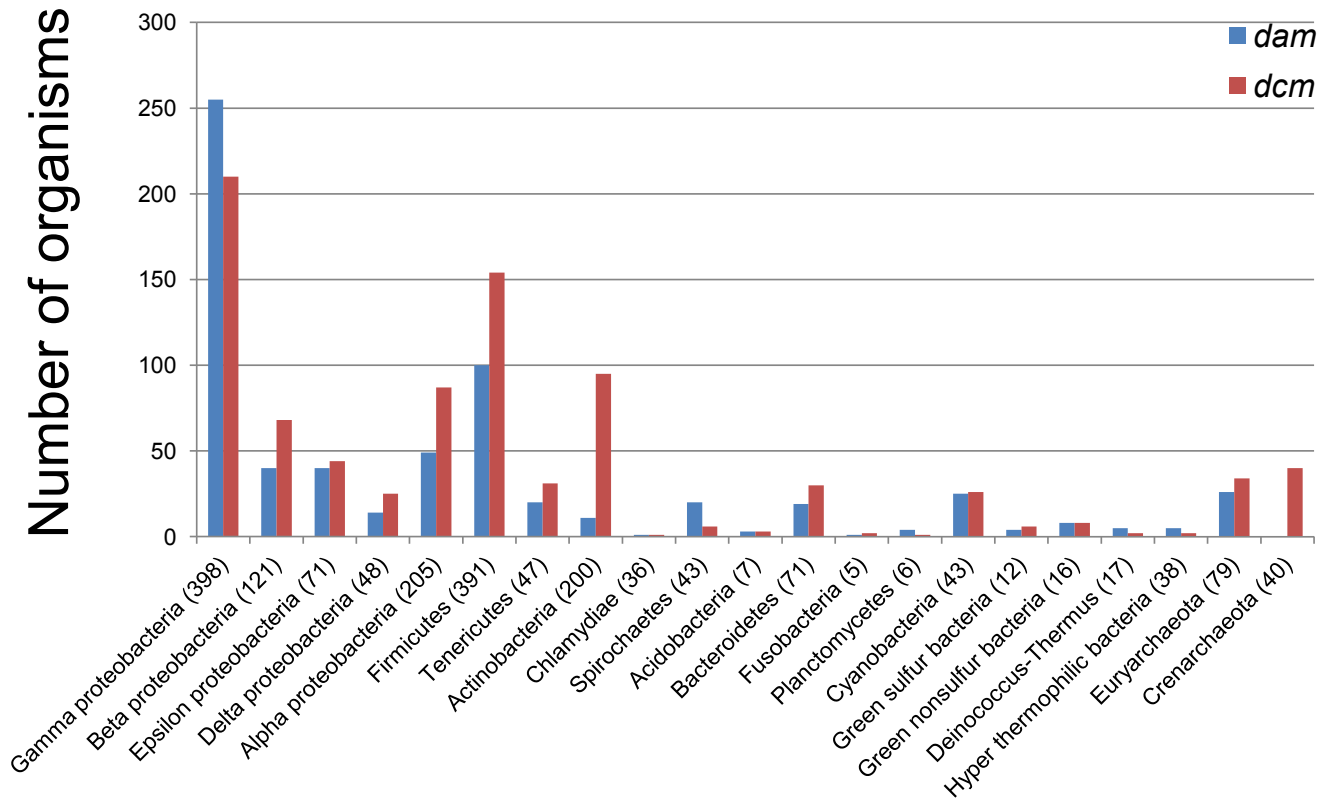


Figure 3. Distribution of homologs of DNA adenine methylase gene (*dam*) and DNA cytosine methylase gene (*dcm*) of *Escherichia coli*. The numbers in parentheses indicate the numbers of organisms analyzed in this study. Taxonomic classification is according to the KEGG (Kanehisa et al., 2012).

In addition, frequencies of palindromic DNA sequence patterns are significantly lower than those of non-palindromic sequence patterns in bacterial and archaeal genomes (Gelfand and Koonin, 1997). Palindrome avoidance has been reported to be intimately correlated with infective behavior of the bacteriophage (Rocha et al., 2001). The low frequency of palindromic sequence patterns has been found in not only single genome sequence but also metagenomic sequence data (Dick et al., 2009). Generally, restriction enzymes recognize palindromic DNA sequences and digest these regions. However, when palindromic DNA sequences are methylated by a DNA methylase, restriction enzymes can no longer digest them (Wilson and Murray, 1991; Bickle and Krüger, 1993). Bacteria protect their palindromic DNAs modified by modification enzymes, but digest bacteriophage palindromic DNAs that are not modified. Thus, the restriction-modification system functions as a virus resistance system (Kobayashi, 2001; Labrie et al., 2010). Palindrome avoidance influences genomic signatures.

Surprisingly, bacteria and archaea possess clustered, regularly interspaced, short palindromic repeats (CRISPRs) and the CRISPR-associated (*cas*) genes as a virus resistance system, which acts as a defense system against viral infections through the use of CRISPR RNA transcripts (Barrangou et al., 2007; Brouns et al., 2008; Marraffini and Sontheimer, 2008; Sorek et al., 2008; Karginov and Hannon, 2010; Labrie et al., 2010). Thus, bacteria and

archaea employ palindromic DNA sequence avoidance as well as palindromic DNA sequences in CRISPR regions, as a defense system against viral infections.

Restriction-modification systems may also influence GC content and genomic signatures. DNA methylase distributions among bacteria and archaea were shown in Figure 3. *Escherichia coli* has DNA adenine methylase gene (*dam*; NCBI gene ID, 947893) and DNA cytosine methylase gene (*dcm*; NCBI gene ID, 946479) (Marinus, 1987). Based on the Kyoto Encyclopedia of Genes and Genomes (KEGG) orthology database (Kanehisa et al., 2012), the distributions of *dam* and *dcm* differ (Figure 3). The differences in their distributions suggest that genomic GC content is related to DNA methylase distribution. For example, most Actinobacteria (genomes with high GC content) lack *dam*, and most Spirochaetes (genomes with low GC content) lack *dcm*. However, certain cases cannot be explained by genome GC content. Most Tenericutes (genomes with low GC content) have both *dam* and *dcm*. In addition, although the GC contents of Crenarchaeota are not high (35–60% among 40 archaea used in this study), all Crenarchaeota possess *dcm* and lack *dam*. On the other hand, restriction enzymes have no structural similarity with DNA methylases or other restriction enzymes (Bickle and Krüger, 1993). It was reported that the restriction-modification system is a mobile element (Kusano et al., 1995; Naito et al., 1995; Kobayashi, 2001).

Genomic GC content influences codon and amino acid usages (Muto and Osawa, 1987; Lobry, 1997; Singer and Hickey, 2000; Knight et al., 2001; Chen et al., 2004; Wan et al., 2004; Lightfield et al., 2011; Schmidt et al., 2012). In addition, the GC contents of horizontally transferred DNA regions have been ameliorated to adjust to host chromosome GC content (Lawrence and Ochman, 1997). It strongly suggests that each bacterium or archaeon has a system for maintaining organism-specific GC content and genomic signature.

### Differences in GC contents between host genome DNA and horizontally transferred DNA regions

Genome size and GC content are weakly correlated in bacteria and archaea (Bentley and Parkhill, 2004; Musto et al., 2006; Mitchell, 2007; Suzuki et al., 2008; Guo et al., 2009; Nishida, 2012). The genomes of obligate host-associated bacteria are short and low GC content (Moran, 2002; Klasson and Andersson, 2004; McCutcheon and Moran, 2012), with exception of *Candidatus Hodgkinia cicadicola* (McCutcheon et al., 2009; Van Leuven and McCutcheon, 2012). In addition, horizontally transferred DNA, plasmid DNA, and virus DNA have lower GC content than host chromosome DNAs do (Rocha and Danchin, 2002). Most of the differences in GC content between plasmids and their host chromosomes are of less than 10% (Nishida, 2012), suggesting that host organisms cannot maintain and regulate plasmids with very different GC content from their own. If bacteria and archaea maintain lower GC content for horizontally transferred regions, this maintenance will compete with their amelioration of GC content to match the GC content of the host genome (Lawrence and Ochman, 1997).

Interestingly, bacteria and archaea have not acquired DNAs with a GC content higher than the GC content of their own genome. Bacterial and archaeal genomes with high AT content are protected from attacks by most viruses. On the other hand, it is difficult for those organisms to use any plasmids. The genome sizes of obligate host-associated bacteria are decreasing (for example, Oshima and Nishida, 2008). However, the genome size reduction is not limited in obligate host-associated bacteria (Nilsson et al., 2005). There is a general bias among bacteria toward genomic deletions rather than insertions (Mira et al., 2001). Plasmids play an important role in additional gene gain uptake into chromosomes (Davison, 1999; Harrison and Brockhurst, 2012). It is possible that obligate host-associated bacteria do not need additional gene uptake. It may therefore be hypothesized that these bacteria maintain a genome with low GC content as a virus resistance system.

### GC content and nucleoid-associated proteins

Nucleoid-associated proteins are related not only to nucleoid structures but also to gene regulation (Dillon and Dorman, 2010). The heat-stable (or histone-like) nucleoid-structural (H-NS) protein in *Salmonella enterica* binds DNA regions with low GC content rather than the remaining chromosomal DNA and inhibits expression of the genes contained in those regions, to which horizontally transferred DNA fragments locate (Lucchini et al., 2006; Navarre et al., 2006). Similar functions for nucleoid-associated proteins have been found in other bacteria (Castang et al., 2008; Gordon et al., 2010; Smits and Grossman, 2010; Yun et al.,

2010). These gene-silencing systems depend on the fact that horizontally transferred DNAs have lower GC content than host chromosome DNAs do (Rocha and Danchin, 2002).

This gene-repression system involving nucleoid-associated proteins is widespread amongst bacteria and archaea, suggesting that the nucleoid-associated proteins may bind to DNA regions with different GC content between different bacterial or archaeal species. For example, in the *Symbiobacterium thermophilus* genome with high GC content (69%), transposase genes, markers of transposable genetic elements, are more frequently found in regions with lower GC content (less than 65% GC content) than in the remaining chromosomal DNA (Nishida and Yun, 2011). Interestingly, nucleoid-associated protein genes are distributed not only throughout bacterial chromosomes but also within plasmids, suggesting that plasmids have carried these genes (Yun et al., 2010; Takeda et al., 2011).

Although nucleoid-associated proteins have different structures, they share the same function (Gordon et al., 2011). Interestingly, core histones, which are structurally different from bacterial nucleoid-associated proteins, prefer AT-rich DNA to GC-rich DNA. This DNA sequence preference plays an important role in nucleosome formation (Segal et al., 2006; Segal and Widom, 2009; Valouev et al., 2011; Nishida et al., 2012b). The interactions between DNA sequence preferences and nucleoid-associated proteins may have played an important role in global regulations of genes among Bacteria, Archaea, and Eukarya during evolution.

### Acknowledgements

I thank Professor Teruhiko Beppu for his valuable comments and encouragements.

### References

- Albrecht-Buehler, G. (2007a). The three classes of triplet profiles of natural genomes. *Genomics* 89, 596-601.
- Albrecht-Buehler, G. (2007b). Inversions and inverted transpositions as the basis for an almost universal "format" of genome sequences. *Genomics* 90, 297-305.
- Barrangou, R., Fremaux, C., Deveau, H., Richards, M., Boyaval, P., Moineau, S., Romero, D.A., and Horvath, P. (2007). CRISPR provides acquired resistance against viruses in prokaryotes. *Science* 315, 1709-1712.
- Bentley, S.D. and Parkhill, J. (2004). Comparative genomic structure of prokaryotes. *Ann. Rev. Genet.* 38, 771-791.
- Bickle, T.A. and Krüger, D.H. (1993). Biology of DNA restriction. *Microbiol. Rev.* 57, 434-450.
- Bohlin, J. (2011). Genome signatures in microbes – properties and applications. *Scientific World J.* 11, 715-725.
- Brouns, S.J., Jore, M.M., Lundgren, M., Westra, E.R., Slijkhuis, R.J., Snijders, A.P., Dickman, M.J., Makarova, K.S., Koonin, E.V., and van der Oost, J. (2008). Small CRISPR RNAs guide antiviral defense in prokaryotes. *Science* 321, 960-964.
- Campbell, A., Mrázek, J., and Karlin, S. (1999). Genome signature comparisons among prokaryote, plasmid, and mitochondrial DNA. *Proc. Natl. Acad. Sci. USA* 96, 9184-9189.
- Castang, S., McManus, H.R., Turner, K.H., and Dove, S.L. (2008). H-NS family members function coordinately in an

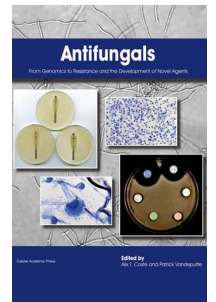
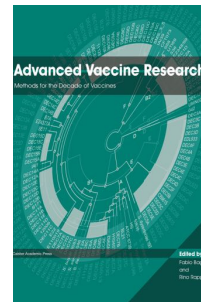
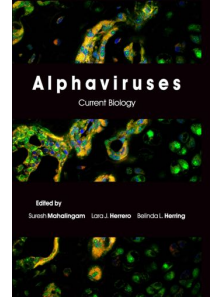
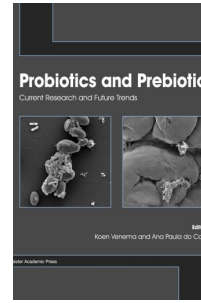
- opportunistic pathogen. *Proc. Natl. Acad. Sci. USA* *105*, 18947–18952.
- Chen, S.L., Lee, W., Hottes, A.K., Shapiro, L., and McAdams, H.H. (2004). Codon usage between genomes is constrained by genome-wide mutational processes. *Proc. Natl. Acad. Sci. USA* *101*, 3480–3485.
- Davison, J. (1999). Genetic exchange between bacteria in the environment. *Plasmid* *42*, 73–91.
- Deschavanne, P.J., Giron, A., Vilain, J., Fagot, G., and Fertil, B. (1999). Genomic signature: characterization and classification of species assessed by chaos game representation of sequences. *Mol. Biol. Evol.* *16*, 1391–1399.
- Dick, G.J., Andersson, A.F., Baker, B.J., Simmons, S.L., Thomas, B.C., Yelton, A.P., and Banfield, J.F. (2009). Community-wide analysis of microbial genome sequence signatures. *Genome Biol.* *10*, R85.
- Dillon, S.C. and Dorman, C.J. (2010). Bacterial nucleoid-associated proteins, nucleoid structure and gene expression. *Nat. Rev. Microbiol.* *8*, 185–195.
- Freese, E. (1962). On the evolution of the base composition of DNA. *J. Theoret. Biol.* *3*, 82–101.
- Gelfand, M.S. and Koonin, E.V. (1997). Avoidance of palindromic words in bacterial and archaeal genomes: a close connection with restriction enzymes. *Nucleic Acids Res.* *25*, 2430–2439.
- Gordon, B.R.G., Li, Y., Cote, A., Weirauch, M.T., Ding, P., Hughes, T.R., Navarre, W.W., Xia, B., and Liu, J. (2011). Structural basis for recognition of AT-rich DNA by unrelated exogenous silencing proteins. *Proc. Natl. Acad. Sci. USA* *108*, 10690–10695.
- Gordon, B.R.G., Li, Y., Wang, L., Sintsova, A., van Bakel, H., Tian, S., Navarre, W.W., Xia, B., and Liu, J. (2010). Lsr2 is a nucleoid-associated protein that targets AT-rich sequences and virulence genes in *Mycobacterium tuberculosis*. *Proc. Natl. Acad. Sci. USA* *107*, 5154–5159.
- Guo, F.B., Lin, H., and Huang, J. (2009). A plot of G + C content against sequence length of 640 bacterial chromosomes shows the points are widely scattered in the upper triangular area. *Chromosome Res.* *17*, 359–364.
- Harrison, E. and Brockhurst, M.A. (2012). Plasmid-mediated horizontal gene transfer is a coevolutionary process. *Trends Microbiol.* *20*, 262–267.
- Hershberg, R. and Petrov, D.A. (2010). Evidence that mutation is universally biased towards AT in bacteria. *PLoS Genet.* *6*, e1001115.
- Hildebrand, F., Meyer, A., and Eyre-Walker, A. (2010). Evidence of selection upon genomic GC-content in bacteria. *PLoS Genet.* *6*, e1001107.
- Kanehisa, M., Goto, S., Sato, Y., Furumichi, M., and Tanabe, M. (2012). KEGG for integration and interpretation of large-scale molecular datasets. *Nucleic Acids Res.* *40*, D109–D114.
- Karginov, F.V. and Hannon, G.J. (2010). The CRISPR systems: small RNA-guided defense in bacteria and archaea. *Mol. Cell* *37*, 7–19.
- Klasson, L. and Andersson, S.G.E. (2004). Evolution of minimal-gene-sets in host-dependent bacteria. *Trends Microbiol.* *12*, 37–43.
- Knight, R.D., Freeland, S.J., and Landweber, L.F. (2001). A simple model based on mutation and selection explains trends in codon and amino-acid usage and GC composition within and across genomes. *Genome Biol.* *2*, research0010.1–13.
- Kobayashi, I. (2001). Behavior of restriction-modification systems as selfish mobile elements and their impact on genome evolution. *Nucleic Acids Res.* *29*, 3742–3756.
- Kusano, K., Naito, T., Hanada, N., and Kobayashi, I. (1995). Restriction-modification systems as genomic parasites in competition for specific sequences. *Proc. Natl. Acad. Sci. USA* *92*, 11095–11099.
- Labrie, S.J., Samson, J.E., and Moineau, S. (2010). Bacteriophage resistance mechanisms. *Nat. Rev. Microbiol.* *8*, 317–327.
- Lawrence, J.G. and Ochman, H. (1997). Amelioration of bacterial genomes: rates of change and exchange. *J. Mol. Evol.* *44*, 383–397.
- Lightfield, J., Fram, N.R., and Ely, B. (2011). Across bacterial phyla, distantly-related genomes with similar genomic GC content have similar patterns of amino acid usage. *PLoS ONE* *6*, e17677.
- Lind, P.A. and Andersson, D.I. (2008). Whole-genome mutational biases in bacteria. *Proc. Natl. Acad. Sci. USA* *105*, 17878–17883.
- Lobry, J.R. (1997). Influence of genomic G + C content on average amino-acid composition of proteins from 59 bacterial species. *Gene* *205*, 309–316.
- Lucchini, S., Rowley, G., Goldberg, M.D., Hurd, D., Harrison, M., and Hinton, J.C. (2006). H-NS mediates the silencing of laterally acquired genes in bacteria. *PLoS Pathog.* *2*, e81.
- Marinus, M.G. (1987). DNA methylation in *Escherichia coli*. *Ann. Rev. Genet.* *21*, 113–131.
- Marraffini, L.A. and Sontheimer, E.J. (2008). CRISPR interference limits horizontal gene transfer in *Staphylococci* by targeting DNA. *Science* *322*, 1843–1845.
- McCutcheon, J.P., McDonald, B.R., and Moran, N.A. (2009). Origin of an alternative genetic code in the extremely small and GC-rich genome of a bacterial symbiont. *PLoS Genet.* *5*, e1000565.
- McCutcheon, J.P. and Moran, N.A. (2010). Functional convergence in reduced genomes of bacterial symbionts spanning 200 My of evolution. *Genome Biol. Evol.* *2*, 708–718.
- McCutcheon, J.P. and Moran, N.A. (2012). Extreme genome reduction in symbiotic bacteria. *Nat. Rev. Microbiol.* *10*, 13–26.
- Mitchell, D. (2007). GC content and genome length in Chargaff compliant genomes. *Biochem. Biophys. Res. Commun.* *353*, 207–210.
- Mira, A., Ochman, H., and Moran, N.A. (2001). Deletional bias and the evolution of bacterial genomes. *Trends Genet.* *17*, 589–596.
- Moran, N.A. (2002). Microbial minimalism: genome reduction in bacterial pathogens. *Cell* *108*, 583–586.
- Musto, H., Naya, H., Zavala, A., Romero, H., Alvarez-Valín, F., and Bernardi, G. (2006). Genomic GC level, optimal growth temperature, and genome size in prokaryotes. *Biochem. Biophys. Res. Commun.* *347*, 1–3.
- Muto, A. and Osawa, S. (1987). The guanine and cytosine content of genomic DNA and bacterial evolution. *Proc. Natl. Acad. Sci. USA* *84*, 166–169.
- Naito, T., Kusano, K., and Kobayashi, I. (1995). Selfish behavior of restriction-modification systems. *Science* *267*, 897–899.

- Navarre, W.W., Porwollik, S., Wang, Y., McClelland, M., Rosen, H., Libby, S.J., and Fang, F.C. (2006). Selective silencing of foreign DNA with low GC content by the H-NS protein in *Salmonella*. *Science* 313, 236-238.
- Nilsson, A.I., Koskiniemi, S., Eriksson, S., Kugelberg, E., Hinton, J.C.D., and Andersson, D.I. (2005). Bacterial genome size reduction by experimental evolution. *Proc. Natl. Acad. Sci. USA* 102, 12112-12116.
- Nishida, H. (2012). Comparative analyses of base compositions, DNA sizes, and dinucleotide frequency profiles in archaeal and bacterial chromosomes and plasmids. *Int. J. Evol. Biol.* 2012, 342482.
- Nishida, H., Abe, R., Nagayama, T., and Yano, K. (2012a). Genome signature difference between *Deinococcus radiodurans* and *Thermus thermophilus*. *Int. J. Evol. Biol.* 2012, 205274.
- Nishida, H., Beppu, T., and Ueda, K. (2011). Whole-genome comparison clarifies close phylogenetic relationships between the phyla Dictyoglomi and Thermotogae. *Genomics* 98, 370-375.
- Nishida, H., Kondo, S., Matsumoto, T., Suzuki, Y., Yoshikawa, H., Taylor, T.D., and Sugiyama, J. (2012b). Characteristics of nucleosomes and linker DNA regions on the genome of the basidiomycete *Mixia osmundae* revealed by mono- and dinucleosome mapping. *Open Biol.* 2, 120043.
- Nishida, H. and Yun, C.-S. (2011). Phylogenetic and guanine-cytosine content analysis of *Symbiobacterium thermophilum* genes. *Int. J. Evol. Biol.* 2011, 634505.
- Oshima, K. and Nishida, H. (2008). Deletion of the genes evolving under *Ureaplasma*-specific selection. *J. Mol. Evol.* 66, 529-532.
- Rocha, E.P. and Danchin, A. (2002). Base composition bias might result from competition for metabolic resources. *Trends Genet.* 18, 291-294.
- Rocha, E.P., Danchin, A., and Viari, A. (2001). Evolutionary role of restriction/modification systems as revealed by comparative genome analysis. *Genome Res.* 11, 946-958.
- Rocha, E.P. and Feil, E.J. (2010). Mutational patterns cannot explain genome composition: Are there any neutral sites in the genomes of bacteria? *PLoS Genet.* 6, e1001104.
- Schmidt, A., Rzanny, M., Schmidt, A., Hagen, M., Schütze, E., and Kothe, E. (2012). GC content-independent amino acid patterns in Bacteria and Archaea. *J. Basic Microbiol.* 52, 195-205.
- Segal, E., Fondufe-Mittendorf, Y., Chen, L., Thåström, A., Field, Y., Moore, I.K., Wang, J.P., and Widom, J. (2006). A genomic code for nucleosome positioning. *Nature* 442, 772-778.
- Segal, E. and Widom, J. (2009). What controls nucleosome positions? *Trends Genet.* 25, 335-343.
- Singer, G.A. and Hickey, D.A. (2000). Nucleotide bias causes a genomewide bias in the amino acid composition of proteins. *Mol. Biol. Evol.* 17, 1581-1588.
- Smits, W.K. and Grossman, A.D. (2010). The transcriptional regulator Rok binds A+T-rich DNA and is involved in repression of a mobile genetic element in *Bacillus subtilis*. *PLoS Genet.* 6, e1001207.
- Sorek, R., Kunin, V., and Hugenholtz, P. (2008). CRISPR – a widespread system that provides acquired resistance against phages in bacteria and archaea. *Nat. Rev. Microbiol.* 6, 181-186.
- Sueoka, N. (1961). Correlation between base composition of deoxyribonucleic acid and amino acid composition of protein. *Proc. Natl. Acad. Sci. USA* 47, 1141-1149.
- Sueoka, N. (1962). On the genetic basis of variation and heterogeneity of DNA base composition. *Proc. Natl. Acad. Sci. USA* 48, 582-592.
- Sueoka, N. (1988). Directional mutation pressure and neutral molecular evolution. *Proc. Natl. Acad. Sci. USA* 85, 2653-2657.
- Suzuki, H., Sota, M., Brown, C.J., and Top, E.M. (2008). Using Mahalanobis distance to compare genomic signatures between bacterial plasmids and chromosomes. *Nucleic Acids Res.* 36, e147.
- Takeda, T., Yun, C.-S., Shintani, M., Yamane, H., and Nojiri, H. (2011). Distribution of genes encoding nucleotide-associated protein homologs in plasmids. *Int. J. Evol. Biol.* 2011, 685015.
- Thomas, S.H., Wagner, R.D., Arakaki, A.K., Skolnick, J., Kirby, J.R., Shimkets, L.J., Sanford, R.A., and Löffler, F.E. (2008). The mosaic genome of *Anaeromyxobacter dehalogenans* strain 2CP-C suggests an aerobic common ancestor to the delta-proteobacteria. *PLoS ONE* 3, e2103.
- Valouev, A., Johnson, S.M., Boyd, S.D., Smith, C.L., Fire, A.Z., and Sidow, A. (2011). Determinants of nucleosome organization in primary human cells. *Nature* 474, 516-520.
- Van Leuven, J.T. and McCutcheon, J.P. (2012). An AT mutational bias in the tiny GC-rich endosymbiont genome of *Hodgkinia*. *Genome Biol. Evol.* 4, 24-27.
- Wan, X.-F., Xu, D., Kleinhofs, A., and Zhou, J. (2004). Quantitative relationship between synonymous codon usage bias and GC composition across unicellular genomes. *BMC Evol. Biol.* 4, 19.
- Wilson, G.G. and Murray, N.E. (1991). Restriction and modification systems. *Ann. Rev. Genet.* 25, 585-627.
- Wu, H., Zhang, Z., Hu, S., and Yu, J. (2012). On the molecular mechanism of GC content variation among eubacterial genomes. *Biol. Direct* 7, 2.
- Yun, C.S., Suzuki, C., Naito, K., Takeda, T., Takahashi, Y., Sai, F., Terabayashi, T., Miyakoshi, M., Shintani, M., Nishida, H., Yamane, H., and Nojiri, H. (2010). Pmr, a histone-like protein H1 (H-NS) family protein encoded by the IncP-7 plasmid pCAR1, is a key global regulator that alters host function. *J. Bacteriol.* 192, 4720-4731.
- Zhang, S.-H. and Wang, L. (2011). A novel common triplet profile for GC-rich prokaryotic genomes. *Genomics* 97, 330-331.
- Zhao, X., Zhang, Z., Yan, J., and Yu, J. (2007). GC content variability of eubacteria is governed by the pol III alpha subunit. *Biochem. Biophys. Res. Commun.* 356, 20-25.

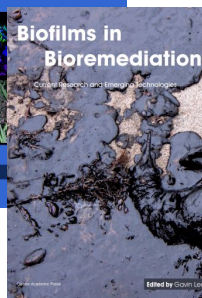
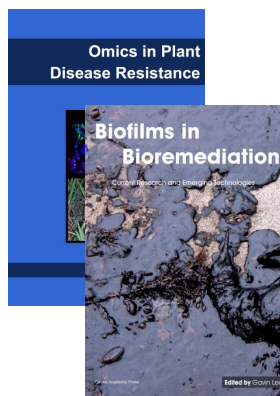
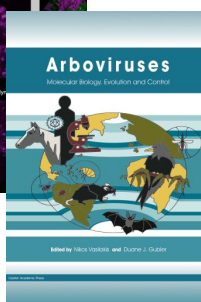
# Further Reading

**Caister Academic Press** is a leading academic publisher of advanced texts in microbiology, molecular biology and medical research. Full details of all our publications at [caister.com](http://www.caister.com)

- **MALDI-TOF Mass Spectrometry in Microbiology**  
Edited by: M Kostrzewa, S Schubert (2016)  
[www.caister.com/malditof](http://www.caister.com/malditof)
- **Aspergillus and Penicillium in the Post-genomic Era**  
Edited by: RP Vries, IB Gelber, MR Andersen (2016)  
[www.caister.com/aspergillus2](http://www.caister.com/aspergillus2)
- **The Bacteriocins: Current Knowledge and Future Prospects**  
Edited by: RL Dorit, SM Roy, MA Riley (2016)  
[www.caister.com/bacteriocins](http://www.caister.com/bacteriocins)
- **Omics in Plant Disease Resistance**  
Edited by: V Bhaduria (2016)  
[www.caister.com/opdr](http://www.caister.com/opdr)
- **Acidophiles: Life in Extremely Acidic Environments**  
Edited by: R Quatrini, DB Johnson (2016)  
[www.caister.com/acidophiles](http://www.caister.com/acidophiles)
- **Climate Change and Microbial Ecology: Current Research and Future Trends**  
Edited by: J Marxsen (2016)  
[www.caister.com/climate](http://www.caister.com/climate)
- **Biofilms in Bioremediation: Current Research and Emerging Technologies**  
Edited by: G Lear (2016)  
[www.caister.com/biorem](http://www.caister.com/biorem)
- **Microalgae: Current Research and Applications**  
Edited by: MN Tsaloglou (2016)  
[www.caister.com/microalgae](http://www.caister.com/microalgae)
- **Gas Plasma Sterilization in Microbiology: Theory, Applications, Pitfalls and New Perspectives**  
Edited by: H Shintani, A Sakudo (2016)  
[www.caister.com/gasplasma](http://www.caister.com/gasplasma)
- **Virus Evolution: Current Research and Future Directions**  
Edited by: SC Weaver, M Denison, M Roossinck, et al. (2016)  
[www.caister.com/virusevol](http://www.caister.com/virusevol)
- **Arboviruses: Molecular Biology, Evolution and Control**  
Edited by: N Vasilakis, DJ Gubler (2016)  
[www.caister.com/arbo](http://www.caister.com/arbo)
- **Shigella: Molecular and Cellular Biology**  
Edited by: WD Picking, WL Picking (2016)  
[www.caister.com/shigella](http://www.caister.com/shigella)
- **Aquatic Biofilms: Ecology, Water Quality and Wastewater Treatment**  
Edited by: AM Romani, H Guasch, MD Balaguer (2016)  
[www.caister.com/aquaticbiofilms](http://www.caister.com/aquaticbiofilms)
- **Alphaviruses: Current Biology**  
Edited by: S Mahalingam, L Herrero, B Herring (2016)  
[www.caister.com/alpha](http://www.caister.com/alpha)
- **Thermophilic Microorganisms**  
Edited by: F Li (2015)  
[www.caister.com/thermophile](http://www.caister.com/thermophile)



- **Flow Cytometry in Microbiology: Technology and Applications**  
Edited by: MG Wilkinson (2015)  
[www.caister.com/flow](http://www.caister.com/flow)
- **Probiotics and Prebiotics: Current Research and Future Trends**  
Edited by: K Venema, AP Carmo (2015)  
[www.caister.com/probiotics](http://www.caister.com/probiotics)
- **Epigenetics: Current Research and Emerging Trends**  
Edited by: BP Chadwick (2015)  
[www.caister.com/epigenetics2015](http://www.caister.com/epigenetics2015)
- **Corynebacterium glutamicum: From Systems Biology to Biotechnological Applications**  
Edited by: A Burkovski (2015)  
[www.caister.com/cory2](http://www.caister.com/cory2)
- **Advanced Vaccine Research Methods for the Decade of Vaccines**  
Edited by: F Bagnoli, R Rappuoli (2015)  
[www.caister.com/vaccines](http://www.caister.com/vaccines)
- **Antifungals: From Genomics to Resistance and the Development of Novel Agents**  
Edited by: AT Coste, P Vandeputte (2015)  
[www.caister.com/antifungals](http://www.caister.com/antifungals)
- **Bacteria-Plant Interactions: Advanced Research and Future Trends**  
Edited by: J Murillo, BA Vinatzer, RW Jackson, et al. (2015)  
[www.caister.com/bacteria-plant](http://www.caister.com/bacteria-plant)
- **Aeromonas**  
Edited by: J Graf (2015)  
[www.caister.com/aeromonas](http://www.caister.com/aeromonas)
- **Antibiotics: Current Innovations and Future Trends**  
Edited by: S Sánchez, AL Demain (2015)  
[www.caister.com/antibiotics](http://www.caister.com/antibiotics)
- **Leishmania: Current Biology and Control**  
Edited by: S Adak, R Datta (2015)  
[www.caister.com/leish2](http://www.caister.com/leish2)
- **Acanthamoeba: Biology and Pathogenesis (2nd edition)**  
Author: NA Khan (2015)  
[www.caister.com/acanthamoeba2](http://www.caister.com/acanthamoeba2)
- **Microarrays: Current Technology, Innovations and Applications**  
Edited by: Z He (2014)  
[www.caister.com/microarrays2](http://www.caister.com/microarrays2)
- **Metagenomics of the Microbial Nitrogen Cycle: Theory, Methods and Applications**  
Edited by: D Marco (2014)  
[www.caister.com/n2](http://www.caister.com/n2)



Order from [caister.com/order](http://www.caister.com/order)