Available online at www.sciencedirect.com
SCIENCE DIRECT
Virus Research 113 (2005) 123–131
Virus Research
www.elsevier.com/locate/virusres

Synonymous codon usage bias in 16 *Staphylococcus aureus* phages: Implication in phage therapy

K. Sau [a], S.K. Gupta [a], S. Sau [b,*], T.C. Ghosh [a,*]

[a] *Bioinformatics Centre, Bose Institute, P1/12, CIT Scheme VII M, Calcutta 700 054, India*
[b] *Department of Biochemistry, Bose Institute, P1/12, CIT Scheme VII M, Calcutta 700 054, India*

Received 26 January 2005; received in revised form 6 May 2005; accepted 10 May 2005
Available online 20 June 2005

Abstract

To reveal the factors influencing architecture of protein-coding genes in staphylococcal phages, relative synonymous codon usage variation has been investigated in 920 protein-coding genes of 16 staphylococcal phages. As expected for AT rich genomes, there are predominantly A and T ending codons in all 16 phages. Both N_c plot and correspondence analysis on relative synonymous codon usage indicates that mutation bias influences codon usage variation in the 16 phages. Correspondence analysis also suggests that translational selection and gene length also influence the codon usage variation in the phages to some extent and codon usage in staphylococcal phages is phage-specific but not *S. aureus*-specific. Further analysis indicates that among 16 staphylococcal phages, 44AHJD, P68 and K may be extremely virulent in nature as most of their genes have high translation efficiency. If this is true, then above three phages may be useful for curing staphylococcal infections.
© 2005 Elsevier B.V. All rights reserved.

Keywords: Staphylococcal phages; Synonymous codon usage; Mutational bias; Translational selection; Phage therapy

1. Introduction

Synonymous codon usage bias had been determined in numerous organisms and it was shown that codon usage varies not only among the genes of a genome but also varies from genome to genome (D'Onofrio and Bernardi, 1992; Gu et al., 2004; Jenkins and Holmes, 2003; Kunisawa et al., 1998; Lynn et al., 2002; Majumdar et al., 1999; Moriyama and Powell, 1998). Several factors such as mutational bias (Jenkins and Holmes, 2003; Levine and Whitemore, 2000), translational selection (Ghosh et al., 2000; Grantham et al., 1981; Gupta and Ghosh, 2001; Ikemura, 1985; Lesnik et al., 2000; Sharp and Cowe, 1991), secondary structure of proteins (Chiusano et al., 2000; Gupta et al., 2000; Oresisc and Shalloway, 1998; Xie and Ding, 1998), replicational and transcriptional selection (McInerney, 1998; Romero et al., 2000), environmental factors (Basak et al., 2004; Lynn et al., 2002), etc. were shown to influence the codon usage bias in the organisms.

The phages of *Staphylococcus aureus* are highly diverse in nature and distributed across the three families of bacteriophages, namely, Podoviridae (Vybiral et al., 2003), Myoviridae (O'Flaherty et al., 2004), and Siphoviridae (Ackermann, 2001; Iandolo et al., 2002; Liu et al., 2004). The genomes of staphylococcal phages so far reported are GC-poor and have highly variable sizes (Ackermann, 2001; Iandolo et al., 2002; Vybiral et al., 2003; O'Flaherty et al., 2004; Liu et al., 2004). Most of them were even shown to carry different toxin – encoding genes (Iandolo et al., 2002; Brüssow et al., 2004; Pantůček et al., 2004). Thus far, little work has been carried out to understand the codon usage bias in *S. aureus* phages though it has immense potentials in enriching the molecular biology of these phages. Besides, such study may give clues about the virulence nature of the lytic staphylococcal phages.

Abbreviations: RSCU, relative synonymous codon usage; CA, correspondence analysis; N_c, effective number of codons; GC_{3s}, the frequency of G + C at the synonymous third codon positions

* Corresponding authors. Tel.: +91 33 2334 6626; fax: +91 33 2334 3886.
E-mail addresses: sau@bic.boseinst.ernet.in (S. Sau), tapash@bic.boseinst.ernet.in (T.C. Ghosh).

0168-1702/$ – see front matter © 2005 Elsevier B.V. All rights reserved.
doi:10.1016/j.virusres.2005.05.001

The information may be useful in selecting the extremely virulent staphylococcal phages suitable for phage therapy. In this paper, we studied the synonymous codon usage bias in the 920 genes of 16 *S. aureus* phages at length. Our analysis showed that the synonymous codon usage pattern significantly varies from phage to phage and the codon usage bias in these phages is dictated by several factors such as mutational pressure, translation selection and gene length. Finally, based on our data and others' data, we proposed that among the 16 staphylococcal phages, 44AHJD, P68 and K may be useful for curing staphylococcal infections.

2. Material and methods

Complete genome sequences of the *S. aureus* phages 77, 44AHJD, K, ϕ11, ϕ12, ϕ13, ϕETA, ϕPV83, ϕPVL, ϕSLT, ϕN315, ϕMu50A, ϕMu50B, ϕSa2mw, ϕSa3mw, and ϕP68, downloaded from NCBI (USA) and EMBL (UK), were utilized to study the synonymous codon usage bias in these phages. The basic nature and status of the above 16 staphylococcal phages were presented in Table 1. Among the above phages, phage K belongs to the family Myoviridae (O'Flaherty et al., 2004), whereas, 44AHJD and P68 belong to the family Podoviridae (Vybiral et al., 2003). All other phages except ϕMu50A, ϕMu50B, ϕSa2mw, ϕSa3mw belong to Siphoviridae. Gene number in the above 16 phages varies from 21 to 118. All the coding sequences of these 16 phages were extracted from the feature table of genome (according to GenBank records) by a program written in C. Coding sequences carrying less than 50 codons were not considered in the study. Thus, a total of 920 protein-coding genes from 16 phages were utilized for all analyses here. All coding sequences of *S. aureus* strain Mu50 (accession no. NC_002758) except those belonging to ϕMu50A and ϕMu50B prophages were also used in this comparative study.

The relative synonymous codon usage (RSCU) in all phages had been determined to study the overall codon usage variation among the genes and genomes. RSCU is defined as the ratio of the observed frequency of codons to the expected frequency if all the synonymous codons for those amino acids are used equally (Sharp and Li, 1987). RSCU values greater than 1.0 indicate that the corresponding codon is more frequently used than expected, whereas the reverse is true for RSCU values less than 1.0.

GC_{3S} is the frequency of (G + C) and A_{3S}, T_{3S}, G_{3S}, and C_{3S} are the frequencies of A, T, G, and C at the synonymous third positions of codons. N_c, the effective number of codons used by a gene, is generally used to measure the bias of synonymous codons and independent of amino acid compositions and codon number (Wright, 1990). The values of N_c range from 20 (when one codon is used per amino acid) to 61 (when all the codons are used with equal probability). Highly biased genes are generally highly expressed (Sharp and Cowe, 1991) and as there is no information of gene expression level of bacteriophages we have considered the highly biased genes as highly expressed in case of staphylococcal phages under study. Hence the sequences in which the N_c values are less than 30 were considered to be highly expressed and those have N_c values greater than 55 were considered as lowly expressed genes. All the above-mentioned parameters and correspondence analysis (Greenacre, 1984) were carried out by the program CodonW 1.3 (available at http://www.molbiol.ox.ac.uk/cu). The copy number of tRNA species in Mu50 and corresponding anticodon sequences were determined by a tRNAscan-SE program (available at http://www.genetics.wustl.edu/eddy/tRNAscan-SE).

Table 1
Basic characteristics of 16 staphylococcal phages

Phage name	Family	Virulent or temperate	GenBank accession no.	Number of genes per phage	Presence of toxin gene in phage	Number of genes per phage carrying ≥50 codons
K	Myoviridae	Virulent	NC_005880	118	No	115
44AHJD	Podoviridae	Virulent	NC_004678	21	No	21
P68	Podoviridae	Virulent	NC_004679	22	No	22
77	Siphoviridae	Temperate	NC_005356	69	NK	65
ϕ11	Siphoviridae	Temperate	NC_004615	53	No	53
ϕ12	Siphoviridae	Temperate	NC_004616	49	No	49
ϕ13	Siphoviridae	Temperate	NC_004617	49	Yes	49
ϕPV83	Siphoviridae	Temperate	NC_002486	65	Yes	62
ϕPVL	Siphoviridae	Temperate	NC_002321	62	Yes	59
ϕSLT	Siphoviridae	Temperate	NC_002661	61	Yes	60
ϕETA	Siphoviridae	Temperate	NC_003288	66	Yes	65
ϕN315	Siphoviridae	Temperate	NC_004740	65	Yes	61
ϕMu50A	NK	Temperate	[a]	63	Yes	61
ϕMu50B	NK	Temperate	[a]	70	No	62
ϕSa2mw	NK	Temperate	[b]	65	Yes	62
ϕSa3mw	NK	Temperate	[b]	58	Yes	54

NK: not known.
[a] Prophages ϕMu50A and ϕMu50B carry genes SAV1940–SAV2002 and SAV847–SAV916 of *S. aureus* Mu50, respectively.
[b] Prophages ϕSa2mw and ϕSa3mw harbor genes MW1378–MW14442 and MW1882–1940 of *S. aureus* MW2, respectively.

Download English Version:

https://daneshyari.com/en/article/9289291

Download Persian Version:

https://daneshyari.com/article/9289291

Daneshyari.com