Genetics and Molecular Biology, 33, 2, 408-410 (2010)

www.sbg.org.br

Short Communication

Translational selection on SHH genes

Mohammadreza Hajjari[1], Behnaz Saffar[2,3] and Atefeh Khoshnevisan

[1]*Department of Genetics, Faculty of Biological Science, Tarbiat Modares University, Tehran, Iran.*
[2]*Department of Genetics, Faculty of Science, Shahrekord University, Shahrekord, Iran.*
[3]*Biotechnology Research Department, Shahrekord University, Shahrekord, Iran.*

Abstract

Codon usage bias has been observed in various organisms. In this study, the correlation between SHH genes expression in some tissues and codon usage features was analyzed by bioinformatics. We found that translational selection may act on compositional features of this set of genes.

Key words: bioinformatics, codon usage bias, correlation, translational selection.

Received: June 15, 2009; Accepted: November 17, 2009.

Synonymous codons are codons which code for the same amino acid. Non-random usage of these codons is a widespread occurrence which has been observed in many organisms such as *Drosophila melanogaster, E.coli* and *Saccharomyces cerevisiae* (Sharp and Li, 1987; Ikemura, 1985; Powell and Moriyama, 1997). One of the models that express such biases is translational selection (Zuckerkandl and Pauling, 1965; Ikemura, 1981; Sørensen *et al.*, 1989; Debry and Marzluff, 1994; Levy *et al.*, 1996). Correlations between synonymous codon usage and gene expression levels are usually attributed to the higher abundance of isoaccepting tRNA for optimal codons to maximize translational efficiency both in unicellular (Dong *et al.*, 1996; Rocha, 2004) and multicellular organisms (Moriyama *et al.*, 1997; Duret, 2000; Kanaya *et al.*, 2001; Lavner and Kotlar, 2005). In vertebrates there has been much debate on this correlation (Wolfe and Sharp, 1993; Musto *et al.*, 2001; Konu and Li, 2002). Moreover, this correlation is not clear in mammals including human beings.

Thus, analyzing codon usage bias in specific tissues and specific subsets of genes is an interesting area of investigation. The Sonic Hedgehog (SHH) pathway is one of the most important developmental pathways which is conserved from flies to human. The proteins that are involved in this pathway are not only critical in embryonic development but have also been implicated in certain cancers. In this study we analyzed the correlation between synonymous codon usage features of SHH pathway genes and their expression in human tissues. The results are expected to provide valuable information on translational efficiency of this pathway in humans. By such information, analyzing cancer mechanisms and causative mutations may become more expedient.

SHH pathway genes were collected from a Hedgehog Signaling Pathway Database. In total 31 genes were selected from SHH signal receiving cells. The NCBI database was used to obtain their CDSs. To avoid statistical errors, the CDSs were aligned by ClustalW. Similarity scores of every two genes were between one and 68. Normalized expression levels in the brain, embryonic tissue, prostate, ovary, testis, liver, muscle, and the eye were retreived from the SOURCE database. By using the FREQSQ program we calculated synonymous codon usage features and percentage of each synonymous codon in each codon family that codes for the same amino acid. Statistical analysis was performed by using the MINITAB13 program. Codons with p values below 0.01 were considered as significant features.

After analyzing 59 synonymous codon usage features (all codons except termination codons, Methionine, and Tryptophan codons) for 31 genes in the SHH pathway, 1829 features were obtained. Then, by analyzing the relationship between gene expression levels in eight tissues and synonymous codon usage features, 13 significant features were noted in four tissues (brain, ovary, testis and the eye) (Table 1). Among significant features, three of "synonymous codon usage" were shared by two tissues. Regression equations were then obtained for significant features (Table 2). For one amino acid (L) we could find two different and significant codon usage features in brain and testis.

The above analysis revealed a correlation between synonymous codon usage and expression level of SHH genes in brain, testis, ovary, and the eye. But this correlation was not observed in embryonic tissue, prostate, liver, and muscle. Thus, translational selection may select synonymous codons in genome sequences. Our analysis indicates

Send correspondence to Behnaz Saffar. P.O. Box 115, Shahrekord University Shahrekord, Iran. E-mail: saffar_b@sku.ac.ir.

Table 1 - Significant features for the correlation between expression levels of SHH genes and synonymous codon usage features in human tissues.

Tissue	Number of expressed genes	Average expression	Highest expressed gene	p value	r_s value	Feature
Brain	31	63.84	PRECAKB	0.001	0.548	CTA(L)
				0.002	0.542	TGT(C)*
				0.002	0.542	TGC(C)*
				0.002	0.542	CCA(P)*
Ovary	21	3.905	GPC1	0.008	-0.563	ATA(I)
				0.004	0.603	CAC(H)
				0.004	-0.606	GGA(G)
Testis	28	15.43	IFT88	0	0.647	TTA(L)
				0.004	0.522	CCA(P)*
				0.002	0.561	ACT(T)
				0.005	0.519	TGT(C)*
				0.005	-0.519	TGC(C)*
Prostate	24	5.304	CSNK1A	-	-	-
Embryonic tissue	22	5.45	GPC4	-	-	-
Eye	26	6.231	GSK3A and PRECAKB	0.008	0.509	TCA(S)
Liver	20	6.65	CSNK1A	-	-	-
Muscle	26	7.375	GLI1	-	-	-

*: Common features.

that SHH gene expression may be regulated at a post-transcriptional level. Common features between two tissues support the assumption that the same mechanism may act in common pathways. For instance, codon usage bias can be attributed to the frequency of isoaccepting tRNAs in a tissue. This may also account for the herein observed correlations, because differences in relative tRNA abundance with a maximum range of tenfold variation have been detected in different human tissues (Ditmar *et al.*, 2006). Furthermore, CTA and TTA, which encode leucin, are usage-biased in the brain and testis. This result indicates the criticality of this amino acid and its tRNAs in the two tissues. There are no significant features for synonymous codon usage in prostate, embryonic tissues, liver, and muscles, wherefore translational selection in these tissues may be absent

Table 2 - Regression equations for significant features.

Tissue	Equation
Brain (4 significant features)	EXP = - 49.9 + 13.7 CTA(L)
	EXP = - 75.6 + 2.79 TGT(C)
	EXP = 203-2.79 TGC(C)
	EXP = - 89.3 + 5.26 CCA(P)
Ovary (3significant features)	EXP = 6.68-0.187 ATA(I)
	EXP = - 1.34 + 0.0954 CAC(H)
	EXP = 8.57-0.165 GGA(G)
Testis (5significant features)	EXP = 3.54 + 1.28 TTA(L)
	EXP = - 5.19 + 0.671 CCA(P)
	EXP = - 13.4 + 1.03 ACT(T)
	EXP = - 3.02 + 0.348 TGT(C)
	EXP = 31.8-0.348 TGC(C)
Eye (1significant feature)	EXP = 1.45 + 0.310 TCA(S)

The average expression in brain and testis was higher than in other tissues. This result indicates translational efficiency and accuracy. In fact, the evolution of codon bias in highly expressed genes is hypothesized to be a result of natural selection for increased protein elongation rates (Bulmer, 1999) or minimized errors in mRNA translation (Akashi, 1994). Furthermore, the frequency of tRNAs or factors which are involved in translation may be different in various tissues. This result has notable implications for understanding the molecular mechanisms of tissue development and cancer. It is important to mention that other compositional features may also affect gene expression in selected tissues. Appropriate analysis can help in understanding molecular mechanisms of gene expression and mRNA translation. Hence, the relationships between genome, transcriptome and proteome may become clearer.

References

Akashi H (1994) Synonymous codon usage in *Drosophila melanogaster*: Natural selection and translational accuracy. Genetics 136:927-935.

Bulmer M (1991) The selection-mutation-drift theory of synonymous codon usage. Genetics 129:897-907.

Debry RW and Marzluff WF (1994) Selection on silent sites in the rodent H3 histone ene family. Genetics 138:191-202.

Ditmar KA, Goodenbour JM and Pan T (2006) Tissue specific differences in human transfer RNA expression. PLoS Genet 2:e221.

Dong H, Nilsson L and Kurland CG (1996) Co-variation of tRNA abundance and codon usage in *Escherichia coli* at different growth rates. J Mol Biol 260:649-663.

Duret L (2000) tRNA gene number and codon usage in the *C. elegans* genome are co-adapted for optimal translation of highly expressed genes. Trends Genet 16:287-289.

Ikemura T (1981) Correlation between the abundance of *Escherichia coli* transfer RNAs and the occurrence of the respective codons in its protein genes: A proposal for a synonymous codon choice that is optimal for the *E. coli* translational system. J Mol Biol 146:1-21.

Ikemura T (1985) Codon usage and tRNA content in unicellular and multicellular organisms. Mol Biol Evol 2:13-34.

Kanaya S, Yamada Y, Kinouchi M, Kudo Y and Ikemura T (2001) Codon usage and tRNA genes in eukaryotes: Correlation of codon usage diversity with translation efficiency and with CG-dinucleotide usage as assessed by multivariate analysis. J Mol Evol 53:290-298.

Konu O and Li MD (2002) Correlations between mRNA expression levels and GC contents of coding and untranslated regions of genes in rodents. J Mol Evol 54:35-41.

Lavner Y and Kotlar D (2005) Codon bias as a factor in regulating expression via translation rate in the Human genome. Gene 345:127-138.

Levy JP, Muldoon RR, Zolotukhin S and Link Jr CJ (1996) Retroviral transfer and expression of a humanized, red-shifted green fluorescent protein gene into human tumor cells. Nat Biotechnol 14:610-614.

Moriyama EN and Powell JR (1997) Codon usage bias and tRNA abundance in *Drosophila*. J Mol Evol 45:514-523.

Musto H, Cruveiller S, D' Onofrio G, Romero H and Bernardi G (2001) Translational selection on codon usage in *Xenopus laevis*. Mol Biol Evol 18:1703-1707.

Powell JR and Moriyama EN (1997) Evolution of codon usage bias in *Drosophila.* Proc Natl Acad Sci USA 94:7784-7790.

Rocha EP (2004) Codon usage bias from tRNA's point of view: Redundancy, specialization, and efficient decoding for translation optimization. Genome Res 14:2279-2286.

Sharp PM and Li WH (1987) The codon adaptation index - A measure of directional synonymous codon usage bias, and its potential applications. Nucleic Acids Res 15:1281-1295.

Sørensen MA, Kurland CG and Pedersen S (1989) Codon usage determines translation rate in *Escherichia coli*. J Mol Biol 207:365-377.

Wolfe KH and Sharp PM (1993) Mammalian gene evolution: Nucleotide sequence divergence between mouse and rat. J Mol Evol 37:441-456.

Zuckerkandl E and Pauling L (1965) Molecules as documents of evolutionary history. J Theor Biol 8:357-366.

Internet Resources

Hedgehog Signaling Pathway Database, http://www.hedgehog.sfsu.edu (April, 2009).

ClustalW, http://www.ebi.ac.uk/clustalw/ (April, 2009).

SOURCE database, http://smd.stanford.edu (April, 2009).

FREQSQ program, http://www.bioinfo.hku.hk/services/analyseq/ cgi-bin/freqsq_in.pl (April, 2009).

NCBI database, http://www.ncbi.nlm.nih.gov (April, 2009).

Associate Editor: Luciano da Fontoura Costa