Pak. J. Bot., 57(6): DOI: http://dx.doi.org/10.30848/PJB2025-6(9)

GENETIC VARIATION AND BIOINFORMATICS ANALYSIS OF THE TEF2 GENE OF *SORDARIA FIMICOLA*

AMNA RASHEED[1], RABIA ARIF[1], UZMA NAUREEN[1], AND MUHAMMAD SALEEM[1]

[1]*Institute of Botany, University of the Punjab Lahore, Pakistan*
**Corresponding author's amna.haris786@gmail.com*

Abstract

The current study aimed to find out genetic variation and post-translational modifications of the eukaryotic translational elongation factor (eEF-1) due to environmental stress. Ribosomal translation factors are directly involved in protein translation and are greatly conserved but several studies have revealed that eEF1A is post-translationally modified. Six parental strains of *Sordaria fimicola* (S1, S2, S3, N5, N6, and N7) were subjected to amplify the TEF2 gene, and a comparison was made with reference species *Saccharomyces cerevisiae*. Different bioinformatics tools, i.e., Clustal W, PMes Server, PAIL and NetNES Servers, NetPhos 3.1, and Phyre2 were used for multiple sequence alignment, to identify methylation sites, acetylation sites on the protein, phosphorylation sites, and Protein 3D structure. The PTMs of *S. fimicola* strains were compared with PTMs of reference species. Out of the total, 329 nucleotides polymorphism was observed on five nucleotide positions. Altogether, a total of 13 sites (8Ser/4Thr/2Y) of phosphorylation modifications were observed in *S. cerevisiae* and *S. fimicola*. The polymorphism was observed at 23 sites in strains of SFS (S1, S2, and S3) while six polymorphic sites were present in NFS strains (NFS). The study reiterated that environmental stress causes genetic variation and ultimately leads to evolution.

Key words: Post-translational modification (PTM), Protein synthesis, *Sordaria fimicola*, Translation elongation factor, Eukaryotic elongation factor eEF1A.

Introduction

The eukaryotic translational elongation factor (eEF-1) complex plays both canonical and non-canonical functions in the cell. During the mechanism of protein synthesis, the protein complex subunits play a fundamental role, as it controls the delivery step of aa-tRNA during the elongation stage of translation. The role of eEF1A is attached with aminoacylated tRNAs and helps its transfer to the A-site of the translating 80S ribosome. Besides its well-established role in the elongation phase of translation, it also controls different cellular processes. The protein complex has two paralogs eEF1A and eEF1B, they are involved in aminoacyl-tRNAs delivery and catalyze them to the ribosomes A-site and act as guanine nucleotide exchange factor (GEF) respectively (Andersen *et al*., 2003 & Khacho *et al*., 2008).

eEF-1A is extensively post-translationally modified by lysyl acetylation, methylation (Cavallius *et al*., 1993), ubiquitination, nitrosylation, glutathionylation, phosphorylation (Trauhj *et al*., 2001), C-terminal methyl esterification (Zobel-Trropp *et al*., 2000), and the attachment of ethanolamine phosphoglycerol (EPG) (Whiteheart *et al*., 1989).

The functional and structural complexity of the proteome is enhanced greatly by spatial and temporal modifications termed as post-translational modifications, subjected to almost all the proteins after their translation. The role of PTMs in cellular processes can be understood by studying, analyzing, and predicting them *In Silico* by using different computational tools which are online available and are user-friendly webservers (Shi, 2019).

Furthermore, the 3D structure of the proteome can also help to predict the modified site of proteins. In the current study, we predict the most common PTMs as Phosphorylation, acetylation, glycosylation, and methylation by using different computational approaches.

Environmental extremes have a remarkable impact on the expression of genetic variation in individuals which leads toward evolution (Hoffman *et al*., 1997). When there is a variation in physical conditions it affects reproductive growth, and when the perseverance of the conditions lead to long-lasting harm, these factors create environmental stress. Stress is an environmental factor that led to evolution and causes variation in living organisms (Parsons, 2005). Significant variation in morphological and physiological characteristics can be accomplished over inducing environmental stress (Weber, 1990 & Weber & Diggins, 1990). The research was undertaken to pursue the effect of stressful conditions on genetic variation and post-translational modifications in the translational elongation factor (eEF-1) complex in *Sordaria fimicola*.

For the present study, six parental strains of *S. fimicola* were used; the strains were earlier collected from Evolution Canyon (EC). Evolution canyon is considered as microclimate to study the relationship between climatic changes and genetic variation; the EC has two slopes; one is tropical, and the other is temperate. The south-facing slope (SFS) is dry, harsh, and has xeric conditions while the north-facing slope (NFS) is mesic, lush green, and moderate in terms of abiotic factors (Nevo, 1995). Strains S1, and S3, were collected from SFS, and N5, N6, and N7 were collected from NFS.

Methodology

Organism: Six various strains of *S. fimicola* (S1, S2, S3, N5, N6, and N7) were collected from molecular genetics and research laboratory, Department of Botany, University of Punjab, Lahore. The strains were earlier collected from Evolution Canyon (EC), Israel. EC has two opposite slopes one is a tropical Asian-African tropical south-facing slope (SFS) and the other is a temperate European north-facing slope (NFS).

DNA extraction: Total genomic DNA extraction was performed by taking perithecia of *Sordaria fimicola*, which were already cultured on potato dextrose agar (PDA) media by following the protocol described by Pietro *et al*., 1995. The quality and quantity of DNA were examined by running the 1% agarose gel electrophoresis and spectrophotometrically for different DNA samples. The gel was visualized under UV light of Gel documentation system (Syngene Germany). The DNA was stored at -20 until further use.

Primer design and PCR amplification: To amplify tef2 gene primers were designed manually and by using primer 3 plus software available at https://primer3plus.com. Primers were used to amplify the candidate gene; of Boiron, Germany has given in (Table 1).

Touch down PCR conditions was used for the amplification purpose, a total volume of 20 µl was prepared by adding 10 µl master mix, 2.5 µl of gDNA, 1 µl of forward primer and I µl of reverse primer, and 5.5 µl of dd H_2O. The PCR amplicons were separated on 1% agarose gel stained with ethidium bromide (BIORAD) along with 1kb DNA ladder (Invitrogen) and visualized under ultraviolet rays in the Gel Doc system. The amplicons were sent to Macrorogen Korea for sequencing. Nucleotides sequences were aligned by using Clustal Omega O software of the tef-2 gene of *S. cerevisiae* (Accession number FR774388) and of *S. fimicola.*

Prediction of PTMS: Various computational tools to predict PTMs of TEF2 gene were used like YinOYang 1.2 (www.cbs.dtu.dk/services/YinOYang/for glycosylation, Netphos 3.1 (www.cbs.dtu.dk/services/NetPhos/) for phosphorylation, PAIL (bdmpail.biocuckoo.org/) for acetylation and ModPred (montana.informatics.indiana.edu/ModPred/faq.html) for methylation on different amino acid residues.

Prediction of protein 3D structure: Phyre2 (Protein Homology/analogy Recognition Engine V 2.0) bioinformatics tool available at http://www.sbg.bio.ic.ac.uk/~phyre2 was used to predict 3D structures and functions of elongation factor complex protein. PyMol molecular system and BioLip were used to visualize structures of protein and to predict ligand-protein interaction.

Results

Single nucleotide polymorphism (SNP) analysis: The sequences of all six strains of *S. fimicola* (S1, S2, S3, N5, N6, and N7) were aligned along with reference species *S. cerevisiae* with the online available tool Clustal Omega 3 to observe polymorphism among various strains of *S. fimicola*. The polymorphism was observed at 23 sites in *S. fimicola* strains, out of which 17 sites were observed in SFS strains (S1, S2, and S3) while six polymorphic sites were present in NFS strains (N5, N6, and N7) the results have been shown in (Figs. 1 & 2).

Prediction of phosphorylation: Protein phosphorylation majorly occurs on serine (S), threonine (T), and tyrosine (Y). For the elongation factor 1 protein of *S. cerevisiae*, phosphorylation was predicted at 13 serine (S), 16 threonines (T), and 3 tyrosine (Y) residues at various positions in the amino acid sequence. Phosphorylation was predicted at 13 serine (S), 17 threonines (T), and 3 tyrosine (Y) residues of S1, S2, and S3 strains. Phosphorylation for N5, N6, and N7 strains was observed at 14 serine (S), 15 threonines (T,), and 3 tyrosine (Y) residues. Results have been shown in (Table 2).

Prediction of *O*-glycosylation: YinOYang 1.2 server results have shown in Table 3 and Graphical representation (Fig. 3) indicates that in *S. cerevisiae O*-glycosylation for elongation factor 1 protein was predicted at 3 serine (S), 5 threonines (T). *O*-glycosylation was predicted at 3 serine (S) and 8threoninese (T) residues at the sequence of amino acids in S1, S2, and S3 strains. While 3 serine (S) and 5 threonine (T) residues were observed for O-glycosylation in N5, N6, and N7 strains.

Table 1. Primers sequences used to amplify tef2 gene.

Sr. No.	Name	Sequence ()	Expected product length (bp)	Tm
1.	TEFF1	ATGGGTAAAGAGAAGTCTCA	421	52
2.	TEFR1	CATTTGACGGAGTCCATCT		53
3.	TEFF2	TGGGACGAATCCATTCCA	420	52
4.	TEFR2	CACCTGGAACACCTTGTTCC		58
5.	TEFF3	GTTCCAGGTGACAACGTTGG	460	58
6.	TEFR3	TTTGAGCAGCCTTGGTAACC		56

Table2. Phosphorylation prediction on serine, threonine and tyrosine residues in *S. fimicola* and *S. cerevisiae.*

Organism	Serine	Threonine	Tyrosine	Kinases
S. cerevisiae	6, 18, 21, 53, 107, 157, 192, 273, 289, 314, 356, 381, 438	22, 72, 82, 88, 106, 142, 172, 240, 267, 277, 284, 285, 341, 430, 442, 449	56, 85, 252	PKC, PKG, unsp, PKA, p38MAPK, cdk5, CKII, INSR, cdc2, DNAPK, PKB
Total	**13**	**16**	**03**	**11**
SFS strains of *S. fimicola*	6, 18, 21, 53, 107, 157, 192, 237, 289, 314, 356, 381, 394	22, 72, 82, 88, 106, 142, 172, 240, 267, 277, 284, 285, 341, 363, 430, 442, 449	56, 85, 252	PKC, PKG, unsp, PKA, p38MAPK, cdk5, CKII, INSR, cdc2, DNAPK, PKB
Total	**13**	**17**	**03**	**11**
NFS strains of *S. fimicola*	6, 18, 21, 53, 107, 157, 192, 225, 237, 289, 314, 356, 381, 438	22, 72, 82, 88, 106, 142, 240, 267, 277, 284, 285, 341, 430, 442, 449	56, 85, 252	PKC, PKG, unsp, PKA, p38MAPK, cdk5, CKII, INSR, cdc2, DNAPK, PKB
Total	**14**	**15**	**03**	**11**

Highlighted areas show polymorphisms on that site; Red highlighted areas show polymorphisms on that site while yellow highlighted areas show conserved sites

```
S3      AACATGATTACTGGTACTTCTCATGCTGACTGTGCTATCTTGATTATTGCAGGTGGTGTC  360
S1      AACATGATTACTGGTACTTCTCATGCTGACTGTGCTATCTTGATTATTGCAGGTGGTGTC  360
S2      AACATGATTACTGGTACTTCTCATGCTGACTGTGCTATCTTGATTATTGCAGGTGGTGTC  360
Ref.    AACATGATTACTGGTACTTCTCAAGCTGACTGTGCTATCTTGATTATTGCTGGTGGTGTC  360
N6      ATCATGATTACTGGTACTTCTCAAGCTGACTGTGCTATCTTGATTATTGCTGGTGGTGTC  360
N5      ATCATGATTACTGGTACTTCTCAAGCTGACTGTGCTATCTTGATTATTGCTGGTGGTGTC  360
N7      ATCATGATTACTGGTACTTCTCAAGCTGACTGTGCTATCTTGATTATTGCTGGTGGTGTC  360
        * ******************** ************************ *********

S3      GGTGAATTCGTAGCCGGTATCTCTAAGGATGGTCAAACCAGAGAACACCCTTTGTTGGCT  420
S1      GGTGAATTCGTAGCCGGTATCTCTAAGGATGGTCAAACCAGAGAACACCCTTTGTTGGCT  420
S2      GGTGAATTCGTAGCCGGTATCTCTAAGGATGGTCAAACCAGAGAACACCCTTTGTTGGCT  420
Ref.    GGTGAATTCGAAGCCGGTATCTCTAAGGATGGTCAAACCAGAGAACACGCTTTGTTGGCT  420
N6      GGTGAATTCGAAGCCGGTATCTCTAAGGATGGTCAAACCAGAGAACACGCTTTGTTGGCT  420
N5      GGTGAATTCGAAGCCGGTATCTCTAAGGATGGTCAAACCAGAGAACACGCTTTGTTGGCT  420
N7      GGTGAATTCGAAGCCGGTATCTCTAAGGATGGTCAAACCAGAGAACACGCTTTGTTGGCT  420
        ********* ************************************ **********

S3      TTCACCTTGGGTGTTAGACAATTGATTGTTGCTCTCAACAAGATGGACTCCGTCAAATGG  480
S1      TTCACCTTGGGTGTTAGACAATTGATTGTTGCTCTCAACAAGATGGACTCCGTCAAATGG  480
S2      TTCACCTTGGGTGTTAGACAATTGATTGTTGCTCTCAACAAGATGGACTCCGTCAAATGG  480
Ref.    TTCACCTTGGGTGTTAGACAATTGATTGTTGCTGTCAACAAGATGGACTCCGTCAAATGG  480
N6      TTCACCTTGGGTGTTAGACAATTGATTGTTGCTGTCAACAAGATGGACTCCGTCAAATGG  480
N5      TTCACCTTGGGTGTTAGACAATTGATTGTTGCTGTCAACAAGATGGACTCCGTCAAATGG  480
N7      TTCACCTTGGGTGTTAGACAATTGATTGTTGCTGTCAACAAGATGGACTCCGTCAAATGG  480
        ******************************** *************************

S3      GAAGCTACCTCCAACGCTCCATGGTACAAGCGTTGGGAAAAGGAAACCAAGGCCGGTGTC  660
S1      GAAGCTACCTCCAACGCTCCATGGTACAAGCGTTGGGAAAAGGAAACCAAGGCCGGTGTC  660
S2      GAAGCTACCTCCAACGCTCCATGGTACAAGCGTTGGGAAAAGGAAACCAAGGCCGGTGTC  660
Ref.    GAAGCTACCACCAACGCTCCATGGTACAAGGGTTGGGAAAAGGAAACCAAGGCCGGTGTC  660
N6      GAAGCTACCACCAACGCTCCATGGTACAAGGGTTGGGAAAAGCAAACCAAGGCCGGTGTC  660
N5      GAAGCTACCACCAACGCTCCATGGTACAAGGGTTGGGAAAAGGAAACCAAGGCCGGTGTC  660
N7      GAAGCTACCACCAACGCTCCATGGTACAAGGGTTGGGAAAAGGAAACCAAGGCCGGTGTC  660
        ********* ******************** *********** *****************

S3      GCTGGTGTTACCACTGAAGTCAAGTCGGTTGAAATGCATCACGAACAATTGGAACAAGGT  900
S1      GCTGGTGTTACCACTGAAGTCAAGTCGGTTGAAATGCATCACGAACAATTGGAACAAGGT  900
S2      GCTGGTGTTACCACTGAAGTCAAGTCGGTTGAAATGCATCACGAACAATTGGAACAAGGT  900
Ref.    GCTGGTGTTACCACTGAAGTCAAGTCCGTTGAAATGCATCACGAACAATTGGAACAAGGT  900
N6      GCTGGTGTTACCACTGAAGTCAACTCCGTTGAAATGCATCACGAACAATTGGAACAAGGT  900
N5      GCTGGTGTTACCACTGAAGTCAACTCCGTTGAAATGCATCACGAACAATTGGAACAAGGT  900
N7      GCTGGTGTTACCACTGAAGTCAACTCCGTTGAAATGCATCACGAACAATTGGAACAAGGT  900
        ********************** ** ********************************

S3      ACCGTCATTGTTTTGAACCATGCAGGTCAAATCACTGCTGGTTACTCACCAGTTTTGGAT  1080
S1      ACCGTCATTGTTTTGAACCATGCAGGTCAAATCACTGCTGGTTACTCACCAGTTTTGGAT  1080
S2      ACCGTCATTGTTTTGAACCATGCAGGTCAAATCACTGCTGGTTACTCACCAGTTTTGGAT  1080
Ref.    ACCGTCATTGTTTTGAACCATGCAGGTCAAATCTCTGCTGGTTACTCTCCAGTTTTGGAT  1080
N6      ACCGTCATTGTTTTGAACCATCCAGGTCAAATCTCTGCTGGTTACTCTCCAGTTTTGGAT  1080
N5      ACCGTCATTGTTTTGAACCATCCAGGTCAAATCTCTGCTGGTTACTCTCCAGTTTTGGAT  1080
N7      ACCGTCATTGTTTTGAACCATCCAGGTCAAATCTCTGCTGGTTACTCTCCAGTTTTGGAT  1080
        ******************** ********** ************ ************

S3      TGTCACACTCCTCACATTGCTTGTAGATTCGACGAATTGTTGGAAAAGAACGACAGAAGA  1140
S1      TGTCACACTCCTCACATTGCTTGTAGATTCGACGAATTGTTGGAAAAGAACGACAGAAGA  1140
S2      TGTCACACTCCTCACATTGCTTGTAGATTCGACGAATTGTTGGAAAAGAACGACAGAAGA  1140
Ref.    TGTCACACTGCTCACATTGCTTGTAGATTCGACGAATTGTTGGAAAAGAACGACAGAAGA  1140
N6      TGTCACACTGCTCACATTGCTTGTAGATTCGACGAATTGTTGGAAAAGAACGACAGAAGA  1140
N5      TGTCACACTGCTCACATTGCTTGTAGATTCGACGAATTGTTGGAAAAGAACGACAGAAGA  1140
N7      TGTCACACTGCTCACATTGCTTGTAGATTCGACGAATTGTTGGAAAAGAACGACAGAAGA  1140
        ********* ************************************************

S3      TCTGGTAAGAAGTTGGAAGACCATCCAAAGTTGTTGAAGTCCGGTGACGCTGCTTTGGTC  1200
S1      TCTGGTAAGAAGTTGGAAGACCATCCAAAGTTGTTGAAGTCCGGTGACGCTGCTTTGGTC  1200
S2      TCTGGTAAGAAGTTGGAAGACCATCCAAAGTTGTTGAAGTCCGGTGACGCTGCTTTGGTC  1200
Ref.    TCTGGTAAGAAGTTGGAAGACCATCCAAAGTTCTTGAAGTCCGGTGACGCTGCTTTGGTC  1200
N6      TCTGGTAAGAAGTTGGAAGACCAACCAAAGTTCTTGAAGTCCGGTGACGCTGCTTTGGTC  1200
N5      TCTGGTAAGAAGTTGGAAGACCAACCAAAGTTCTTGAAGTCCGGTGACGCTGCTTTGGTC  1200
N7      TCTGGTAAGAAGTTGGAAGACCAACCAAAGTTCTTGAAGTCCGGTGACGCTGCTTTGGTC  1200
        ********************** ******** **************************
```

Fig. 1. Pairwise alignment of nucleotide sequence of Tef2 gene in *S. fimicola* and *S. cerevisiae* red lines indicates the Single nucleotide polymorphisms (SNPs).

```
S1      MGKEKSHINLVVIGHVDSGKSTTTGHLIYKCGGIDKRTIEKFEKEAAELGKGSFKYAWVL  60
S2      MGKEKSHINLVVIGHVDSGKSTTTGHLIYKCGGIDKRTIEKFEKEAAELGKGSFKYAWVL  60
S3      MGKEKSHINLVVIGHVDSGKSTTTGHLIYKCGGIDKRTIEKFEKEAAELGKGSFKYAWVL  60
Ref.    MGKEKSHINVVVIGHVDSGKSTTTGHLIYKCGGIDKRTIEKFEKEAAELGKGSFKYAWVL  60
N7      MGKEKSHINVVVIGHVDSGKSTTTGHLIYKCGGIDKRTIEKFEKEAAELGKGSFKYAWVL  60
N5      MGKEKSHINVVVIGHVDSGKSTTTGHLIYKCGGIDKRTIEKFEKEAAELGKGSFKYAWVL  60
N6      MGKEKSHINVVVIGHVDSGKSTTTGHLIYKCGGIDKRTIEKFEKEAAELGKGSFKYAWVL  60
        *********:**********************************************************

S1      DKLKAERERGFTIDIALWKFETPKYQVTVIDAPGHRDFIKNMITGTSHADCAILIIAGGV  120
S2      DKLKAERERGFTIDIALWKFETPKYQVTVIDAPGHRDFIKNMITGTSHADCAILIIAGGV  120
S3      DKLKAERERGFTIDIALWKFETPKYQVTVIDAPGHRDFIKNMITGTSHADCAILIIAGGV  120
Ref.    DKLKAERERGITIDIALWKFETPKYQVTVIDAPGHRDFIKNMITGTSQADCAILIIAGGV  120
N7      DKLKAERERGITIDIALWKFETPKYQVTVIDAPGHRDFIKIMITGTSQADCAILIIAGGV  120
N5      DKLKAERERGITIDIALWKFETPKYQVTVIDAPGHRDFIKIMITGTSQADCAILIIAGGV  120
N6      DKLKAERERGITIDIALWKFETPKYQVTVIDAPGHRDFIKIMITGTSQADCAILIIAGGV  120
        **********:****************************** ******:************

S1      GEFVAGISKDGQTREHPLLAFTLGVRQLIVALNKMDSVKWDESRFQEIVKETSNFIKKVG  180
S2      GEFVAGISKDGQTREHPLLAFTLGVRQLIVALNKMDSVKWDESRFQEIVKETSNFIKKVG  180
S3      GEFVAGISKDGQTREHPLLAFTLGVRQLIVALNKMDSVKWDESRFQEIVKETSNFIKKVG  180
Ref.    GEFEAGISKDGQTREHALLAFTLGVRQLIVAVNKMDSVKWDESRFQEIVKETSNFIKKVG  180
N7      GEFEAGISKDGQTREHALLAFTLGVRQLIVAVNKMDSVKWDESRFQEIVNETSNFIKKVG  180
N5      GEFEAGISKDGQTREHALLAFTLGVRQLIVAVNKMDSVKWDESRFQEIVKETSNFIKKVG  180
N6      GEFEAGISKDGQTREHALLAFTLGVRQLIVAVNKMDSVKWDESRFQEIVKETSNFIKKVG  180
        *** ********** **********:**********:**********

S1      YNPKTVPFVPISGWNGDNMIEATSNAPWYKRWEKETKAGVVKGKTLLEAIDAIEQPSRPT  240
S2      YNPKTVPFVPISGWNGDNMIEATSNAPWYKRWEKETKAGVVKGKTLLEAIDAIEQPSRPT  240
S3      YNPKTVPFVPISGWNGDNMIEATSNAPWYKRWEKETKAGVVKGKTLLEAIDAIEQPSRPT  240
Ref.    YNPKTVPFVPISGWNGDNMIEATTNAPWYKGWEKETKAGVVKGKTLLEAIDAIEQPSRPT  240
N7      YNPKTVPFVPISGWNGDNMIEATTNAPWYKGWEKETKAGVVKGKSLLEAIDAIEQPSRPT  240
N5      YNPKTVPFVPISGWNGDNMIEATTNAPWYKGWEKETKAGVVKGKSLLEAIDAIEQPSRPT  240
N6      YNPKTVPFVPISGWNGDNMIEATTNAPWYKGWEKQTKAGVVKGKSLLEAIDAIEQPSRPT  240
        **********************:****** ***:*********:**********

S1      DKPLRLPLQDVYKIGGIGTVPVGRVETGVIKPGMVVTFAPAGVTTEVKSVEMHHEQLEQG  300
S2      DKPLRLPLQDVYKIGGIGTVPVGRVETGVIKPGMVVTFAPAGVTTEVKSVEMHHEQLEQG  300
S3      HKPLRLPLQDVYKIGGIGTVPVGRVETGVIKPGMVVTFAPAGVTTEVKSVEMHHEQLEQG  300
Ref.    DKPLRLPLQDVYKIGGIGTVPVGRVETGVIKPGMVVTFAPAGVTTEVKSVEMHHEQLEQG  300
N7      DKPLRLPLQDVYKIGGIGTVPVGRVETGVIKPGMVVTFAPAGVTTEVNSVEMHHEQLEQG  300
N5      DKPLRLPLQDVYKIGGIGTVPVGRVETGVIKPGMVVTFAPAGVTTEVNSVEMHHEQLEQG  300
N6      DKPLRLPLQDVYKIGGIGTVPVGRVETGVIKPGMVVTFAPAGVTTEVNSVEMHHEQLEQG  300
        **********************************************:************

S1      VPGDNVGFNVKNVSVKEIRRGNVCGDAKNDPPKGCASFNATVIVLNHAGQITAGYSPVLD  360
S2      VPGDNVGFNVKNVSVKEIRRGNVCGDAKNDPPKGCASFNATVIVLNHAGQITAGYSPVLD  360
S3      VPGDNVGFNVKNVSVKEIRRGNVCGDAKNDPPKGCASFNATVIVLNHAGQITAGYSPVLD  360
Ref.    VPGDNVGFNVKNVSVKEIRRGNVCGDAKNDPPKGCASFNATVIVLNHPGQISAGYSPVLD  360
N7      VPGDNVGFNVKNVSVKEIRRGNVCGDAKNDPPKGCASFNATVIVLNHPGQISAGYSPVLD  360
N5      VPGDNVGFNVKNVSVKEIRRGNVCGDAKNDPPKGCASFNATVIVLNHPGQISAGYSPVLD  360
N6      VPGDNVGFNVKNVSVKEIRRGNVCGDAKNDPPKGCASFNATVIVLNHPGQISAGYSPVLD  360
        ********************************************** ***:*********

S1      CHTPHIACRFDELLEKNDRRSGKKLEDHPKLLKSGDAALVKFVPSKPMCVEAFSEYPPLG  420
S2      CHTPHIACRFDELLEKNDRRSGKKLEDHPKLLKSGDAALVKFVPSKPMCVEAFSEYPPLG  420
S3      CHTPHIACRFDELLEKNDRRSGKKLEDHPKLLKSGDAALVKFVPSKPMCVEAFSEYPPLG  420
Ref.    CHTAHIACRFDELLEKNDRRSGKKLEDHPKFLKSGDAALVKFVPSKPMCVEAFSEYPPLG  420
N7      CHTAHIACRFDELLEKNDRRSGKKLEDQPKFLKSGDAALVKFVPSKPMCVEAFSEYPPLG  420
N5      CHTAHIACRFDELLEKNDRRSGKKLEDQPKFLKSGDAALVKFVPSKPMCVEAFSEYPPLG  420
N6      CHTAHIACRFDELLEKNDRRSGKKLEDQPKFLKSGDAALVKFVPSKPMCVEAFSEYPPLG  420
        *** ***********************:**:*****************************

S1      RFAVRDMRQTVAVGVIKSVDMTEKAAKVTKAAQKAAKK*  458
S2      RFAVRDMRQTVAVGVIKSVDMTEKAAKVTKAAQKAAKK*  458
S3      RFAVRDMRQTVAVGVIKSVDMTEKAAKVTKAAQKAAKK*  458
Ref.     RFAVRDMRQTVAVGVIKSVDKTEKAAKVTKAAQKAAKK-  458
N7      RFAVRDMRQTVAVGVIKSVDKTEKAAKVTKAAQKAAKK*  458
N5      RFAVRDMRQTVAVGVIKSVDKTEKAAKVTKAAQKAAKK*  458
N6      RFAVRDMRQTVAVGVIKSVDKTEKAAKVTKAAQKAAKK*  458
        ******************* *****************
```

Fig. 2. Pairwise alignment of the amino acid sequence of EF1 protein in six strains of *S. fimicola* and *S. cerevisiae*. Alphabets in red representing the PTMs sites on Ser/Thr/Y and Lysine residue for potential acetylation, glycosylation and phosphorylation. Alphabets in Grey representing polymorphic sites.

Table 3. Prediction of *O*-glycosylation, Acetylation and Methylation in various strains of *S. fimicola* and *S. cerevisiae.*

Organisms	O-glycosylation	Acetylation	Methylation
	Residues Serine (S) and Threonine (T)	Residue Lysine (K)	Residue Arginine (R)
S. cerevisiae	107S, 185T, 204T, 237S, 284T, 285T, 352S, 449T	23, 39, 54, 58, 180, 225, 227, 245, 319, 387, 396, 404, 409, 440, 444, 447, 450, 453, 457, 460, 461	7,11,18
Total	**8**	**21**	**3**
SFS	107S, 185T, 204S,216T ,237S, 284T, 285T, 352T, 449T	4, 6, 21, 37, 52, 56, 178, 223, 225, 243, 317, 385, 394, 402, 407, 448, 451, 455, 458, 459	2,4,7
Total	**9**	**20**	**3**
NFS	107S, 185T, 204T,237S, 284T, 285T, 352S, 449T	4, 6, 21,37, 52, 56, 178 223, 225,243,317, 384, 385, 394, 402, 407, 438, 442, 445, 448, 451, 455, 458, 459	2,4,7
Total	**8**	**23**	**3**

Highlighted areas show polymorphisms on that site; Red highlighted areas show polymorphisms on that site while yellow highlighted areas show conserved sites

Prediction of methylation: Methylation at three arginine residues (R7, R11, and R18) was predicted in *S. cerevisiae* reference species. While (R2, R4, and R7) positions were investigated for methylation in all six strains of *S. fimicola.* R7 site seems to be a conserved site in all strains of *S. fimicola* and reference organisms.

Prediction of acetylation: (PAIL) server results have shown in (Table 3), acetylation at lysine (K) residue was investigated at 21 positions in *S. cerevisiae*. In the S1, S2, and S3 strains of *S. fimicola*, acetylation was found at 20 lysine residues and 24 sites were observed in the N5, N6, and N7 strains (Table 2). Only 225 lysines (k) position was found to be conserved in *S. cerevisiae* and *S. fimicola*.

Prediction of 3D protein structure: Phyr2 (Protein Homology/analogy Recognition Engine V 2.0) online server tool available at http://www.sbg.bio.ic.ac.uk/ ~phyre2 was used to visualize the 3D structure of elongation factor alpha 1 protein of S1 strain, N7 strains, and reference strain of *S. cerevisia*e. Graphical dissimilarities in the orientation of alpha and beta-helix were found in all the experimental strains due to the occurrence of variation in amino acid sequences (Fig. 4). The motifs shown in the helical form are α-helix, motifs in arrow form are β-sheets and it also shows the loop regions for the attachment of ligand.

Discussion

According to different studies effect of environmental stress can be positively associated with genetic variation in individuals (Parsons, 2005). In the present study genetic variation was observed in SFS strains that encounter environmental stress as compared to NFS strains. The six strains of S. *fimicola* S1, S2, S3, N5, N6, and N7 were used to amplify the tef-2 gene to investigate natural polymorphisms. Point mutations at four positions (112, 454, 674, and 1293) were found between *S. fimicola* sequences (FR774388) and the reference species.

Natural polymorphisms were at nineteen other positions. The 28C, 211T, 324T, 371T, 671G, 867G, 922G, 1048A, 1088G, 1172G, and 1322T variants were observed only in SFS strains while variations at position the 302T variant were present only in N5 strain and 643C in N6, 510G in N7 while 780C only in S2 strains. A total of 23 point mutations were observed in the exonic region of the tef-2 gene in which SFS strains showed more polymorphic sites these results are in accordance with Arif *et al*., (2017) who studied single nucleotide polymorphism in frequency clock and Mating Type a-1 proteins encoding genes and results showed that more genetic variation was observed in SFS strains than NFS strains. Similar findings were also reported by Hosid *et al*., (2008), which revealed that there is a relationship between genetic variation and environmental stress in the soil fungus *Emericella nidulans*. Environmental stresses contribute to genetic variation and change the gene frequencies, and ultimately lead towards evolution (Hoffman & Parsons, 1991).

It is reported that translation elongation factors are extremely conserved among various species and in addition to protein synthesis they might play a crucial role in other biological processes (Dever *et al*., 2018). EF-2 catalyzes the translocation of peptidyl-tRNA on the ribosome (Riis *et al*., 1990). Chen *et al*., (2019) worked on characterization and expression of translation elongation factor 2 from *Dermatophagoides farina*.

The structure and dynamics of proteome altered due to post-translational modifications which occur in almost every kind of protein after its synthesis. Various types of PTMs such as acetylation, phosphorylation, glycosylation, or methylation can be studied by using different computational tools.

Phosphorylation is the ubiquitous kind as it was reported that 30% of the total protein undergo the process of phosphorylation at least at one position. Protein kinases are the enzymes involved in the process of phosphorylation while protein phosphatases remove the chemical group as it is a kind of reversible PTM (Pinna, 1996 & Cohen, 2002). The present study reported a total of 32 phosphorylation sites of TEF2 protein in *S. cerevisiae* which included 13 Serine residues, 16 Threonine residues and 3 Tyrosine residues. In south-facing strains (S1, S2, S3), 13 Serine residues, 17 Threonine residues, and 3 Tyrosine residues were predicted as phosphorylation sites and variation occurred at 394S, 142T, 363T residues which were not found in reference organism but present in SFS strains (Table 2). In North-facing strains (N5, N6, N7), 14 Serine residues, 15 Threonine residues, and 3 Tyrosine residues were investigated as phosphorylation sites and variation occurred at 225S residue which was present in NFS strains but not found in reference organism (Table 1). Chen *et al*., (2019) reported 31 phosphorylation sites of TEF2 in *Dermatophagoides farina* which included 17 Serine residues 8 Threonine residues and 6 Tyrosine residues.

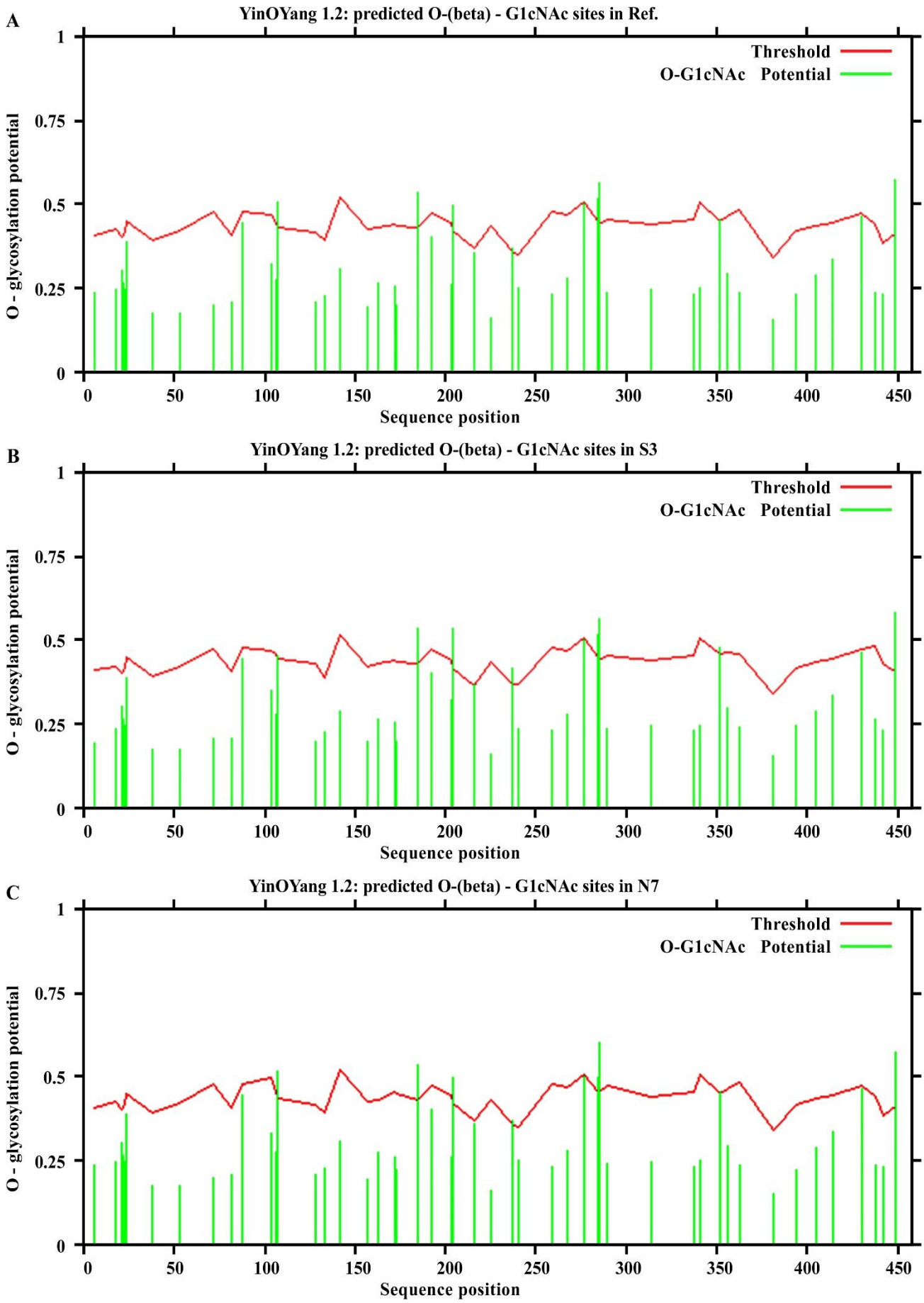

Fig. 3. Graphical representation of Prediction of Potential Glycosylation and YinOYang sites in *S. cerevisiae* (A), SFS strains (B) and NFS strains (C) of *S. fimicola* at 0.5 thresholds.

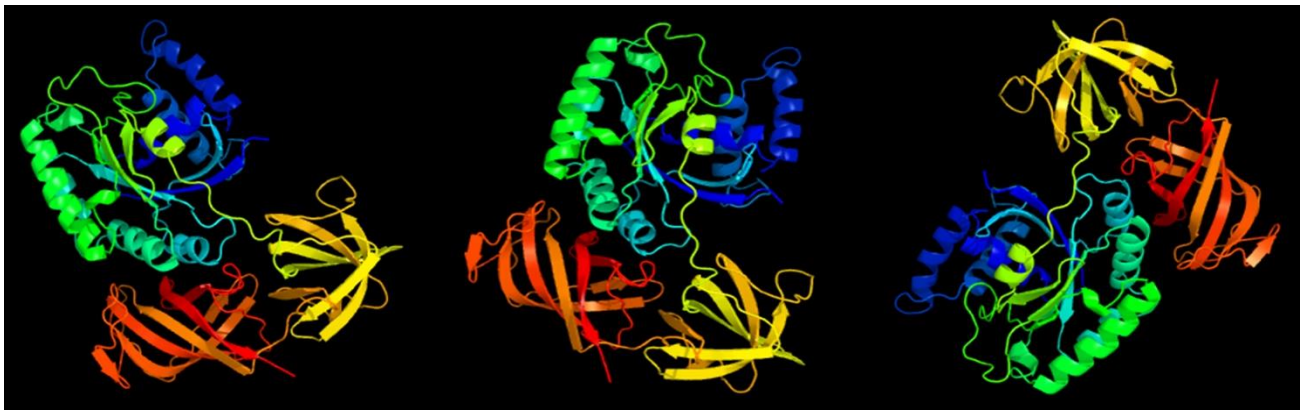

Fig. 4. 3D structure of translation elongation factor alpha 1 protein in *S. cerevisiae* (Left), SFS (Center) and NFS (Right) of *S. fimicola* Image coloured by rainbow N → C terminus *S. cerevisiae* = Model dimensions (Å): X: 77.549 Y: 51.010 Z: 69.994 SFS= Model dimensions (Å): X: 77.549 Y: 51.010 Z: 69.994 NFS= Model dimensions (Å): X: 78.134 Y: 51.010 Z: 69.994.

Present work investigated O-glycosylation at Serine and Threonine residues in tef2 protein of *S. fimicola*. Eight *O-glycosylated* sites were predicted in *S. cerevisiae* as well as in NFS strains while nine *O-glycosylated* sites were investigated in SFS strains. Six *O-glycosylated* sites (107S, 185T, 237S, 284T, 285T, 449T) were found to be conserved among reference organisms, SFS strains, and NFS strains. Gonzalez *et al*., (2012) studied *O-glycosylation* in secretory proteins of eight fungal genomes. Bukhari *et al*., (2020) also reported glycosylation sites in *S. fimicola* by using YinOYang 1.2 server.

Lysine acetylation is a ubiquitous kind of PTMs of proteins and is known to play a key role in broad-spectrum cellular functions (Kim *et al*., 2006 & Zhao *et al*., 2010). In the TEF2 protein of *S. cerevisiae,* acetylation was predicted at 21 Lysine residues. In SFS strains, Lysine acetylation was predicted at 20 sites, and in NFS strains, Lysine acetylation ion was predicted at 24 sites. Single acetylation site (225K) was found to be conserved in the organism, SFS strains, and NFS strains.

In *S. cerevisiae,* there are 166 methyltransferases found when methyltransferase genes are searched in the *S. cerevisiae* genome database (White, 2019). A great number of the identified methyltransferases in the yeast *S. cerevisiae* modify ribosomal components, indicating that methylation of translational components is significant for cellular functions (Clarke, 2013). The current study reported that Arginine methylation at 3 sites of *S. cerevisiae*, SFS, and NFS strains in TEF2 protein. Single methylation site 7R was found to be conserved between *S. cerevisiae*, SFS strains, and NFS strains.

Conclusion

The present work was performed to detect polymorphism and post-translational modifications by using different computational techniques in the TEF2 gene in *S. fimicola*. Significant differences in the gene sequences and their reflections upon post-translational modifications of proteins were observed. So, it is found that genetic variation may occur under stressful environmental conditions in nucleotide sequence of tef 2 genes. This is shown by the presence of some different and some conserved PTM sites for translation elongation factor alpha 1 in each strain of *S. fimicola*.

References

Andersen, G.R., P. Nissen and J. Nyborg. 2003. Elongation factors in protein biosynthesis. *Trends Biochem. Sci*., 28(8): 434-441.

Arif, R., F. Akram, T. Jamil, H. Mukhtar, S.F. Lee and M. Saleem. 2017. Genetic variation and its reflection on posttranslational modifications in frequency clock and mating type a-1 proteins in *Sordaria fimicola*. *Biomed Res. Int.*, (1): 1268623.

Bukhari, S.H., I. Mobeen, U. Naureen, F. Akram, R. Arif, M.G. Shahid and M. Saleem. 2020. Analysis of genetic polymorphisms and post translational modifications of cytochrome C-1 in Sordaria fimicola. *Int. J. Agric. Biol*., 23(3): 675-680.

Cavallius, J., W. Zoll, K. Chakraburtty and W.C. Merrick. 1993. Characterization of yeast EF-1: non-conservation of post-translational modifications. *Biochim. Biophys. Acta.*, 1163(1): 75-80.

Chen, D., Q. Fu, J. Lin, C. Hu, N. Huang, K.X. Chang and Z. Liu. 2019. Gene synthesizing, expression and immunogenicity characterization of recombinant translation elongation factor 2 from Dermatophagoides farinae. *Mol. Med. Rep.*, 20(6): 5324-5334.

Clarke, S.G. 2013. Protein methylation at the surface and buried deep: thinking outside the histone box. *Trends Biochem. Sci.*, 38(5): 243-252.

Cohen, P. 2002. The origins of protein phosphorylation. *Nat. Cell Biol*., 4(5): 127-130.

Dever T.E., J.D. Dinman and R. Green. 2018. Translation Elongation and Recoding in Eukaryotes. *Cold Spring Harb. Perspect Biol*., 10(8): a032649.

Gonzalez, M., N. Brito and C. Gonzalez. 2012. High abundance of Serine/Threonine-rich regions predicted to be hyper-O-glycosylated in the secretory proteins coded by eight fungal genomes. *BMC Microbiol.*, 12(1): 1-10.

Hoffman, A.A. and P.A. Parsons. 1994. Evolutionary genetics and environmental stress. *J. Evol. Biol.*, 7(5): 634-634.

Hosid, E., I. Grishkan, Z. Frenkel, S.P. Wasser, E. Nevo and A. Korol. 2008. Diversity of microsatellites in natural populations of ascomycetous fungus, *Emericella nidulans*, in Israel on local and regional scales. *Mycol. Prog*. (7): 99-109.cohenst

Khacho, M., K. Mekhail, K. Pilon-Larose, A. Pause, J. Cote and S. Lee. 2008. eEF1A is a novel component of the mammalian nuclear protein export machinery. *Mol. Biol. Cell.*, 19(12): 5296-5308.

Kim, S.C., R. Sprung, Y. Chen, Y. Xu, H. Ball, J. Pei and Y. Zhao. 2006. Substrate and functional diversity of lysine acetylation revealed by a proteomics survey. *Mol. Cell*., 23(4): 607-618.

Nevo, E. 1995. Asian, African and European biota meet at 'Evolution Canyon' Israel: local tests of global biodiversity and genetic diversity patterns. *Proc. R. Soc. Lond.*, 262(1364): 149-155.

Parsons, P.A. 2005. Environments and evolution: interactions between stress, resource inadequacy and energetic efficiency. *Biol. Rev*., 80(4): 589-610.

Pietro S, T.M. Fulton, J. Chunwongesm and S.D. Tanksley. 1995. Extraction of high-quality DNA for Genome Sequencing. *Mol. biol. Res*., 13: 207. doi:10.1007/BF02670897.

Pinna, L.A., and M. Ruzzene, 1996. How do protein kinases recognize their substrates?. *Biochim. Biophys. Acta. Mol. Cell. Res*., 1314(3): 191-225.

Riis, B., S.I. Rattan, B.F. Clark and W.C. Merrick. 1990. Eukaryotic protein elongation factors. *Trends Biochem. Sci*., 15(11): 420-424.

Shi, S., L. Wang, M. Cao, G. Chen and J. Yu. 2019. Proteomic analysis and prediction of amino acid variations that influence protein posttranslational modifications. *Brief. Bioinform.*, 20(5): 1597-1606.

Traugh, J.A. 2001. Insulin, phorbol ester and serum regulate the elongation phase of protein synthesis. *Prog. Mol. Subcell. Biol*., 26: 33-48.

Weber, K.E. and L. Diggins. 1990. Increased selection response in larger populations. II. Selection for ethanol vapor resistance in Drosophila melanogaster at two population sizes. *Genetics*, 125(3): 585-597.

White, J. 2019. A Quest for purpose: investigating the functional implication of the methylation of proteins involved in translation. University of California, Los Angeles.

Whiteheart, S.W., P.O.N.N.I.A.H. Shenbagamurthi, L.I.N.G. Chen, R.J. Cotter and G.W. Hart. 1989. Murine elongation factor 1α (EF-1α) is posttranslationally modified by novel amide-linked ethanolamine-phosphoglycerol moieties: addition of ethanolamine-phosphoglycerol to specific glutamic acid residues on EF-1α. *J. Biol. Chem.*, 264(24): 14334-14341.

Zhao, S., W. Xu, W. Jiang, W. Yu, Y. Lin, T. Zhang, J. Yao, L. Zhou, Y. Zeng, H. Li and Y. Li. 2010. Regulation of cellular metabolism by protein lysine acetylation. *Science*, 327(5968): 1000-1004.

Zobel-Thropp, P., M.C. Yang, L. Machado and S. Clarke. 2000. A novel post-translational modification of yeast elongation factor 1A: methyl esterification at the C terminus. *J. Biol. Chem*., 275(47): 37150-37158.

(Received for publication 11 January 2024)