Pak. J. Bot., 57(5): 1691-1700, 2025. DOI: http://dx.doi.org/10.30848/PJB2025-5(23)

MOLECULAR AND COMPUTATIONAL CHARACTERIZATION OF VARIATION IN TRANSLATION ELONGATION FACTOR METHYLTRANSFERASE *EFM6* OF *SORDARIA FIMICOLA* AND ITS FFECT ON POST-TRANSLATIONAL MODIFICATIONS

MAIMOONA ILYAS[1,2*] AND MUHAMMAD SALEEM[1*]

[1]*Institute of Botany, University of the Punjab, Lahore, Pakistan*
[2]*Department of Genetics and Genome Biology, University of Leicester, Leicester LE1 7RH, UK*
**Corresponding author's email: maimoona.ilyas111@gmail.com, saleem.botany@pu.edu.pk*

Abstract

Post-translational modifications (PTMs) are important alterations of the proteome that make species more compatible with changing environments. The current study was designed to reveal the importance of PTMs and genetic polymorphism of the elongation factor methyltransferase 6 (*efm6*) in six different strains of *Sordaria fimicola* through different molecular techniques and computational tools. The *efm6* is involved in the methylation of S-adenosyl-L-methionine (SAM) at the N-terminus of the protein which is an important modification for the survival of the species. *S. fimicola* strains were collected, three from South south-facing slope (S1, S2, S3) and three from North north-facing slope (N5, N6, N7) of Evolution Canyon, Israel. DNA was extracted from all six strains and sequences were obtained after amplification of genes. A total of 56 nucleotide variations in all six strains were observed compared to the reference *Sordaria macrospora*. There were 7 inter-strain variations observed when strains were compared with each other. Seven inter-strain polymorphic sites were observed in one of SFS strains S3. The predicted putative 3D structure of the EFM6 of *S. fimicola* revealed that it is consistent with the known structure of SAM. Acetylation sites were predicted at 7 similar positions in both SFS and NFS strains. Glycosylation was predicted on asparagine residues, with N-glycosylation at 5 sites in all strains of *S. fimicola* while O-glycosylation was predicted on 3 Ser residues in all strains. Potential phosphorylation was determined with 15 sites in all strains respectively. Protein network analysis showed that *efm6* has 11 biologically active interactions with other genes that work as a group inferred from *S. macrospora*. This whole molecular and bioinformatics analysis generated valuable knowledge that stressed environmental conditions are responsible for genetic variations among species and despite inter-species heterogeneity and complexity, the EFM6 is conserved in almost all eukaryotes.

Key words: PTMs of EFM6, Genetic polymorphism of SFS and NFS strains, EFM 6 gene in *S. fimicola*, Non-histone Methyltransferase.

Introduction

Gene expression and regulation are continuously reshaping with emerging epigenetic modulators as methylation of lysine (Lys) and arginine (Arg) residues plays a crucial role in the transfer of methyl groups to pivotal biomolecules like DNA, RNA and proteins (Boriak-Sjodin & Swinger, 2016; Huang, 2019). Protein methylation in eukaryotic cells is rising as one of the most influential post-translational modifications (Khoury *et al*., 2011). Protein methylation helps in understanding protein-protein interactions as it shows high evolutionary conservation (Erce *et al*., 2012). Mostly known methyltransferases specifically target histone proteins but many new methyltransferases have been reported in recent decades that are specific for non-histones and ribosomal proteins as well (Clarke, 2013). In non-histone proteins, methylation shows interplay with modifications like phosphorylation and acetylation (Jenuwein *et al*., 2001) and it has also been reported for protein interaction codes but still, there is a knowledge gap in understanding the pathways of methyltransferases that catalyze protein methylation (Gu & Zhu, 2012).

Protein methylation primarily ensues on Arg and Lys residues but is also known to occur on cysteine, histidine, glutamine, glutamate, asparagine and N- and C termini of proteins (Schaner *et al*., 2010; Webb *et al*., 2010). Although N-terminal methylation on various protein residues is known for long periods enzymes responsible for this methylation have been reported recently (Stock *et al*., 1987). Elongation factor methyltransferases (EFMs) catalyze the reaction between non-histone protein substrate usually ribosomal subunits/translation elongation factors and S-adenosyl-L-methionine (SAM)-the methyl donor, subsequently producing methylated protein and S-adenosyl-L-homocysteine (SAH) (Petrossian & Clark, 2011; Dzialo *et al*., 2014). *Saccharomyces cerevisiae* is considered a model organism for studying methylation of translation elongation factors as it has many modification sites and these sites are conserved in higher eukaryotes as well (Cavallius *et al*., 1993; Couttas *et al*., 2012). There are three evolutionary conserved elongation factors (EF1A, EF2 and EF3) that have been recognized in *S. cerevisiae*. These three proteins lead tRNA to ensure correct codon match at different active sites of ribosomes. EF1A directs exact codon matching between aminoacyl-tRNA and mRNA, while EF2 and EF3 assist in appropriate translocation and elimination of peptidyl-tRNA and deacylated-tRNAs, respectively (Couttas *et al*., 2012). Although methyltransferases involved in catalyzing these modifications have been recognized but practical significance of these modifications is still largely undiscovered (Lipson *et al*., 2010). *Ynl024c* in *Saccharomyces cerevisiae* has been renamed as *efm6* as it helps in methylation of Lys-390 residue in eEF1A in wild type of yeast (Jakabsson *et al*., 2015). It has also been reported as *80A10* (80A10.400) found in cosmid contig of *Neurospora crassa* while in *Sordaria macrospora* as (*SMAC_06518*), based on structural topology and its functional homology both *80A10* and *SMAC_06518* are named as *efm6* (NCBI, UniProt). The *efm6* belongs to the Class I superfamily of methyltransferases having a non-SET domain of the Rossmann-like fold of seven β-sheets connected with α-helices. It provides specific nucleotide binding sites not only for small Lys, Arg residues and co-

factors but also for larger molecules like DNA and RNA (Li *et al*., 2019**).** EFM6 is localized in the cytoplasm where it ensures rRNA-mediated translation and assists in important cellular functions including cytoskeletal organization along with accompanying ER to induce certain drugs under stressed conditions (Jakobsson *et al*., 2017).

The valley "Evolution Canyon" (EC), Israel, has an amalgam of two distinct ecologies as one slope facing the southern side has harsh xeric tropical conditions, while the north-facing slope near Europe has cold, shady and temperate conditions (Nevo, 2012). Harsh ecological conditions constantly encourage gene mutations which eventually produce new genetic recombination and help in species adaptation (Hoffman & Hercus, 2000; Saleem *et al*., 2001). The current study was designed to predict post-translational modifications and in-silico proteomic analysis of the *efm6* gene in six different EC strains of *Sordaria fimicola* as it has been revealed that epigenetic changes and polymorphism both are required for species survival under harsh ecological conditions.

Methodology

Genomic DNA extraction: The official review board of the University of the Punjab (PU), Lahore, Pakistan, permitted this study. After approval, all Northern (N5, N6, N7) and Southern (S1, S2, S3) EC strains of *S. fimicola* were identified and collected by (Nevo *et al*., 2006) were cultured and sub-cultured in the Molecular Genetics Lab, Institute of Botany, PU. For this purpose, potato dextrose agar by (Sigma Aldrich) was used 3.9g/100 ml to prepare the growth media. The fully grown cultures were obtained after 14 days of incubation; afterward, cultures were reserved at 20°C for long-standing use. Then all six strains of *S. fimicola* were subjected to nuclear genome extraction through a modified fungal DNA extraction method (Peitro *et al*., 1995) followed by quantification *via* agarose gel electrophoresis.

PCR and gene sequencing: Forward (ATGTCCGCCC GCTCCCGC) and Reverse (TTATTTGCGGGTAAT CGCTACAGG) primers were designed to amplify the *efm6* with the help of SnapGene viewer. Conventional PCR was done at melting temperature (Tm) 65°C. The size of obtained PCR products was confirmed by running PCR amplicons against a 100 bp ladder on 1.5% agarose gel. After confirmation PCR products were cleaned up by ethanol precipitation and sent for Sanger sequencing to Source Bioscience based in Cambridge, UK. After the confirmation of sequences from blast, *efm6* of all strains were submitted to the GenBank and following accession numbers were assigned (S1- OR800283, S2- OR800284, S3- OR800285, N5- OR800286, N6- OR800287, N7- OR800288).

Bioinformatics and in silico proteomic study: Bioinformatics and in silico proteomic analysis of all six sequenced *efm6* were performed with the help of different online available tools. Results of sequencing were analyzed by MEGA-X for multiple sequence alignment. To identify polymorphism at the amino acids level, multiple protein sequence alignment was made from NCBI-COBALT (ncbi.nlm.nih.gov/tools/cobalt/). PTMs were observed through online servers the NetPhos 3.1 server (http://www.cbs.dtu.dk/services/NetPhos/) to observe phosphorylation at threonine (Thr), tyrosine (Tyr) and serine (Ser) residues. The GPS-MSP was used to predict methylation on Lys/ Arg residues. The PAIL (http://bdmpail.biocuckoo.org/prediction.php) server was exploited to study acetylation at internal Lys residues. Glycosylation sites were explored with the help of (http://www.cbs.dtu.dk/services/NetNGlyc/) for N-glycan sites and (http://www.cbs.dtu.dk/services/NetOGlyc/) for O-Glycan sites. Furthermore, protein structures were made by SWISS-Model (AlphaFold2) and the STRING database (https://string-db.org/) was screened to show functional protein interactions and expression of associated proteins in related organisms.

Results

Nucleotide multiple sequence alignment (NMSA): All the SFS (S1, S2, S3) and NFS (N5, N6 and N7) strains of *Sordaria fimicola* were subjected to multiple sequence alignment through MEGA-X to observe natural genetic variations with *Sordaria macrospora* which was used as reference. The nucleotide multiple sequence alignment (NMSA) of *efm6* gene of all six *Sordaria fimicola* strains was obtained from MEGA-X by aligning them against *efm6* (SMAC_06518) of *Sordaria macrospora*. A total no. of 58 nucleotide variations was observed when the *S. fimicola* strains were compared with *S. macrospora* throughout the alignment at positions 4, 5, 12, 15, 18, 21, 22, 24, 27, 30, 36, 39, 42, 43, 45, 78, 84, 96, 99, 105, 145, 159, 193, 222, 224 – 229, 231 – 234, 237, 239, 240, 241, 243, 245, 246, 248, 291, 308, 324, 375, 378, 390, 399, 408, 414, 415, 457, 526, 565, 619, 653 and 658. Inter-strain variations were also observed between the *S. fimicola* strains at 7 positions and all of them were observed in S3 one of SFS strains while rest of all five strains showed similarity with consensus sequence at these sites (Fig. 1).

The multiple sequence alignment data generated by MEGA-X revealed that the *efm6* of *S. fimicola* is conserved and genetic variations were observed at a total no. of 58 sites with respect to *S. macrospora* and 7 inter-strain variations. One of SFS strains (S3) showed more polymorphic sites as compared to the rest of strains. The findings from the NMSA indicated that while the efm6 is conserved, stressed environmental factors are naturally causing variations in SFS strains of *S. fimicola*.

Protein multiple sequence alignment (PMSA): The protein multiple sequence alignment (PMSA) of EFM6 domain of SFS and NFS strains of *Sordaria fimicola* was obtained from COBALT NCBI by aligning them with SMAC_06518 protein sequence of *S. macrospora*. At 20 positions including 26, 49, 65, 75, 76, 77, 78, 79, 80, 81, 82, 83, 103, 138, 152, 176, 189, 207, 218, and 219 amino acid codes were changed in all *S. fimicola* strains when they were compared with *S. macrospora* resulted from 58 nucleotide variations. The inter-strain variations were observed only at two positions in S3 strain of S. fimicola including 76^{th} and 78^{th} while the remaining part of the domain was highly conserved (Fig. 2).

Despite the fact that 58 nucleotide variations altered the codes of 20 amino acids that were seen during species comparison, the EFM6 is highly conserved, as were the other 213 amino acids.

```
SMAC_06518      ATGAGCGCCCGATCCCGATCATTGTCACCGGAGTTTGATCCATTGGCCTTCGACCAGGAT    60
S3_OR800285     ATGTCCGCCCGCTCGCGCTCCCTCTCCCCGGAGTTCGACCCGCTCGCCTTCGACCAGGAT    60
S1_OR800283     ATGTCCGCCCGCTCCCGCTCCCTCTCCCCCGAGTTCGACCCCCTCGCCTTCGACCAGGAT    60
N6_OR800287     ATGTCCGCCCGCTCCCGCTCCCTCTCCCCCGAGTTCGACCCCCTCGCCTTCGACCAGGAT    60
S2_OR800284     ATGTCCGCCCGCTCCCGCTCCCTCTCCCCCGAGTTCGACCCCCTCGCCTTCGACCAGGAT    60
N5_OR800286     TGTCCGCCCGCTCCCGCTCCCTCTCCCCCGAGTTCGACCCCCTCGCCTTCGACCAGGAT     60
N7_OR800288     ATGTCCGCCCGCTCCCGCTCCCTCTCCCCCGAGTTCGACCCCCTCGCCTTCGACCAGGAT    60
                *** ****** ** ** **  * ** ** ***** ** **  * ***************

SMAC_06518      TTCACTCCTCTCCCAGAGTACAAAGCTGCTACCGATACGGCTCTTGACTTTAATGGACTC    120
S3_OR800285     TTCACTCCTCTCCCAGATTACAAGGCTGCTACCGACACAGCTCTCGACTTTAATGGACTC    120
S1_OR800283     TTCACTCCTCTCCCAGATTACAAGGCTGCTACCGACACAGCTCTCGACTTTAATGGACTC    120
N6_OR800287     TTCACTCCTCTCCCAGATTACAAGGCTGCTACCGACACAGCTCTCGACTTTAATGGACTC    120
S2_OR800284     TTCACTCCTCTCCCAGATTACAAGGCTGCTACCGACACAGCTCTCGACTTTAATGGACTC    120
N5_OR800286     TTCACTCCTCTCCCAGATTACAAGGCTGCTACCGACACAGCTCTCGACTTTAATGGACTC    120
N7_OR800288     TTCACTCCTCTCCCAGATTACAAGGCTGCTACCGACACAGCTCTCGACTTTAATGGACTC    120
                ********************* ***** ********** ** ***** ***************

SMAC_06518      CTGCCAGAGCCTTTGAAGCTCCACCAAGATTTAAGGACAGGATGCGGAGGGCAACTTTGG    180
S3_OR800285     CTGCCAGAGCCTTTGAAGCTCCACGAAGATTTAAGGACGGGATGCGGAGGGCAACTTTGG    180
S1_OR800283     CTGCCAGAGCCTTTGAAGCTCCACGAAGATTTAAGGACGGGATGCGGAGGGCAACTTTGG    180
N6_OR800287     CTGCCAGAGCCTTTGAAGCTCCACGAAGATTTAAGGACGGGATGCGGAGGGCAACTTTGG    180
S2_OR800284     CTGCCAGAGCCTTTGAAGCTCCACGAAGATTTAAGGACGGGATGCGGAGGGCAACTTTGG    180
N5_OR800286     CTGCCAGAGCCTTTGAAGCTCCACGAAGATTTAAGGACGGGATGCGGAGGGCAACTTTGG    180
N7_OR800288     CTGCCAGAGCCTTTGAAGCTCCACGAAGATTTAAGGACGGGATGCGGAGGGCAACTTTGG    180
                ************************ ************ ********************

SMAC_06518      CCGGCCGGAATGACCCTCGCCAAGCATATGCTGCGCTACCACGCTGATAAGCTGCAAAAG    240
S3_OR800285     CCGGCCGGAATGGCCCTCGCCAAGCATATGCTGCGCTACCATGGACCGCACTTACATATC    240
S1_OR800283     CCGGCCGGAATGGCCCTCGCCAAGCATATGCTGCGCTACCATGGATTACACTCACATATC    240
N6_OR800287     CCGGCCGGAATGGCCCTCGCCAAGCATATGCTGCGCTACCATGGATTACACTCACATATC    240
S2_OR800284     CCGGCCGGAATGGCCCTCGCCAAGCATATGCTGCGCTACCATGGATTACACTCACATATC    240
N5_OR800286     CCGGCCGGAATGGCCCTCGCCAAGCATATGCTGCGCTACCATGGATTACACTCACATATC    240
N7_OR800288     CCGGCCGGAATGGCCCTCGCCAAGCATATGCTGCGCTACCATGGATTACACTCACATATC    240
                *********** **************************** *    *  ** *

SMAC_06518      GCTCGGATACTCGAGATTGGCGCCGGCGGTGGCCTTGTCGGGCTTGCAGTCGCCAAAGCC    300
S3_OR800285     TCCCCCAGACTCGAGATTGGCGCCGGCGGTGGCCTTGTCGGGCTTGCAGTTGCCAAAGCC    300
S1_OR800283     TCCCCCAGACTCGAGATTGGCGCCGGCGGTGGCCTTGTCGGGCTTGCAGTTGCCAAAGCC    300
N6_OR800287     TCCCCCAGACTCGAGATTGGCGCCGGCGGTGGCCTTGTCGGGCTTGCAGTTGCCAAAGCC    300
S2_OR800284     TCCCCCAGACTCGAGATTGGCGCCGGCGGTGGCCTTGTCGGGCTTGCAGTTGCCAAAGCC    300
N5_OR800286     TCCCCCAGACTCGAGATTGGCGCCGGCGGTGGCCTTGTCGGGCTTGCAGTTGCCAAAGCC    300
N7_OR800288     TCCCCCAGACTCGAGATTGGCGCCGGCGGTGGCCTTGTCGGGCTTGCAGTTGCCAAAGCC    300
                * *  * ****************************************** *********

SMAC_06518      TGCAGCTATGAAACTCCCATGTATATCACAGACCAACTGGAGATGGAAGAGCTCATGGCA    360
S3_OR800285     TGCAGCTTTGAAACTCCCATGTACATCACAGACCAACTGGAGATGGAAGAGCTCATGGCA    360
S1_OR800283     TGCAGCTTTGAAACTCCCATGTACATCACAGACCAACTGGAGATGGAAGAGCTCATGGCA    360
N6_OR800287     TGCAGCTTTGAAACTCCCATGTACATCACAGACCAACTGGAGATGGAAGAGCTCATGGCA    360
S2_OR800284     TGCAGCTTTGAAACTCCCATGTACATCACAGACCAACTGGAGATGGAAGAGCTCATGGCA    360
N5_OR800286     TGCAGCTTTGAAACTCCCATGTACATCACAGACCAACTGGAGATGGAAGAGCTCATGGCA    360
N7_OR800288     TGCAGCTTTGAAACTCCCATGTACATCACAGACCAACTGGAGATGGAAGAGCTCATGGCA    360
                ******* ************** ***********************************

SMAC_06518      CACAACATCACCCTGAATGGCCTCGACGATAAGGTCAAATCCATGATTCTCAACTGGGGC    420
S3_OR800285     CACAACATCACCCTCAACGGCCTCGACGACAAGGTCAAGTCCATGATCCTCAATAGGGGC    420
S1_OR800283     CACAACATCACCCTCAACGGCCTCGACGACAAGGTCAAGTCCATGATCCTCAATAGGGGC    420
N6_OR800287     CACAACATCACCCTCAACGGCCTCGACGACAAGGTCAAGTCCATGATCCTCAATAGGGGC    420
S2_OR800284     CACAACATCACCCTCAACGGCCTCGACGACAAGGTCAAGTCCATGATCCTCAATAGGGGC    420
N5_OR800286     CACAACATCACCCTCAACGGCCTCGACGACAAGGTCAAGTCCATGATCCTCAATAGGGGC    420
N7_OR800288     CACAACATCACCCTCAACGGCCTCGACGACAAGGTCAAGTCCATGATCCTCAATAGGGGC    420
                ************** ** *********** ******** ******** *****  *****

SMAC_06518      GAGCCACTCCCAGCGGAAATCGTTGCTCTCAAGCCCAATACTATTCTGGCTGCTGATTGC    480
S3_OR800285     GAGCCACTCCCAGCGGAAATCGTTGCTCTCAAGCCCGATACTATTCTGGCTGCTGATTGC    480
S1_OR800283     GAGCCACTCCCAGCGGAAATCGTTGCTCTCAAGCCCGATACTATTCTGGCTGCTGATTGC    480
N6_OR800287     GAGCCACTCCCAGCGGAAATCGTTGCTCTCAAGCCCGATACTATTCTGGCTGCTGATTGC    480
S2_OR800284     GAGCCACTCCCAGCGGAAATCGTTGCTCTCAAGCCCGATACTATTCTGGCTGCTGATTGC    480
N5_OR800286     GAGCCACTCCCAGCGGAAATCGTTGCTCTCAAGCCCGATACTATTCTGGCTGCTGATTGC    480
N7_OR800288     GAGCCACTCCCAGCGGAAATCGTTGCTCTCAAGCCCGATACTATTCTGGCTGCTGATTGC    480
                ********************************** ***********************
```

```
SMAC_06518    GTCTACTTCGAACCCGCCTTCCCGCTCCTCCTCCAGACCCTGAAGGGTCTATTGGCCCTT  540
S3_OR800285   GTCTACTTCGAACCCGCCTTCCCGCTCCTCCTCCAGACCCTGAAGGATCTATTGGCCCTT  540
S1_OR800283   GTCTACTTCGAACCCGCCTTCCCGCTCCTCCTCCAGACCCTGAAGGATCTATTGGCCCTT  540
N6_OR800287   GTCTACTTCGAACCCGCCTTCCCGCTCCTCCTCCAGACCCTGAAGGATCTATTGGCCCTT  540
S2_OR800284   GTCTACTTCGAACCCGCCTTCCCGCTCCTCCTCCAGACCCTGAAGGATCTATTGGCCCTT  540
N5_OR800286   GTCTACTTCGAACCCGCCTTCCCGCTCCTCCTCCAGACCCTGAAGGATCTATTGGCCCTT  540
N7_OR800288   GTCTACTTCGAACCCGCCTTCCCGCTCCTCCTCCAGACCCTGAAGGATCTATTGGCCCTT  540
              ********************************************* ************

SMAC_06518    GAACCCAACGCCACCGTTTACTTCTGTTTCAAGAAGAGACGCCGCGCAGACATGCAGTTT  600
S3_OR800285   GAACCCAACGCCACCGTTTACTTCAGTTTCAAGAAGAGACGCCGCGCAGACATGCAGTTT  600
S1_OR800283   GAACCCAACGCCACCGTTTACTTCAGTTTCAAGAAGAGACGCCGCGCAGACATGCAGTTT  600
N6_OR800287   GAACCCAACGCCACCGTTTACTTCAGTTTCAAGAAGAGACGCCGCGCAGACATGCAGTTT  600
S2_OR800284   GAACCCAACGCCACCGTTTACTTCAGTTTCAAGAAGAGACGCCGCGCAGACATGCAGTTT  600
N5_OR800286   GAACCCAACGCCACCGTTTACTTCAGTTTCAAGAAGAGACGCCGCGCAGACATGCAGTTT  600
N7_OR800288   GAACCCAACGCCACCGTTTACTTCAGTTTCAAGAAGAGACGCCGCGCAGACATGCAGTTT  600
              ********************** *********************************

SMAC_06518    TTCAAGGCTGCCAGAAAGACCTTCAAGATCACCGAGCTTGAGGACGAAGACCAGCCCGTG  660
S3_OR800285   TTCAAGGCTGCCAGAAAGTCCTTCAAGATCACCGAGCTTGAGGACGAAGACCGGCCCATG  660
S1_OR800283   TTCAAGGCTGCCAGAAAGTCCTTCAAGATCACCGAGCTTGAGGACGAAGACCGGCCCATG  660
N6_OR800287   TTCAAGGCTGCCAGAAAGTCCTTCAAGATCACCGAGCTTGAGGACGAAGACCGGCCCATG  660
S2_OR800284   TTCAAGGCTGCCAGAAAGTCCTTCAAGATCACCGAGCTTGAGGACGAAGACCGGCCCATG  660
N5_OR800286   TTCAAGGCTGCCAGAAAGTCCTTCAAGATCACCGAGCTTGAGGACGAAGACCGGCCCATG  660
N7_OR800288   TTCAAGGCTGCCAGAAAGTCCTTCAAGATCACCGAGCTTGAGGACGAAGACCGGCCCATG  660
              ***************** ********************************* **** **
```

Fig. 1. Nucleotide Multiple sequence alignment (NMSA) of *efm6* of all SFS (S1- OR800283, S2- OR800284, S3- OR800285) and NFS (N5- OR800286, N6- OR800287, N7- OR800288) strains of *Sordaria fimicola* constructed by MEGA-X. The NMSA of 660bp showed the genetic polymorphism at total no. of 56 and 7 sites in inter-species inter-strain comparisons respectively. The gray colored highlighted nucleotides represented the genetic polymorphism between *S. fimicola* and *S. macrospora* (the consensus sequence) while the blue colored highlighted nucleotides showed inter-strain polymorphism of *S. fimicola* strains.

Predicted 3D structure of *S. fimicola*: The predicted 3-D structure EFM6 of all SFS and NFS *S. fimicola* strains were prepared by using SWISS-MODEL (AlphaFold) and Phyre2 (Protein fold recognition) web tools. The (AF-F7W4K1-F1) of EFM6 of *S. macrospora* was used as reference model for structure comparison. The 3D protein structure was modeled through heuristic modeling, based on percentage identity and alignment coverage where 92% residues were aligned to the best templates with more than 90% confidence score to use it as reference. The models were retrieved with more than 70% sequence identity with *Posdospora anserina* and *Madurella mtcetomatis* (filamentous ascomycetes) fungi (Fig. 3). EFM6 also showed structural similarity (26.49%) with human methyltransferase like proteins METTL21D, fold library id (7oat.1. C) and METTL21B, fold library id (4qpn.1. A). Furthermore, the secondary structure prediction and domain analysis by Phyre 2 revealed that EFM6 domain has 25% alpha helix, 26% β-pleated sheets and 7% transmembrane helix (Fig. 3).

The EFM6 protein interaction network of *S. fimicola*: The input of the EFM6 sequences of *S. fimicola* in the STRING network database showed 91.3% to 91.8% identity with 427.9% Bitscore and 1.8 -120 e-value to the protein interaction network of *S. macrospora*. The STRING analysis was performed to annotate the efm6 in *S. fimicola* strains. All EFM6 of *S. fimicola* are 91.3% to 91.8% similar to F7W4K1 protein in many organisms particularly to ascomycetes and with preferred name protein lysine N-terminal methyltransferase and annotated it as S-adenosyl-L-methionine-dependent protein-lysine N-methyltransferase that methylates elongation factor 1-alpha. It belongs to the class I-like SAM-binding methyltransferase superfamily, METTL21 family, EFM6 subfamily (Table 1). The network of interacting proteins obtained had 11 no. of nodes with 20 edges, in which the colored nodes showed the first shell of interaction and white nodes indicated the second shell of interaction of EFM6 with other proteins. The filled nodes of interacting proteins presented information about the predicted 3D protein structure of that particular protein as shown in cluster-1 SMAC_00190 in *S. macrospora* first shell of interactions while rest of proteins in this network did not have predicted secondary structure. The result of the protein network of the EFM6 of *S. macrospora* is significantly or partially active and biologically it is connected with other proteins as a group. Different colors of lines of edges indicated both physical and functional association of interacting proteins, the blue colored line between two nodes indicated interaction based on curated database. Bright blue and black colors presented protein co-occurrence and co-expression, respectively. Green, sky blue and purple colored lines showed protein neighborhood, protein homology and experimentally determined data, respectively. From the resemblance of the EFM6 protein of *S. fimicola* and *S. macrospora* we can speculate that EFM6 in *S. fimicola* may also have such an interaction network as all proteins presented in the STRING network were putatively homologue to each other. EFM5 and EFM7 were more closely related to EFM6 as they all share SAM-dependent methyltransferase at N- terminus and were involved in mono, di and tri-methylation of glycine and lysine residues of elongation factor-1- alpha (Fig. 4).

```
Query_10001   1   MSARSRSLSPEFDPLAFDQDFTPLPEYKAATDTALDFNGLLPEPLKLHQDLRTGCGGQLWPAGMTLAKHMLRYHADKLQK  80
Query_10002   1   MSARSRSLSPEFDPLAFDQDFTPLPDYKAATDTALDFNGLLPEPLKLHEDLRTGCGGQ-WPAGMALAKHMLRYHGLHSHI  79
Query_10003   1   MSARSRSLSPEFDPLAFDQDFTPLPDYKAATDTALDFNGLLPEPLKLHEDLRTGCGGQLWPAGMALAKHMLRYHGLHSHI  80
Query_10004   1   MSARSRSLSPEFDPLAFDQDFTPLPDYKAATDTALDFNGLLPEPLKLHEDLRTGCGGQLWPAGMALAKHMLRYHGPHLHI  80
Query_10005   1   MSARSRSLSPEFDPLAFDQDFTPLPDYKAATDTALDFNGLLPEPLKLHEDLRTGCGGQLWPAGMALAKHMLRYHGLHSHI  80
Query_10006   1   MSARSRSLSPEFDPLAFDQDFTPLPDYKAATDTALDFNGLLPEPLKLHEDLRTGCGGQLWPAGMALAKHMLRYHGLHSHI  80
Query_10007   1   MSARSRSLSPEFDPLAFDQDFTPLPDYKAATDTALDFNGLLPEPLKLHEDLRTGCGGQLWPAGMALAKHMLRYHGLHSHI  80

Query_10001  81   ARILEIGAGGGLVGLAVAKACSYETPMYITDQLEMEELMAHNITLNGLDDKVKSMILNWGEPLPAEIVALKPNTILAADC  160
Query_10002  80   SPRLEIGAGGGLVGLAVAKACSFETPMYITDQLEMEE-MAHNITLNGLDDKVKSMILNRGEPLPAEIVALKPDTILAADC  158
Query_10003  81   SPRLEIGAGGGLVGLAVAKACSFETPMYITDQLEMEELMAHNITLNGLDDKVKSMILNRGEPLPAEIVALKPDTILAADC  160
Query_10004  81   SPRLEIGAGGGLVGLAVAKACSFETPMYITDQLEMEELMAHNITLNGLDDKVKSMILNRGEPLPAEIVALKPDTILAADC  160
Query_10005  81   SPRLEIGAGGGLVGLAVAKACSFETPMYITDQLEMEELMAHNITLNGLDDKVKSMILNRGEPLPAEIVALKPDTILAADC  160
Query_10006  81   SPRLEIGAGGGLVGLAVAKACSFETPMYITDQLEMEELMAHNITLNGLDDKVKSMILNRGEPLPAEIVALKPDTILAADC  160
Query_10007  81   SPRLEIGAGGGLVGLAVAKACSFETPMYITDQLEMEELMAHNITLNGLDDKVKSMILNRGEPLPAEIVALKPDTILAADC  160

Query_10001 161   VYFEPAFPLLLQTLKGLLALEPNATVYFCFKKRRRADMQFFKAARKTFKITELEDEDQPVFTRQGLFLYAITRk  234
Query_10002 159   VYFEPAFPLLLQTLKDLLALEPNATVYFSFKKRRRADMQFFKAARKSFKITELEDEDRPMFTRQGLFLYAITR-  231
Query_10003 161   VYFEPAFPLLLQTLKDLLALEPNATVYFSFKKRRRADMQFFKAARKSFKITELEDEDRPMFTRQGLFLYAIP--  232
Query_10004 161   VYFEPAFPLLLQTLKDLLALEPNATVYFSFKKRRRADMQFFKAARKSFKITELEDEDRPMFTRQGLFLYAITA-  233
Query_10005 161   VYFEPAFPLLLQTLKDLLALEPNATVYFSFKKRRRADMQFFKAARKSFKITELEDEDRPMFTRQGLFLYAITR-  233
Query_10006 161   VYFEPAFPLLLQTLKDLLALEPNATVYFSFKKRRRADMQFFKAARKSFKITELEDEDRPMFTRQGLFLYAITA-  233
Query_10007 161   VYFEPAFPLLLQTLKDLLALEPNATVYFSFKKRRRADMQFFKAARKSFKITELEDEDRPMFTRQGLFLYAITR-  233
```

Fig. 2. The protein multiple sequence alignment (PMSA) attained from Constraint based multiple sequence alignment tool (NCBI-COBALT) to check the conserved domain and local sequence similarity. EFM6 of SFS (S1-Query_10002, S2-Query_10003, S3-Query_10004) and NFS (N5-Query_10005, N6-Query_10006, N7-Query_10007) strains of *Sordaria fimicola* aligned with *Sordaria macrospora* (SMAC_06518-Query_10001) used as consensus sequence for comparison. The amino acid code changed at 20 sites in all *S. fimicola* strains when compared with consensus. Only two amino acids were changed in (S3- OR800285) that showed EFM6 is highly conserved domain.

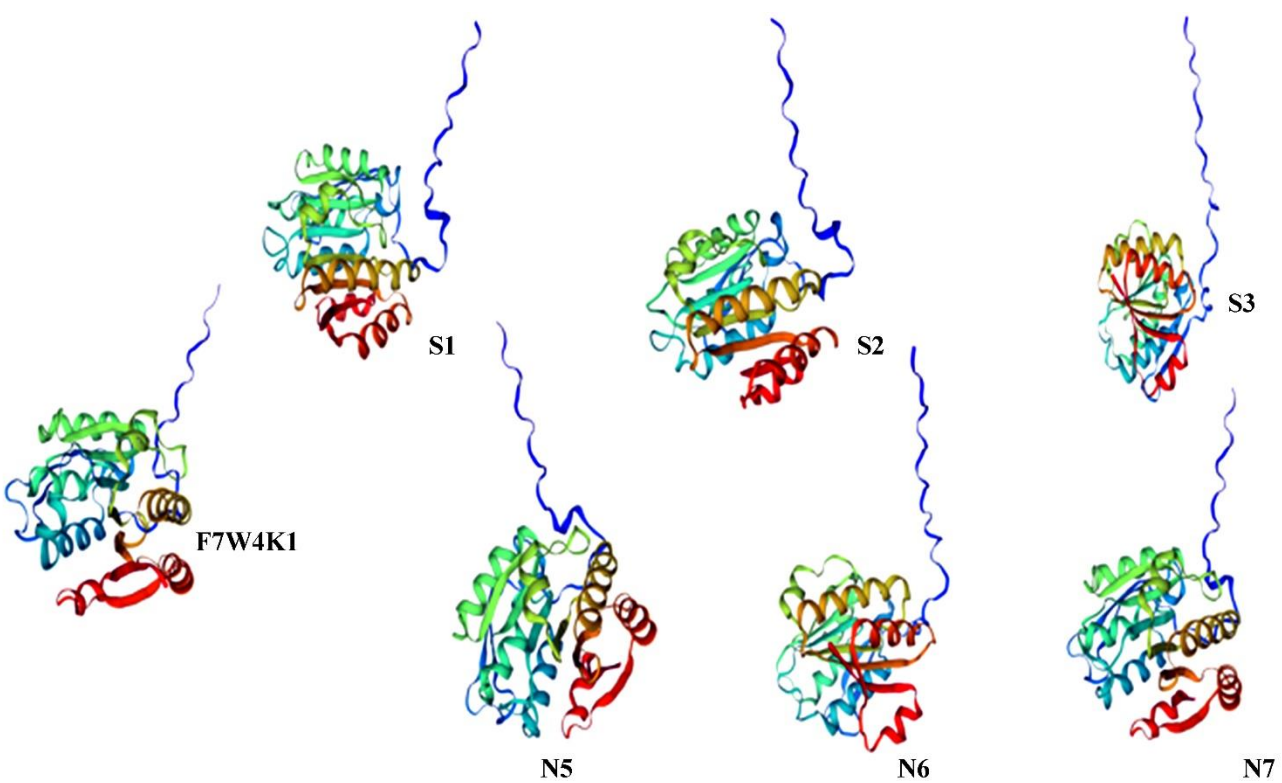

Fig. 3. The putative 3-D protein models EFM6 of all SFS (S1, S2, S3) and NFS (N5, N6, N7) strains of *Sordaria fimicola* inferred from SWISS-Model (AlphaFold2) by using F7W4K1 of *Sordaria macrospora* as reference model for comparison. The structural assessment of all EFM6 models of *S. fimicola* have Molprobity score 1.04, along with 95% of amino acids lie in Ramachandran favored regions with 0.68 clash score.

Table 1. Obtained from STRING analysis of EFM6 of all strains of *Sordaria fimicola* to show the fact that *efm6* sequenced from *S. fimicola* strains codes for EFM6 protein and belongs to SAM-binding Mtases superfamily, METTL21 family and EFM6 subfamily.

Query item	SRING id	Identity	Bitscore	E-value	Annotation
S1_OR800283	F7W4K1	91.4	427.9%	1.8e-120	SAM-binding Mtases superfamily, METTL21 family, EFM6 subfamily
S2_OR800284	F7W4K1	91.3	422.5%	7.3e-119	SAM-binding Mtases superfamily, METTL21 family, EFM6 subfamily
S3_OR800285	F7W4K1	91.8	429.9%	4.6e-121	SAM-binding Mtases superfamily, METTL21 family, EFM6 subfamily
N5_OR800286	F7W4K1	91.4	427.9%	1.8e-120	SAM-binding Mtases superfamily, METTL21 family, EFM6 subfamily
N6_OR800287	F7W4K1	91.4	426.0%	6.7e-120	SAM-binding Mtases superfamily, METTL21 family, EFM6 subfamily
N7_OR800288	F7W4K1	91.4	427.9%	1.8e-120	SAM-binding Mtases superfamily, METTL21 family, EFM6 subfamily

Table 2. for details of four different types of post-translational modification sites in both *Sordaria macrospora* and *Sordaria fimicola* strains highlighted regions same sites in both species.

	Residues sits of PTMs		
	Sordaria macrospora	***S. fimicola** (SFS Strains)**	***S. fimicola** (NFS Strains)**
Acetylation	77K, 99K, 131K, 192K, 202K, 206K, 209K, 234K	99K, 131K, 133K, 151K, 175K, 191K, 192K, 202K, 206K, 209K	99K, 131K, 133K, 151K, 175K, 191K, 192K, 202K, 206K, 209K
Methylation	131K, 151K, 233K	46K, 72R, 83R, 99K, 131K, 151K, 206K, 233R	46K, 72R, 83R, 99K, 131K, 151K, 206K, 233R
N-Glycosylation	38N, 122N, 126N, 138N,153N, 183N	38N, 122N, 126N, 138N, 183N	38N, 122N, 126N, 138N, 183N
O-Glycosylation	5S, 22T	2S, 5S, 7S, 22T	2S, 5S, 7S
Phosphorylation	2S, 5S, 7S, 9S, 22T, 27Y, 65T, 105T, 108Y, 110T, 124T, 162Y, 207T, 211T, 232T	2S, 5S, 7S, 9S, 22T, 27Y, 78T, 81S, 105T, 134S, 162Y, 173T, 189S, 207T, 211T	2S, 5S, 7S, 9S, 22T, 27Y, 78T, 81S, 105T, 134S, 162Y, 173T, 189S, 207T, 211T

K= Lysine, N= Asparagine, NFS= North facing slope, PTM= Post-translational modification, R= Arginine, S= Serine, **S*= *Sordaria*, SFS= South facing slope, T= Threonine, Y= Tyrosine. Yellow-colored highlighted positions are similar in both species

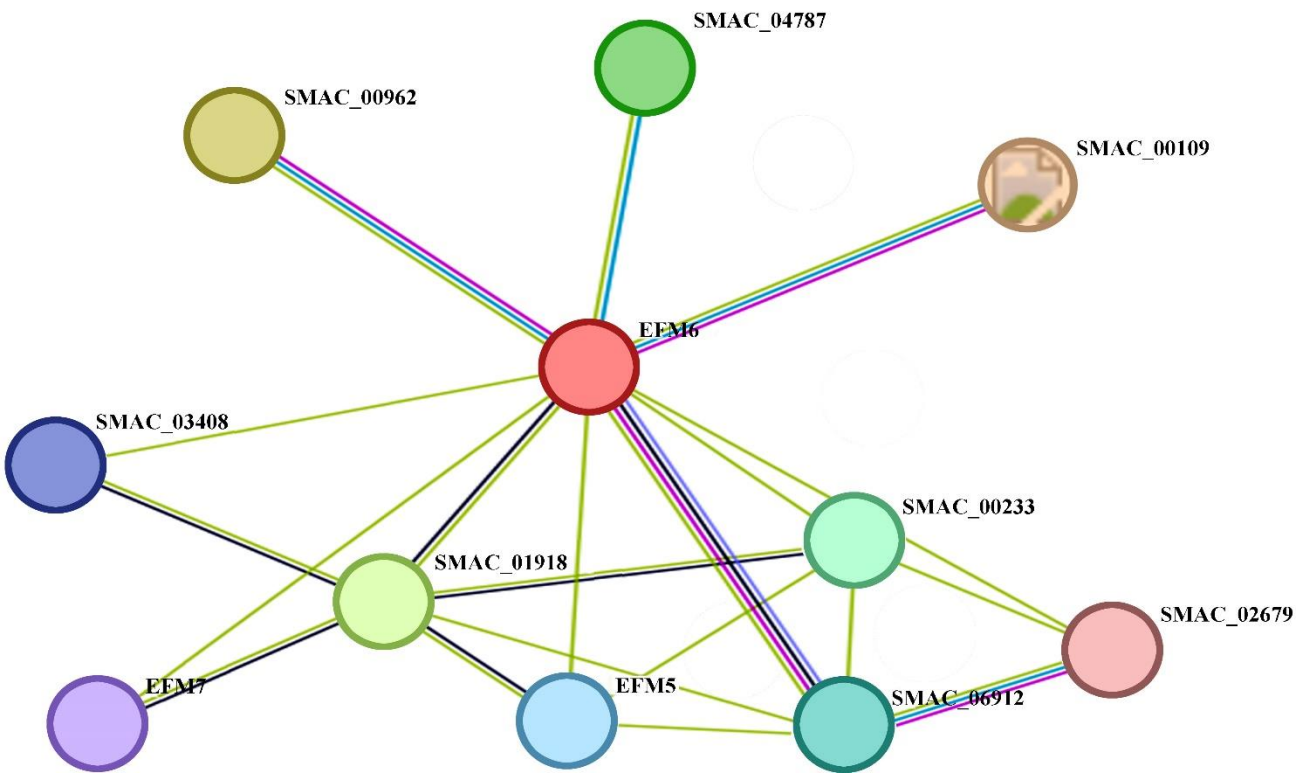

Fig. 4. The predicted biological interaction network of EFM6 of *S. fimicola* was observed in *S. macrospora* based on functional and physical protein association. Cluster 1 covered EFM5, EFM7, SMAC_01918 and SMAC_03408. Cluster 2 comprised of EFM6, SMAC_00109, SMAC_00233, SMAC_00962 and SMAC_04787. Cluster 3 covered only two genes SMAC_02679 and SMAC_06912. The predicted no. of nodes and edges are 11 and 20 respectively with average node degree = 3.46, average local clustering coefficient = 0.851and PPI enrichment p-value: 0.00472. The black color edges show the gene co-expression that can be seen in cluster 2 and EFM6 protein has gene co-expression with SMAC_01918 (EFM4). The parrot color edges show the information obtained from text mining. The purple-colored edges depict the data obtained from biochemical/experimental analysis of putative homologs interacting in other organisms.

Post translational modifications (PTMs): Five types of PTMs of EFM6 of *Sordaria macrospora* and all six strains of *Sordaria fimicola* were predicted showed in Table 2. Acetylation on internal Lys residues was predicted on net 10 positions of EFM6 of *Sordaria fimicola* strains. In *S. macrospora* a total no. of 8 sites were observed which were similar to *S. fimicola* except Lys^{77} and Lys^{234} and *S. fimicola* showed 3 different potential sites for acetylation (Lys^{191}, Lys^{151} and Lys^{175}) which were absent in *S. macrospora*. All three types of methylations (mono, di and tri) were predicted on total no. of eight sites in SFS strains including 5 lysine and 3 arginine residues and on net seven sites in NFS strains counting 5 lysine and 2 arginine residues of *S. fimicola*. The Lys^{72} was found to be absent in NFS strains. In *S. macrospora* only 2 lysine and 1 arginine methylated sites were observed identical to all *S. fimicola* strains including 131K, 151K and 233R.

Net N-glycosylation was predicted on total no. of 5 Asn residues of all SFS and NFS strains of *S. fimicola* for EFM6. These all sites were similar to *S. macrospora* except 153N. As indicated in (Table 2) net O-glycosylation prediction was detected on a total of 4 sites, comprising 3 Ser and 1 Thr) in SFS strains of *S. fimicola* and only 3 Ser residues in NFS strains. Only two positions, Ser5 and Thr22, which were also found in *S. fimicola*, were predicted in the case of *S. macrospora*. In comparison to NFS strains, SFS strains displayed more potential sites for O-glycosylation. Nine Ser, six Thr, and three Tyr residues make up the total of eighteen sites on the EFM6 protein that are predicted to be phosphorylated across all *S. fimicola* strains. Except for S3, where Ser78 and Ser81 were missing, all the strains displayed comparable sites. Fifteen sites in total were found in *S. macrospora*; all of these, with the exception of Thr232, were also found in *S. fimicola* strains. In contrast, *S. macrospora* strains lacked Ser78, Ser81, Ser134, Thr173, and Ser189. Overall NFS strains showed more acetylated, o-glycosylated and phosphorylated sites as compared to SFS strains when compared with each other.

Discussion

Understanding of biodiversity in micro scale environment is essential to study natural genetic variations that occur with the passage of time and make the populations more adaptive in changing harsh abiotic conditions (Hoffmann & Parsons, 1991). Genetic diversity also helps to study phylogenetic distances among different species of the same population and between the strains of the same species which consequently assist in protein profiling of a particular gene for studies (Nevo *et al*., 2012). The multiple sequence alignment showed genetic polymorphism at 56 sites while a major part of the gene was conserved similarly multiple protein sequence alignment revealed a net 20 variation sites and the remaining part of the functional domain was identified as conserved region. The SFS strains showed more polymorphic sites as compared to the NFS strains, which reveals that these polymorphic sites are formed due to natural genetic variation under stressed environmental conditions. Saleem *et al*., (2001) studied the relationship between recombinants formed after crossing over and variations occurring at the gene level due to unfavorable surroundings among all *S. fimicola* strains. They concluded that all SFS strains have high rates of mutations due to harsh xeric environments as compared to NFS strains which are more conserved and collected from relatively moderate temperate environments. They also found that all the strains have highly conserved genome as when they crossed them with each other the matting results were not satisfactory.

efm6 is directly involved in the methylation of non-histone proteins during translation and it is a conserved domain from prokaryotes to higher eukaryotes is also proved by our results as 3D protein structure has also been first time reported in *S. fimicola* species. Borgo *et al*., (2022) forwarded a study on genetic screening of genes involved in the modification of outer membrane proteins in *Rickettsia parkeri* through two protein lysine methyltransferases to shield against autophagy. Owings *et al*., (2016) conducted a study on *Pseudomonas aeruginosa* and reported that elongation factor methyltransferase (EftM) is involved in trimethylation of elongation factor-Tu domain. The EftM tri-methylated Lys 5 of EF-Tu and this PTM improves surface adherence of bacteria to epithelial cells without halting the canonical function of the EF-Tu domain.

The presence of 7β-strands (7βS) indicated EFM6 belongs to non-histone proteins which are localized in cytoplasm and involved in translation and found to be highly conserved (Petrossian & Clarke, 2009). Flanes *et al*., (2016) published a review about non-histone KMtases; in their study they reported that non-histone methyltransferases structurally warped in 7βS were found to be different from set domain family. Most of them are involved in methylation of ribosomal proteins, lipids and secondary metabolites. There are so many non-histone lysine methyltransferases that have been discovered and characterized from yeast to human and were placed in the lysine methyltransferase family16 having a highly conserved domain. Another novel human 7βS methyltransferase has been discovered and characterized (Jakobsson *et al*., 2017). Researchers used different enzymology essays and gene knockout techniques to study the canonical role of non-histone methyltransferases in translation and found that new KMT4 is involved in methylation of Lys 36 residue of eukaryotic elongation factor 1 alpha (eEF1A). Gene knockout experiments revealed that the absence of KMT4 produced translational defects in the eEF1A domain.

Jakbosson *et al*., 2018 again published an article in which they reported another 7βS human methyltransferase METTL13. This enzyme was involved in the methylation of Lys 55 and the N-terminus of the eEF1A and eEF1A2 domains. They also found that METTL13 is very specific in its function as it allocates specific codons to the ribosomal complex during translation and any malfunctioning of this enzyme can alter the translation dynamics. Buuh *et al*., (2017) reported the effects of different pivotal PTMs on non-histone proteins such as

methylation, acetylation, phosphorylation, glycosylation and many others that are involved in the regulation of different cellular pathways important for survival and stability. They also reported that any misinterpretation or defect in these pathways can be proven hazardous and can result in abnormal cellular functioning of an organism.

The PTMs done by methyltransferases on Lys residues have functional significance in many cellular events. These modifications on different sites at the tail and globular parts of proteins are accurately precisely guided by the action of methylases (writers) and demethylases (erasers) with the help of effector proteins (readers) which identify particular Lys residues containing methyl group (Hyun *et al*., 2017). Post-translational modifications are important epigenetic marks that can enhance protein functional accuracy by altering its structure and consequently make it more diverse (Arif *et al*., 2019).

Phosphorylation is one the most important modifications involved in cell differentiation, growth, DNA damage repair activation system and signal transduction. Kinases and phosphatases are the two important enzymes involved in phosphorylation, repression or over expression of these enzymes can activate oncogenic activity of cells and leads towards tumor formation (Aridto *et al*., 2017; Bhal *et al*., 2021). Phosphorylation most commonly occurs on serine and threonine residues and least commonly on tyrosine comprising only 1.8% of total phosphorylated sites as reported in our study (Schwartz & Murray, 2011; Roskoski, 2012; Nishi *et al*., 2014). Phosphorylation on Tyr residues is domain specific as it only occurs in the presence of specific kinase epidermal growth factor receptor (EGFR) in eukaryotes. Histidine and aspartate residues are also sometimes phosphorylated but these modifications are usually weak and unstable (Huether & McCance, 2014).

Methylation is also another important PTM that mostly occurs on lysine and arginine residues. These residues are usually mono, di and tri methylated depending upon the targeted substrate. Methylation of proteins is essential for the regulation of many biological and physiological processes like cell metabolism and transcriptional regulation (Bedford & Richard, 2005; Chen *et al*., 2006; Web *et al*., 2010). Deng *et al*., (2017) used a large-scale data set of 5421 methylated lysine and arginine residues in 2592 proteins and predicted Lys/Arg methylated specific sites through GPS-MPS webserver that were difficult to identify with experimental techniques.

Acetylation occurs mostly on the N-terminus and internal lysine residues of proteins and is involved in the regulation of both physiological and pathological processes at cellular level. It is also involved in metabolism and signaling and these acetylated proteins are also known as acetylome (Choudhary *et al*., 2014; Dialo *et al*., 2018). Glycosylation is as important as other PTMs are, such as it is involved in the making of cell membrane frontier molecules like glycoproteins and glycolipids important for a cell's stability and function. Glycosylation mostly occurs on ser/thr residues in case of O-glycosylation and on asparagine in case of N-glycosylation. A lot of studies are available about histone glycosylation events but research work about ribosomal proteins is under progress and studies showed that this modification is very important in cancer studies as these epigenetic marks are directly or indirectly involved in tumor progression (Indellicota & Trinchera, 2021).

Zhang *et al*., (2015) reported a study on methylation of non-histone proteins and found that it is an important PTM for gene expression in eukaryotes. Methylated sites of non-histone proteins are highly conserved and most every so often are further modified by other PTMs including phosphorylation and acetylation (Reimand *et al*., 2015 and Zhang *et al*., 2018). In non-histone proteins, more than 60% of methylation oftenly occurs on ten phosphorylation sites and afterward, these sites are further modified through acetylation another type of PTM. (Evich *et al*., 2016). In comparison, methylation is a less complex chemical alteration than glycosylation and phosphorylation and, in this way, all these PTMs interact with each other's role and modify the same sites in the same region of the protein (Qamar *et al*., 2018 and Ferrari *et al*., 2020).

Conclusion and Suggestions

The novel non-histone 7β-strands *efm6* non-histone methyltransferase has been identified for the first time in six different strains of coprophilous fungus, *Sordaria fimicola* and studied at both genetic and proteomic levels to understand its role in methylation of proteins during translation. All SFS strains showed more polymorphic sites in contrast to NFS strains that were found to be more conserved and revealed the fact that a more severe xeric environment can accelerate natural genetic variation. The *efm6* was also conserved and specifically involved in the methylation of the eukaryotic elongation factor proved by its predicted protein structure and biological interactions. Moreover, predicted PTMs of EFM6 can also enhance its catalytic activity and improve its function by modifying its structure which should be endorsed by experimental techniques that can produce valuable data to study the effects of non-histone methyltransferases involved in translation of eukaryotic elongation factor eEF1A as *Sordaria fimicola* has also been used as model organism.

Funding and Acknowledgement

The funding for the current study was provided by the Higher Education Commission (HEC) Pakistan, and all research work was started in the Molecular Genetics Lab under the supervision of Prof. Dr. Muhammad Saleem, at the Institute of Botany, University of the Punjab, Pakistan and completed in the lab of Genetic Architecture, School of Genetics and Genome Biology, University of Leicester, UK. We are extremely grateful to Prof. Dr. Louis Edward for giving permission to perform this research work in his lab and for reviewing and commenting on the manuscript.

References

Ardito, F., M. Giuliani, D. Perrone, G. Troiano and L. Lo Muzio. 2017. The crucial role of protein phosphorylation in cell signaling and its use as targeted therapy. *Int. J. Mol. Med*., 40(2): 271-280.

Arif, R., S.H. Bukhari, M. Ishfaq, M.G. Shahid, S.F. Lee and M. Saleem. 2019. Genetic variation and post-translational modifications of cytochrome c oxidase-1 (COX1) in different strains of *Sordaria fimicola*. *Int. J. Agric. Biol*., 21: 1055-1062.

Bahl, S., H. Ling, N.P. Acharige, I. Santos-Barriopedro, M.K.H. Pflum and E. Seto. 2021. EGFR phosphorylates HDAC1 to regulate its expression and anti-apoptotic function. *Cell Death Dis*., 12(5): 1-17.

Bedford, M.T. and S. Richard. 2005. Arginine methylation: an emerging regulator of protein function. *Mol. Cell*., 18(3): 263-272.

Borgo, G.M., T.P. Burke, C.J. Tran, N.T. Lo, P. Engström and M.D. Welch. 2022. A patatin-like phospholipase mediates Rickettsia parkeri escape from host membranes. *Nat. Comm*., 13(1): 1-15.

Boriack-Sjodin, P.A. and K.K Swinger. 2016. Protein methyltransferases: A distinct, diverse, and dynamic family of enzymes. *Biochem*, 55(11): 1557-1569.

Buuh, Z.Y., Z. Lyu and R.E. Wang. 2017. Interrogating the roles of post-translational modifications of non-histone proteins: miniperspective. *J. Med. Chem*., 61(8): 3239-3252.

Cavallius, J., W. Zoll, K. Chakraburtty and W.C. Merrick. 1993. Characterization of yeast EF-1α: Non-conservation of post-translational modifications. *Biochem. Biophys. Acta, Protein Struct. Mol. Enzymol.*, 1163(1): 75-80.

Chen, H., Y. Xue, N. Huang, X. Yao and Z. Sun. 2006. MeMo: a web tool for prediction of protein methylation modifications. *Nucl. Acids Res*., 34(2): W249-W253.

Choudhary, C., B.T. Weinert, Y. Nishida, E. Verdin and M. Mann. 2014. The growing landscape of lysine acetylation links metabolism and cell signaling. *Nat. Rev. Mol. Cell Biol.*, 15(8): 536-550.

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

Couttas, T.A., M.J. Raftery, M.P. Padula, B.R. Herbert and M.R Wilkins. 2012. Methylation of translation-associated proteins in *Saccharomyces cerevisiae*: Identification of methylated lysines and their methyltransferases. *Proteomics.*, 12(7): 960-972.

Deng, W., Y. Wang, L. Ma, Y. Zhang, S. Ullah and Y. Xue. 2017. Computational prediction of methylation types of covalently modified lysine and arginine residues in proteins. *Brief. Bioinform.*, 18(4): 647-658.

Diallo, I., M. Seve, V. Cunin, F. Minassian, J.F. Poisson, S. Michelland and S. Bourgoin-Voillard. 2019. Current trends in protein acetylation analysis. *Expert Rev. Proteom.*, 16(2): 139-159.

Dzialo, M.C., K.J. Travaglin, S. Shen, K. Roy, G.F. Chanfreau, J.A. Loo and S.G. Clarke. 2014. Translational roles of elongation factor 2 protein lysine methylation. *J. Biol. Chem.*, 289(44): 30511-30524.

Erce, M.A., C.N. Pang, G. Hart-Smith and M.R. Wilkins. 2012. The methylproteome and the intracellular methylation network. *Proteomics*, 12(4-5): 564-586.

Falnes, P.Ø., M.E. Jakobsson, E. Davydova, A. Ho and J. Małecki. 2016. Protein lysine methylation by seven-β-strand methyltransferases. *Biochem. J.*, 473(14): 1995-2009.

Ferrari, K.J., S. Amato, R. Noberini, C. Toscani, D. Fernández-Pérez, A.D. Rossi and Pasini. 2021. Intestinal differentiation involves cleavage of histone H3 N-terminal tails by multiple proteases. *Nucl. Acids Res.*, 49(2): 791-804.

Gu, B. and W.G. Zhu. 2012. Surf the post-translational modification network of p53 regulation. *Int. J. Boil. Sci.*, 8(5): 672.

Hoffmann, A.A. and M.J. Hercus. 2000. Environmental stress as an evolutionary force. *Bioscience*, 50(3): 217-226.

Hoffmann, A.A. and P.A. Parsons. 1991. *Evolutionary genetics and environmental stress*. Oxford University Press.

Huang, R. 2019. Chemical Biology of Protein N-Terminal Methyltransferases. *Chembiochem*, 20(8): 976-984.

Huether, S.E. and K.L. McCance. 1994. Pathophysiology: The biologic basis for disease in adults and children. *Dimen. Crit. Care Nurs.*, 13(6): 315.

Hyun, K.G., Y.S. Noh and J.J. Song. 2017. *Arabidopsis* FRIGIDA stimulates EFS histone H3 Lys36 methyltransferase activity. *Plant Cell Rep.*, 36(7): 1183-1185.

Indellicato, R. and M. Trinchera. 2021. Epigenetic regulation of glycosylation in cancer and other diseases. *Int. J. Mol. Sci.*, 22(6): 2980.

Jakobsson, M.E., E. Davydova, J. Małecki, A. Moen and P.Ø. Falnes. 2015. Saccharomyces cerevisiae eukaryotic elongation factor 1A (eEF1A) is methylated at Lys-390 by a METTL21-like methyltransferase. *PLoS One*, 10(6): 0131426.

Jakobsson, M.E., J. Małecki, B.S. Nilges, A. Moen, S.A. Leidel and P.Ø. Falnes. 2017. Methylation of human eukaryotic elongation factor alpha (eEF1A) by a member of a novel protein lysine methyltransferase family modulates mRNA translation. *Nucl. Acids Res.*, 45(14): 8239-8254.

Jakobsson, M.E., J.M. Małecki, L. Halabelian, B.S. R. Nilges, R. Pinto, S. Kudithipudi, S. Munk, E. Davydova, F.R. Zuhairi, C.H. Arrowsmith and A. Jeltsch. 2018. The dual methyltransferase METTL13 targets N terminus and Lys55 of eEF1A and modulates codon-specific translation rates. *Nat. Comm.*, 9(1): 1-15.

Jenuwein, T. and C.D. Allis. 2001. Translating the histone code. *Sci.*, 293(5532): 1074-1080.

Khoury, G.A., R.C. Baliban and C.A. Floudas. 2011. Proteome-wide post-translational modification statistics: frequency analysis and curation of the swiss-prot database. *Sci. Rep.*, 1(1): 1-5.

Larsen, S.C., K.B. Sylvestersen, A. Mund, D. Lyon, M. Mullari, M.V. Madsen and M.L. Nielsen. 2016. Proteome-wide analysis of arginine monomethylation reveals widespread occurrence in human cells. *Sci. Signal.*, 9(443): 9-9.

Li, X., C. Wang, H. Jiang and C. Luo. 2019. A patent review of arginine methyltransferase inhibitors (2010–2018). *Expert. Opin. Ther. Pat.*, 29(2): 97-114.

Lipson, R.S., K.J. Webb and S.G. Clarke. 2010. Two novel methyltransferases acting upon eukaryotic elongation factor 1A in *Saccharomyces cerevisiae*. *Arch. Biochem. Biophys.*, 500(2): 137-143.

Nevo, E. 2006. " Evolution Canyon": a microcosm of life's evolution focusing on adaptation and speciation. *Isr. J. Ecol. Evol.*, 52(3-4): 501-506.

Nevo, E., Y.B. Fu, T. Pavlicek, S. Khalifa, M. Tavasi and A. Beiles. 2012. Evolution of wild cereals during 28 years of global warming in Israel. *PNAS.*, 109(9): 3412-3415.

Nishi, H., A. Shaytan and A.R. Panchenko. 2014. Physicochemical mechanisms of protein regulation by phosphorylation. *Front. Genet.*, 5: 270.

Owings, J.P., E.G. Kuiper, S.M. Prezioso, J. Meisner, J.J. Varga, N. Zelinskaya, E.B. Dammer, D.M. Duong, N.T. Seyfried, S. Albertí and G.L. Conn. 2016. Pseudomonas aeruginosa EftM is a thermoregulated methyltransferase. *J. Biol. Chem.*, 291(7): 3280-3290.

Petrossian, T.C. and S.G. Clarke. 2009. Bioinformatic identification of novel methyltransferases. *Epigenomics.*, 1(1): 163-175.

Petrossian, T.C. and S.G. Clarke. 2011. Uncovering the human methyltransferasome. *Mol. Cell. Proteom.*, 10(1): M110.000976 P1.

Qamar, S., G. Wang, S.J. Randle, F.S. Ruggeri, J.A. Varela, J.Q. Lin, E.C. Phillips, A. Miyashita., D. Williams, F. Ströhl and W. Meadows. 2018. FUS phase separation is modulated by a molecular chaperone and methylation of arginine cation-π interactions. *Cell*, 173(3): 720-734.

Reimand, J., O. Wagih and G.D. Bader. 2015. Evolutionary constraint and disease associations of post-translational modification sites in human genomes. *PLoS Gen.*, 11(1): 1004919.

Roskoski, J.R. 2012. ERK1/2 MAP kinases: structure, function, and regulation. *Pharmacol. Res.*, 66(2): 105-143.

Saleem, M., B.C. Lamb and E. Nevo. 2001. Inherited differences in crossing over and gene conversion frequencies between wild strains of *Sordaria fimicola* from "Evolution Canyon". *Genetics*, 159(4): 1573-1593.

Schaner Tooley, C.E., J.J. Petkowski, T.L. Muratore-Schroeder, J.L Balsbaugh, J. Shabanowitz, J., M. Sabat, W. Minor, D.F. Hunt and I.G. Macara. 2010. NRMT is an α-N-methyltransferase that methylates RCC1 and retinoblastoma protein. *Nature*, 466(7310): 1125-1128.

Schwartz, P.A. and B.W. Murray. 2011. Protein kinase biochemistry and drug discovery. *Bioorg. Chem.*, 39(5-6): 192-210.

Stock, A., S. Clarke, C. Clarke and J. Stock. 1987. N-terminal methylation of proteins: structure, function and specificity. *FEBS Lett.*, 220(1): 8-14.

Webb, K.J., R.S. Lipson, Q. Al-Hadi, J.P. Whitelegge and S.G. Clarke. 2010. Identification of protein N-terminal methyltransferases in yeast and humans. *Biochem.*, 49(25): 5225-5235.

Zhang, M., J.Y. Xu, H. Hu, B.C. Ye and M. Tan. 2018. Systematic proteomic analysis of protein methylation in prokaryotes and eukaryotes revealed distinct substrate specificity. *Proteomics*, 18(1): 1700300.

Zhang, X., Y. Huang and X. Shi. 2015. Emerging roles of lysine methylation on non-histone proteins. *CMLS.*, 72: 4257-4272.

(Received for publication 02 December 2023)