INTERNATIONAL JOURNAL OF AGRICULTURE & BIOLOGY
ISSN Print: 1560–8530; ISSN Online: 1814–9596
19–1195/2020/23–5–935–942
DOI: 10.17957/IJAB/15.1372
http://www.fspublishers.org

Full Length Article

Genetic and Post-Translational Modification Analysis of Translational Associated Protein RKM4 in *Sordaria fimicola*

Iqra Mobeen[1*], Rabia Arif[1], Amna Rasheed[1], Faiza Akram[1], Memuna Ghafoor Shahid[2] and Muhammad Saleem[1]
[1]*Department of Zoology, University of Education, Lower Mall Campus, Lahore, Pakisan*
[2]*Department of Botany, Government College University, Lahore Pakistan*
[*]For correspondence: iqramobeen905@gmail.com; iqra.phd.botany@pu.edu.pk
Received 24 July 2019; Accepted 12 December 2019; Published 03 March 2020

Abstract

Ribosomal-lysine N-methyltransferases are involved in important biological processes and methylate specific lysine residues of different subunits of the translational machinery. The current study investigates natural genetic variations and post-translational modifications in the *RKM4* gene in six parental strains of *Sordaria fimicola*. For this purpose, DNA was extracted and *RKM4* regions of all strains of *S. fimicola* were amplified by TD-PCR (touchdown PCR) conditions. Polymorphisms are observed on twelve different sites in the *RK-MTases* genes and six non-synonymous substitutions are found. Different bioinformatics tools were used to predict the post-translational modifications after translating the DNA sequence into the protein sequence by EMBOSS Transeq. The current study reports 28 phosphorylation sites for the S_1, S_2, N_5 strains, 26 sites for the S_3, N_6 strains and 27 sites for the N_7 strain. Seven potential lysine acetylation sites for the S_1, S_2, S_3 and five lysine acetylation sites for the N_5, N_6 and N_7 strains of *S. fimicola* are identified. Methylation at arginine and lysine residues is predicted using the PMes server. Only arginine residues exhibit methylation for *RKM4* protein methyltransferase, which includes a total of one and six arginine methylation sites for *S. fimicola* and *Saccharomyces cerevisiae,* respectively. In conclusion, the present study indicates that protein methyltransferases are meant for methylation of lysine residues and they undergo diverse types of post-translational modifications in filamentous fungi. © 2020 Friends Science Publishers

Keywords: Acetylation; Methyltransferases; Polymorphism; Phosphorylation; Methylation

Introduction

Protein methyltransferases (MTases) are the enzymes, which are meant for methylation of protein (Boriack-Sjodin and Swinger 2016). The SET domain MTases catalyze the reaction between a protein substrate and S-adenosyl-L-methionine (SAM), yielding a methylated protein and S-adenosyl- L-homocysteine (SAH). S-adenosyl-L-methionine (SAM) acts as an important cofactor for the transfer of the methyl group to biological molecules like DNA, RNA and proteins (Petrossian and Clarke 2011). SET domain MTases are a new family of methyltransferases, which specifically methylate lysine residues of a large number of different proteins (Yeates 2002). SET domain family was named after the *Drosophila* genes in which it was first discovered; *Su (var)*, Enhancer of zeste, and Trithorax (Jenuwein *et al.* 1998). Later on, it was revealed that these genes encode histone lysine methyltransferases (Rea *et al.* 2000; Nishioka *et al.* 2002). SET domain proteins have now been found in all eukaryotic organisms. The domain, which is approximately 130 amino acids long, was characterized in 1998 (Dillon *et al.* 2005).

At the molecular level, RKM4 involves in the transcription cofactor activity, histone methyltransferase activity, zinc ion, and tetra-pyrrole binding, oxidoreductase activity and acts on other nitrogenous compounds as donors. On the other hand, RKM4 also plays an important role in many biological processes, which include; Organ development, negative regulation of cellular macromolecule biosynthetic process and RNA metabolic process, negative regulation of gene expression, role in histone methylation, generation of precursor metabolites and energy, establishment of localization and peptidyl-lysine mono-methylation (Yang and Zhang 2015; Zhang *et al.* 2017).

Methylation is an extremely important post-translational modification (PTM) of proteins. Methylation has been widely studied concerning the "histone code" gene expression regulation (Black *et al.* 2012). The recent studies show that the methylation of protein also has a potential role in the non-histone proteins (Erce *et al.* 2012; Low and Wilkins 2012; Clarke 2013). Dozens of methylation sites have been explored in Ascomycetes (Plank *et al.* 2015; Yagoub *et al.* 2015) and a hundred and thousand sites in human through the methylproteome enrichment studies (Bremang *et al.* 2013; Guo *et al.* 2014).

The current study is focused on the investigation of

To cite this paper: Mobeen I, R Arif, A Rasheed, F Akram, MG Shahid, M Saleem (2020). Genetic and post-translational modification analysis of translational associated protein RKM4 in *Sordaria fimicola*. *Intl J Agric Biol* 23:935–942

genetic variations in the *RKM4* gene of different strains of *S. fimicola,* collected from "Evolution Canyon", Israel. The "Evolution Canyon", Israel has two contrasting slopes: south-facing slope (SFS) and the north-facing slope (NFS). It presents a microscale environment for the study of genetic variations in different organisms due to its diverse environmental conditions (Nevo 2012). Genetic variations are caused by recombination, spontaneous mutations, gene conversion, and environmental stress. These are key driving factors for evolution and species adaptation (Hoffmann and Hercus 2000; Saleem *et al.* 2001). The strains of *S. fimicola* from the south-facing slope (SFS) bear a high frequency of mutations and gene conversion than the strains from the north-facing slope (NFS), (Arif *et al.* 2019; Jamil *et al.* 2019). Therefore, the SFS strains undergo more genetic variations than the NFS strains of *S. fimicola* (Saleem *et al.* 2001). The fact is that SFS has xeric, harsh conditions and NFS has mild environmental conditions (Nevo 2012). So environmental stress is a major cause of genetic variations (Saleem *et al.* 2001). Genetic variations happening in the DNA ultimately pass into the proteins and then these effect post-translational modifications of proteins. The differences in the position of modified sites in the same protein among different strains of *S. fimicola* and in the *S. cerevisiae* are the reflections of genetic variations (Arif *et al.* 2017a, b). Another purpose of this study is to predict the possible post-translational modifications, 3D structures, and functions of RKM4 protein. Very little work has been done on post-translational modifications of the RKM4. To bridge this knowledge gap, we have used different bioinformatics tools to investigate post-translational modifications in this study. Although the bioinformatics tools are reliable, there is a need to study the post-translational modifications of the RKM4 protein experimentally to authenticate this study.

Materials and Methods

Sub-culturing of experimental organism

The Molecular Genetics Laboratory of Department of Botany, University of the Punjab, Lahore, provided the stock cultures of parental strains of *S. fimicola*. Originally, these strains were isolated from "Evolution Canyon", Israel by Prof. Nevo's Colleagues. The S_1, S_2, S_3 strains collected from the south-facing slope (SFS) and the N_5, N_6, N_7 strains were collected from the north-facing slope (NFS). The sub-culturing of these strains was carried out on PDA (potato dextrose agar) media under sterile conditions. The mature fungal growth was obtained after 9 days by incubating the samples at 20°C in the refrigerated incubator.

Extraction of genomic DNA

Genomic DNA from all parental strains of *S. fimicola* was extracted by the modified Spano *et al.* (1995) method (without a phenolic wash) followed by resolving DNA fragments by 1% agarose gel electrophoresis stained with 0.3μl ethidium bromide. 1 kb DNA ladder was used and the gel was photographed under UV light in the gel documentation system. Afterward, primers specific to the *RKM4* region were designed using the Primer-BLAST tool from the NCBI server (https://blast.ncbi.nlm.nih.gov/Blast.cgi) to amplify the *RKM4* gene in all strains of *S. fimicola*.

Touchdown PCR conditions for amplification of *RKM4* gene

Touchdown PCR (TD-PCR) conditions were used for the amplification of genes to study the possible genetic variations and potential sites for post-translational modifications of different strains of *S. fimicola*. 15 μL PCR reaction mixture was composed of 2 μL DNA sample, 1 μL forward primer, 1 μL reverse primer, 10 μL 2X Amp Master Mix, and 1 μL ddH2O. PCR took109 minutes and 40 cycles for the complete amplification of the gene. The time required for each step of the PCR and other conditions are given in Table 1. The amplified product was resolved at 1% agarose gel electrophoresis followed by visualization under UV light in the gel documentation system and the PCR product was sent to Macrogen Korea for sequencing. Afterward, the sequences translated into protein sequences by EMBOSS Transeq online server (https://www.ebi.ac.uk/Tools/st/emboss_transeq/Protein).

3D structures prediction and visualization

I-TASSER was used to predict the 3D structures and functions of RKM4 proteins. The confidence of each model is quantitatively measured by the C-score value that is calculated on the base of the significance of threading template alignments and the convergence parameters of the structure assembly simulations. The protein structures were visualized in PyMol molecular system. The ligand-protein interaction predictions were carried out by using BioLip, which is a ligand-protein binding database.

Tools used for prediction of post-translational modifications

Different online bioinformatics tools were used for the prediction of post-translational modifications. The PMes Server (bioinfo.ncu.edu.cn/inquiries_PMeS.aspx) was used for the prediction of methylation at lysine and arginine residues and NetPhos 3.1 Server (http://www.cbs.dtu.dk/services/NetPhos/) for the prediction of phosphorylation at threonine (T), tyrosine (Y) and serine (S) residues. The PAIL (http://bdmpail.biocuckoo.org/prediction.php) and the NetNES Servers (http://www.cbs.dtu.dk/services/NetNES/) were used for the prediction of acetylation at arginine (R) residues and nuclear export signals, respectively.

Results

Multiple sequence alignment

The extracted DNA from different strains (S_1, S_2, S_3, N_5, N_6, and N_7) of *Sordaria fimicola* were subjected to amplification of the *RKM4* gene using touchdown PCR conditions. *RKM4* regions with 900 base pairs length were amplified in all studied strains of *S. fimicola*. After sequencing, the sequences of the *RKM4* gene of different strains of *S. fimicola* were aligned with *S. cerevisiae* (reference strain) by online clustal omega alignment tool to observe genetic variations among different strains of *S. fimicola*.

We obtained 12 different polymorphic sites in the *RKM4* regions of six strains of *S. fimicola* with respect to the *S. cerevisiae*. Out of 12 polymorphic sites, six non-synonymous substitutions were observed in the *RKM4* region. Non-synonymous substitutions are those substitutions, which change the coding amino acid. At first polymorphic site in the SFS strains, T was substituted with A at the second base of a codon, resulting in the change of ATC codon into AAC, which changed the Isoleucine (I) into asparagine (N). At a second polymorphic site in the S_2, S_3 and N_5 strains, A was substituted with G, resulting in the change of codon from GAG to GGG, which changed the encoding amino acid from glutamate (E) to glycine (G). At the third site in SFS strains, AT was substituted with CG, changed the codon from GAT to GCG and changed the encoded amino acid from aspartate (D) to alanine (A). In fourth polymorphic site in NFS strains, T was replaced with A at first base of the codon, where TTT is converted into ATT and changed the amino acid from phenylalanine (F) to isoleucine (I). In SFS strains at fifth polymorphic site, G was substituted with A at third base of the codon (ATG-ATA), resulted in the change of methionine (M) into isoleucine (I). At tenth site in the S_3 strain, the substitution of T with A at second base of codon was occurred (TTT-TAT), which substituted the tyrosine (Y) with phenylalanine (F). Other polymorphic sites did not change the coding amino acids, hence known as synonymous substitutions (Fig. 1–2).

Analysis of 3D structures and ligand-protein interactions

The 3D cartoon models of RKM4 protein for *S. cerevisiae* and *S. fimicola* are shown in Fig. 3. The motifs shown in red color are α-helix, motifs in yellow color are β-sheets and motifs in green lines are expressing coils. Both 3D protein structures are different at loop regions and have a difference in coiling. The ligand-protein interaction is shown at 3D models of protein with ligand binding site residues for *S. cerevisiae* and *S. fimicola* in Fig. 4–5, respectively. *S. cerevisiae* has three ligands; SAM, Zn^{+2} and (R, R)-Butane-2, 3-diol, and each of the ligands has its binding site residues. SAM has binding site residues; E80, G81, L82, S221, R222, D239, L240, I241, N242, H243, Y287, Y300, and F302; while the Y287 provides a catalytic binding site. The binding site residues of Zinc are; C65, C68, H86, and C90. (R, R)-Butane-2, 3-diol has four binding site residues; Y41, Y54, C55 and T220 (Fig. 4a–f).

S. fimicola has two SAM and lysine ligands. SAM binding site residues include; V72, A73, G74, Y75, A222, Y223, D248, I249, L250, N251, H252, Y285, Y297, and F299; while lysine has six binding site residues; A222, S224, F225, Q226, Y285, and Y297 (Fig. 5a–d).

Table 1: Touch Down PCR conditions

Stage 1	Step	Temperature (°C)	Time
1	Denaturation	95	3 min
2	Denaturation	95	30 s
3	Annealing	Tm + 10	45 s
4	Elongation	72	60 s
Repeat steps 2-4 for 15 times			
Stage 2	Step	Temperature (°C)	Time
5	Denaturation	95	30 s
6	Annealing	Tm or (Tm – 5)	45 s
7	Elongation	72	60 s
Repeat 5-7 steps for 25 times			
Termination	Step	Temperature (°C)	Time
8	Elongation	72	5 min
9	Stop reaction	4	15 min
10	Hold	23	Until removed from machine

Prediction of post-translational modifications

Prediction of phosphorylation: For the RKM4 protein of *S. cerevisiae*, phosphorylation was predicted at 26 serine (S), 18 threonine (T) and 7 tyrosine (Y) residues at different sites in the amino acid sequence. Phosphorylation was observed at 14 serine (S), 8 threonine (T) and 6 tyrosine (Y) residues of S_1 and S_2 strains. Phosphorylation for S_3 and N_5 strains was found at 13 serine (S), 7 threonine (T) and 8 tyrosine (Y) residues. For N_6 strain, it was predicted at 13 serine (S), 6 threonine (T) and 7 tyrosine (T) residues. 13 serine (S), 6 threonine (T) and 8 tyrosine (Y) residues of N_7 strain were phosphorylated (Table 2).

Prediction of methylation: In this study, it is reported that only arginine residues of RKM4 have undergone methylation. Methylation at six arginine residues (R98, R213, R243, R388, R390, and R445) was investigated in *S. cerevisiae*. Only one arginine residue R62 was found to have the potential for methylation in all studied strains of *S. fimicola* (Table 3).

Prediction of acetylation: In *S. cerevisiae*, acetylation was investigated at 18 lysine (K) residues. In the S_1, S_2 and S_3 strains of *S. fimicola*, acetylation was found at five lysine residues and seven sites in the N_5, N_6 and N_7 strains (Table 3).

Prediction of nuclear export signals (NES): We have reported two nuclear export signals at positions 359L, 140L and three nuclear export signals at positions 60L, 276I, 279I of RKM4 in *S. cerevisiae* and S_1, S_2 and S_3 strains,

```
Ref     CTGAACGAAACCGGATCGTGGGAAGGTTTAATTATCTGTATACTTTATGAGATGGAAGTT 60
N5      CTGAACGAAACCGGATCGTGGGAAGGTTTAATTATCTGTATACTTTATGGGATGGAAGTT 60
N6      CTGAACGAAACCGGATCGTGGGAAGGTTTAATTATCTGTATACTTTATGAGATGGAAGTT 60
N7      CTGAACGAAACCGGATCGTGGGAAGGTTTAATTATCTGTATACTTTATGAGATGGAAGTT 60
S3      CTGAACGAAACCGGATCGTGGGAAGGTTTAATTAACTGTATACTTTATGGGATGGAAGTT 60
S1      CTGAACGAAACCGGATCGTGGGAAGGTTTAATTAACTGTATACTTTATGAGATGGAAGTT 60
S2      CTGAACGAAACCGGATCGTGGGAAGGTTTAATTAACTGTATACTTTATGGGATGGAAGTT 60
        ********************************** ************** **********
Ref     AACGCGTTAATTTTTTGGGATGATAATGAACTGCAACTTTTAAAACCATCACTTGTCCTT 180
N5      AACGCGTTAATTTTTTGGGATGATAATGAACTGCAACTTTTAAAACCATCACTTGTCCTT 180
N6      AACGCGTTAATTTTTTGGGATGATAATGAACTGCAACTTTTAAAACCATCACTTGTCCTT 180
N7      AACGCGTTAATTTTTTGGGATGATAATGAACTGCAACTTTTAAAACCATCACTTGTCCTT 180
S3      AACGCGTTAATTTTTTGGGATGCGAATGAACTGCAACTTTTAAAACCATCACTTGTCCTT 180
S1      AACGCGTTAATTTTTTGGGATGCGAATGAACTGCAACTTTTAAAACCATCACTTGTCCTT180
S2      AACGCGTTAATTTTTTGGGATGCGAATGAACTGCAACTTTTAAAACCATCACTTGTCCTT 180
        **********************  ************************************
Ref     CAAATTGGTGGAGAATTTTCACGTGTAGCGACGTCCTTCGAGTTTGATAATTTTGCCTAT 300
N5      CAAATTGGTGGAGAATTTTCACGTGTAGCGACGTCCTTCGAGATTGATAATTTTGCCTAT 300
N6      CAAATTGGTGGAGAATTTTCACGTGTAGCGACGTCCTTCGAGATTGATAATTTTGCCTAT 300
N7      CAAATTGGTGGAGAATTTTCACGTGTAGCGACGTCCTTCGAGATTGATAATTTTGCCTAT 300
S3      CAAATTGGTGGAGAATTTTCACGTGTAGCGACGTCCTTCGAGTTTGATAATTTTGCCTAT 300
S1      CAAATTGGTGGAGAATTTTCACGTGTAGCGACGTCCTTCGAGTTTGATAATTTTGCCTAT 300
S2      CAAATTGGTGGAGAATTTTCACGTGTAGCGACGTCCTTCGAGTTTGATAATTTTGCCTAT 300
        ****************************************** *****************
Ref     ATGATTCCACTTGCTGATATGTTGAATGCAGATACCAGTAAATGCAACGCTAATTTAACT 480
N5      ATGATTCCACTTGCTGATATGTTGAATGCAGATACCAGTAAATGCAACGCTAATTTAACT 480
N6      ATGATTCCACTTGCTGATATGTTGAATGCAGATACCAGTAAATGCAACGCTAATTTAACT 480
N7      ATGATTCCACTTGCTGATATGTTGAATGCAGATACCAGTAAATGCAACGCTAATTTAACT 480
S3      ATGATTCCACTTGCTGATATATTGAATGCAGATACCAGTAAATGCAACGCTAATTTAACT 480
S1      ATGATTCCACTTGCTGATATATTGAATGCAGATACCAGTAAATGCAACGCTAATTTAACT 480
S2      ATGATTCCACTTGCTGATATATTGAATGCAGATACCAGTAAATGCAACGCTAATTTAACT 480
        ******************** ***************************************
Ref     TACGACTCTAACTGTTTAAAGATGGTTGCTTTGAGGGATATTGAAAAGAATGAACAAGTG 540
N5      TACGACTCTAACTGTTTAAAGATGGTTGCTTTGAGGGATATTGAAAAGAATGAACAAGTG 540
N6      TACGACTCTAACTGTTTAAAGATGGTTGCTTTGAGGGATATTGAAAAGAATGAACAAGTG 540
N7      TACGACTCTAACTGTTTAAAGATGGTTGCTTTGAGGGATATTGAAAAGAATGAACAAGTG 540
S3      TACGACTCTAACTGTTTAAAGATGGTTGCTTTGAGGGACATTGAAAAGAATGAACAAGTG 540
S1      TACGACTCTAACTGTTTAAAGATGGTTGCTTTGAGGGACATTGAAAAGAATGAACAAGTG 540
S2      TACGACTCTAACTGTTTAAAGATGGTTGCTTTGAGGGACATTGAAAAGAATGAACAAGTG 540
        ************************************** *********************
Ref     TGGGACGGTTCGAAGTATGATTTTGGAGAAGTGTTACTTGAAAATATTGTCGAGGCGTTA 660
N5      TGGGACGGTTCGAAGTATGATTTTGGAGAAGTGTTACTTGAAAATATTGTCGAGGCGTGA 660
N6      TGGGACGGTTCGAAGTATGATTTTGGAGAAGTGTTACTTGAAAATATTGTCGAGGCGTGA 660
N7      TGGGACGGTTCGAAGTATGATTTTGGAGAAGTGTTACTTGAAAATATTGTCGAGGCGTGA 660
S3      TGGGACGGTTCGAAGTATGATTTTGGAGAAGTGTTACTTGAAAATATTGTCGAGGCGTTA 660
S1      TGGGACGGTTCGAAGTATGATTTTGGAGAAGTGTTACTTGAAAATATTGTCGAGGCGTTA 660
S2      TGGGACGGTTCGAAGTATGATTTTGGAGAAGTGTTACTTGAAAATATTGTCGAGGCGTTA 660
        ********************************************************** *
Ref     AAAGAGACTTTTGAGACGAATACTGAATTTTTGGACAGGTGTATTGATATCTTACGCAAT 720
N5      AAAGAGACTTTTGAGACGAATACTGAATTTTTGGACAGGTGTATTGATATCTTACGCAAT 720
N6      AAAGAGACTTTTGAGACGAATACTGAATTTTTGGACAGGTGTATTGATATCTTACGCAAT 720
N7      AAAGAGACTTTTGAGACGAATACTGAATTTTTGGACAGGTGTATTGATATCTTACGCAAT 720
S3      AAAGAGACTTTTGAGACGAATACTGAATTTTTGGACAGGTGTATTGATATCTTACGCAAT 720
S1      AAAGAGACTTTTGAGACGAATACTGAGTTTTTGGACAGGTGTATTGATATCTTACGCAAT 720
S2      AAAGAGACTTTTGAGACGAATACTGAGTTTTTGGACAGGTGTATTGATATCTTACGCAAT 720
        ************************** *********************************
Ref     TATAATAATGGTGAATTGTTGCCTCAACTAATACTTTTGGTCCAAATCTTGACAATTCTT 840
N5      TATAATAATGGTGAATTGTTGCCTCAACTAATACTTTTGGTCCAAATCTTGACAATTCTT 840
N6      TATAATAATGGTGAATTGTTGCCTCAACTAATACTTTTGGTCCAAATCTTGACAATTCTT 840
N7      TATAATAATGGTGAATTGTTGCCTCAACTAATACTTTTGGTCCAAATCTTGACAATTCTT 840
S3      TTTAATAATGGTGAATTGTTGCCTCAACTAATACTTTTGGTCCAAATCTTGACGATTCTT 840
S1      TATAATAATGGTGAATTGTTGCCTCAACTAATACTTTTGGTCCAAATCTTGACAATTCTT 840
S2      TATAATAATGGTGAATTGTTGCCTCAACTAATACTTTTGGTCCAAATCTTGACAATTCTT 840
        * *************************************************** ******
Ref     TGCCAAATTCCAGGTTTATGCAAACTGGACATAAAAGCAATGGAAAGGCAAGTGGAGAGA 900
N5      TGCCAAATTCCAGGTTTATGCAAACTGGACATAAAGGCAATGGAAAGGCAAGTGGAGAGA 900
N6      TGCCAAATTCCAGGTTTATGCAAACTGGACATAAAGGCAATGGAAAGGCAAGTGGAGAGA 900
N7      TGCCAAATTCCAGGTTTATGCAAACTGGACATAAAGGCAATGGAAAGGCAAGTGGAGAGA 900
S3      TGCCAAATTCCAGGTTTATGCAAACTGGACATAAAAGCAATGGAAAGGCAAGTGGAGAGA 900
S1      TGCCAAATTCCAGGTTTATGCAAACTGGACATAAAAGCAATGGAAAGGCAAGTGGAGAGA 900
S2      TGCCAAATTCCAGGTTTATGCAAACTGGACATAAAAGCAATGGAAAGGCAAGTGGAGAGA 900
```

Fig. 1: Multiple sequence alignment of different strains of *S. fimicola* with respect to the *S. cerevisiae* to observe genetic diversity among different strains of *S. fimicola* for *RKM4* gene
Keywords: Symbol (*) showing fully conserved sites, space and highlighted regions showing polymorphic sites

respectively. In N_5, N_6 and N_7 strains, four NES sites (53L, 60L, 275I, and 278I) were observed. 60 L residue has been found common for all strains of *S. fimicola* and other sites are present in close proximity with respect to one another. This shows that the 60 L site is conserved in all strains of *S. fimicola* (Table 3).

Discussion

To the best of our knowledge, the *RKM4* gene is first time reported in *S. fimicola*. In the current study, genetic variations investigated in the *RKM4* gene of *S. fimicola*. The SFS strains have nine polymorphic sites and the NFS strains have three polymorphic sites (Fig. 1). Non-synonymous substitutions were observed at six sites in the *RKM4* region, resulted in the change of coding amino acid (Fig. 2). As SFS have xeric and more stressful conditions, due to this SFS strains have more genetic variations than NFS strains. This reveals that environmental stress has a role in the creation of genetic variations as reported by other geneticists. Arif *et al.* (2017a) investigated polymorphism in the *S. fimicola* with the help of SSR marker and identified that SFS strains have more variations than NFS strains. Hosid *et al.* (2008) reported high levels of polymorphism with the help of SSR marker in the soil fungus *Emericella nidulans* from a stressful environment and low levels of polymorphism in the fungus collected from an arid environment. Moreover, genetic variations help the species to better survive in the fluctuating and stressful environment and are major causes of evolution (Hoffmann and Hercus 2000; Saleem *et al.* 2001). Environmental conditions encounter organisms with natural selection by manipulating parental and genetic variants and thus genetic variations become a requirement for evolution as they determine the evolutionary potential of a population (Arber 2000).

```
N6      --------------------LNETGSWEGLIICILYEMEVLQERSRWAPYFKVWNKPSDM      40
N7      --------------------LNETGSWEGLIICILYEMEVLQERSRWAPYFKVWNKPSDM      40
N5      --------------------LNETGSWEGLIICILYGMEVLQERSRWAPYFKVWNKPSDM      40
S2      --------------------LNETGSWEGLINCILYGMEVLQERSRWAPYFKVWNKPSDM      40
S3      --------------------LNETGSWEGLINCILYGMEVLQERSRWAPYFKVWNKPSDM      40
Ref     --------------------LNETGSWEGLIICILYEMEVLQERSRWAPYFKVWNKPSDM      40
S1      --------------------LNETGSWEGLINCILYEMEVLQERSRWAPYFKVWNKPSDM      40
                            *********** **** ***********************

N6      NALIFWDDNELQLLKPSLVLERIGKKEAKEMHERIIKSIKQIGGEFSRVATSFEIDNFAY      100
N7      NALIFWDDNELQLLKPSLVLERIGKKEAKEMHERIIKSIKQIGGEFSRVATSFEIDNFAY      100
N5      NALIFWDDNELQLLKPSLVLERIGKKEAKEMHERIIKSIKQIGGEFSRVATSFEIDNFAY      100
S2      NALIFWDANELQLLKPSLVLERIGKKEAKEMHERIIKSIKQIGGEFSRVATSFEFDNFAY      100
S3      NALIFWDANELQLLKPSLVLERIGKKEAKEMHERIIKSIKQIGGEFSRVATSFEFDNFAY      100
Ref     NALIFWDDNELQLLKPSLVLERIGKKEAKEMHERIIKSIKQIGGEFSRVATSFEFDNFAY      100
S1      NALIFWDANELQLLKPSLVLERIGKKEAKEMHERIIKSIKQIGGEFSRVATSFEFDNFAY      100
        ******* ********************************************** *****

N6      IASIILSYSFDLEMQDSSVNENEEEETSEEELENERYLKSMIPLADMLNADTSKCNANLT      160
N7      IASIILSYSFDLEMQDSSVNENEEEETSEEELENERYLKSMIPLADMLNADTSKCNANLT      160
N5      IASIILSYSFDLEMQDSSVNENEEEETSEEELENERYLKSMIPLADMLNADTSKCNANLT      160
S2      IASIILSYSFDLEMQDSSVNENEEEETSEEELENERYLKSMIPLADILNADTSKCNANLT      160
S3      IASIILSYSFDLEMQDSSVNENEEEETSEEELENERYLKSMIPLADILNADTSKCNANLT      160
Ref     IASIILSYSFDLEMQDSSVNENEEEETSEEELENERYLKSMIPLADMLNADTSKCNANLT      160
S1      IASIILSYSFDLEMQDSSVNENEEEETSEEELENERYLKSMIPLADILNADTSKCNANLT      160
        ********************************************** *************

N6      KETFETNTEFLDRCIDILRNNANIQEFLEGEEIVLDSYDCYNNGELLPQLILLVQILTIL      279
N7      KETFETNTEFLDRCIDILRNNANIQEFLEGEEIVLDSYDCYNNGELLPQLILLVQILTIL      279
N5      KETFETNTEFLDRCIDILRNNANIQEFLEGEEIVLDSYDCYNNGELLPQLILLVQILTIL      279
S2      KETFETNTEFLDRCIDILRNNANIQEFLEGEEIVLDSYDCYNNGELLPQLILLVQILTIL      280
S3      KETFETNTEFLDRCIDILRNNANIQEFLEGEEIVLDSYDCFNNGELLPQLILLVQILTIL      280
Ref     KETFETNTEFLDRCIDILRNNANIQEFLEGEEIVLDSYDCYNNGELLPQLILLVQILTIL      280
S1      KETFETNTEFLDRCIDILRNNANIQEFLEGEEIVLDSYDCYNNGELLPQLILLVQILTIL      280
        **************************************** *******************
```

Fig. 2: Multiple sequence alignment of amino acid sequence of RKM4 protein of different strains of *S. fimicola* with respect to the reference strain *S. cerevisiae* amino acid sequence to spot genetic diversity
Symbol (*) showing conserved sites, space and highlighted regions showing polymorphic sites

On the other hand, genetic variation also has its reflection on post-translational modifications of proteins as described by Arif *et al.* (2017b) in Frequency Clock and Mating Type a-1 proteins of parental strains of *S. fimicola* from EC Israel. Therefore, the difference in the positions of modified sites for the same protein in the different strains of *S. fimicola* is the reflection of genetic variations (Table 2 and 3). Genetic variations of nucleotide sequence ultimately translated into the protein and this produces the proteins with unique PTM sites, which lead towards the functional diversity of proteins. Thus, post-translational modifications

(PTMs) are very important because they change the configuration of proteins and this affects their catalytic functions. Therefore, it is necessary to study them to see how the PTMs play their role in maintaining the biological functions of proteins (Lothrop *et al.* 2013). There are many types of PTMs of proteins, while the present study mainly focused upon phosphorylation, methylation, and acetylation of RKM4 methyltransferase of *S. cerevisiae* and *S. fimicola.*

Methyltransferases are involved in important biological processes and methylate specific lysine (Lys) and N-terminal residues of different subunits of the translational machinery (Porras-Yakushi *et al.* 2007; Lipson *et al.* 2010; Hamey *et al.* 2016). SAM-dependent MTase, RKM4 mono-methylates 60S ribosomal protein L42 (RPL42A and RPL42B) at 'Lys-55' (Webb *et al.* 2008; Lipson *et al.* 2010). A second SET domain methyltransferase Rkm2 is also identified, which is responsible for tri-methylating the ribosomal protein L12ab (Rpl12ab) at lysine 10. The second site of lysine methylation for Rpl12ab at position 3 by RKM2 is identified (Porras-Yakushi *et al.* 2006; Webb *et al.* 2008; Gardner *et al.* 2011).

The RKM4 has many important molecular and biological functions, which were predicted by I-TASSER in this study (Yang and Zhang 2015; Zhang *et al.* 2017). Uslupehlivan *et al.* (2018) predicted the 3D structure and functions of the prion protein of sheep (*Ovis aries*) by I-TASSER. Likewise, Rong *et al.* (2019) used I-TASSER to predict the membrane-spanning helices and topology models for Δ17 fatty acid desaturases from *Rhizophagus irregularis* and *Octopus bimaculoides*.

Methylation predominantly found on lysine (Lys) and arginine (Arg) residues in eukaryotes (Clarke, 2013). In the present study, methylation was predicted at arginine residues of RKM4 MTase of *S. cerevisiae* and *S. fimicola*. The six-arginine residues of RKM4 MTase of *S. cerevisiae* have the potential for methylation. Only one site (R62) was predicted for RKM4 protein, which is common for the SFS and the NFS strains of *S. fimicola* (Table 3). Winter *et al.* (2017) observed methylation at 528th lysine residue of the RKM4 protein of *S. cerevisiae* using mass spectrometry.

Protein phosphorylation plays an important role in the regulatory and signaling processes of the cell. This affects up to 30% of the proteome and essential in the regulation of cellular functions, protein degradation, and stabilization. In addition, phosphorylation networks are also essential backbones of the communication system within cells (Manning *et al.* 2002; Ficarro *et al.* 2002). In the present study, a total 51 phosphorylated sites were investigated for the RKM4 protein of *S. cerevisiae* (Table 2). Winter *et al.* (2017) have identified 3 serine phosphorylation sites at S24, S420, and S446 positions and one threonine phosphorylation site at 480 amino acid position of RKM4 MTase of *S. cerevisiae*. Serine modification at S420 and threonine modification at T480 was also predicted in the present study beside other sites in *S. cerevisiae* (Table 2).

In S_1, S_2, and S_3 strains of *S. fimicola*, phosphorylation

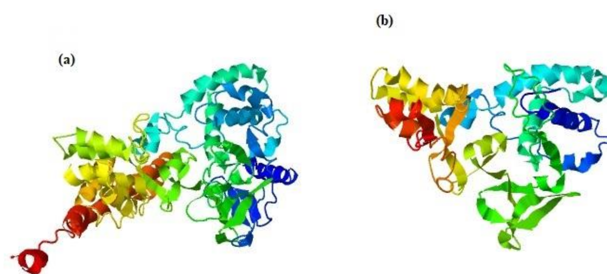

Fig. 3: 3D structure models of RKM4 protein of (**a**) *S. cerevisiae* (**b**) *S. fimicola* generated by I-TASSER and visualized by PyMol. Motifs shown in red color indicate α-helix, yellow indicate β-sheet and motifs shown in green represent coil structure

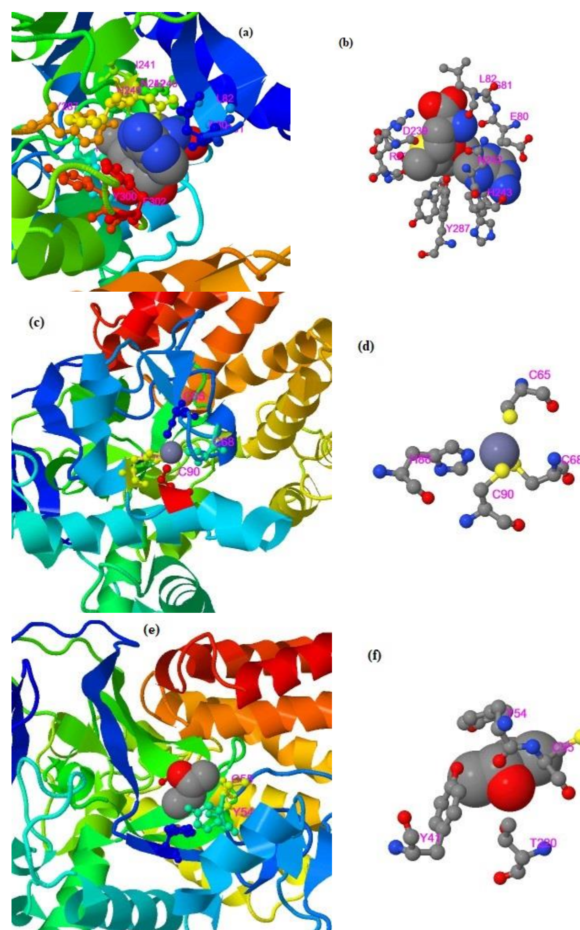

Fig. 4: Ligand-protein interaction of RKM4 protein of *S. cerevisiae.* RKM4 protein has three ligands; SAM and lysine generated from BioLip (ligand-protein binding database). (**a**). SAM ligand attachment with protein. (**b**). SAM ligand with its binding site residue, while Y287 is catalytic site residue (**c**). Zinc ligand attachment with protein. (**d**). Lysine ligand with its binding site residues. (**e**). (R-R)-Butan-2,3-diol attachment with protein. (**f**). (R-R)-Butan-2,3-diol with its binding site residues shown in violet color

reported at 28 sites. For N_6 and N_7 strains of *S. fimicola*, 26 and 27 sites were predicted, respectively (Table 2). Zhu *et*

Table 2: Phosphorylation predicted sites with their protein kinases for RKM4 protein of *S. cerevisiae* and different strains of *S. fimicola.* Numbers in third column are showing the phosphorylation positions on serine, threonine and tyrosine residues of RKM4. The numbers in the others columns (last four) are showing the positions, where the specific protein kinase involved in the phosphorylation of its respective residue i.e. serine, threonine, and tyrosine

Organism	Residues	Phosphorylation sites	Protein kinases			
			CKII	Unsp	PKC	PKA
S. cerevisiae	Serine	5, 24, 59, 60, 63, 67, 75, 86, 118, 137, 158, 172, 187, 189, 197, 198, 208, 270, 337, 420, 425, 429, 446, 450, 485, 486 Total = 26 sites	5, 189, 197, 198, 207, 208, 337, 450	24, 63, 75, 86, 158, 172, 198, 207, 208, 420, 425, 429, 446, 485, 486	63, 75, 158, 187, 446, 486	59, 60, 137, 429
	Threonine	17, 44, 52, 65, 84, 171, 207, 232, 240, 303, 306, 308, 396, 401, 431, 457, 472, 480 Total = 18 sites	17, 84, 171, 308, 431	171, 232, 396, 457, 480	44, 52, 240, 401, 472	-
	Tyrosine	217, 241, 261, 278, 286, 338, 341 Total = 7 sites	-	217, 261, 286, 341	241, 278	-
S1, S2	Serine	6, 38, 57, 78, 87, 92, 107, 109, 117, 118, 128, 190, 204, 257 Total = 14 sites	109, 117, 118, 128, 190, 257	6, 92, 118, 128, 204	78, 107	38, 57
	Threonine	4, 91, 127, 152, 160, 223, 226, 228 Total = 8 sites	4, 127, 226, 228	91, 152	160	-
	Tyrosine	137, 161, 181, 199, 206, 261 Total = 6 sites	-	137, 161, 181, 198, 206, 261	-	-
S3, N5	Serine	6, 38, 57, 78, 92, 107, 109, 117, 118, 128, 190, 204, 256 Total = 13 sites	109, 117, 256	6, 92, 118, 128, 204	78, 107	38, 57
	Threonine	4, 91, 127, 152, 160, 225, 227 Total = 7 sites	4, 227	91, 127, 152	160	-
	Tyrosine	100, 137, 161, 181, 196, 206, 257, 260 Total = 8 sites	225	137, 161, 181, 198, 206, 260	-	-
N6	Serine	6, 38, 57, 78, 92, 107, 109, 117, 118, 128, 190, 204, 256 Total = 13 sites	109, 117, 118, 128 190, 256	6, 92, 118, 128, 204, 78	78, 107	38, 57
	Threonine	4, 91, 127, 152, 160, 225 Total= 6 sites	4, 91, 127, 225	91, 127, 152	160	
	Tyrosine	100, 137, 181, 196, 206, 257, 260 Total = 7 sites	-	137, 161, 181, 198, 206, 260	-	-
N7	Serine	6, 38, 57, 78, 92, 107, 109, 117, 118, 128, 190, 204, 256 Total = 13 sites	109, 117, 118, 128, 190, 256	6, 78, 92, 118, 128, 204	78, 107	38, 57
	Threonine	4, 91, 127, 152, 160, 225 Total = 6 sites	4, 91, 127, 225	91, 127, 152	160	-
	Tyrosine	100,137, 161, 181, 198, 206, 257, 260 Total = 8 sites	-	137, 161, 181, 198, 206, 260	-	-

Table 3: Table showing predicted methylation, acetylation and nuclear export signals (NES) sites for RKM4 protein of *S. cerevisiae* and all studied strains of *S. fimicola.* Numbers are showing methylation positions on arginine (R), acetylation positions on lysine (K) and NES positions on leucine (L) and isoleucine (I)

Organism	Residue	Methylation sites	Residue	Acetylation sites	NES sites	NES potential
S. cerevisiae	Arginine (R)	98, 213, 243, 388, 390, 445 Total = 6 sites	Lysine (K)	30, 46, 49, 77, 146, 149, 219, 234, 247, 368, 403, 423, 426, 445, 484,488, 493, 494 Total = 18 sites	359-L 140-L	0.575 0.529
S1, S2, S3	Arginine (R)	62 Total = 1 site	Lysine (K)	66, 69, 139, 154, 167 Total = 5 sites	60-L 276-I 279-I	0.624 0.507 0.562
N5, N6, N7	Arginine (R)	62 Total = 1 site	Lysine (K)	66, 69, 139, 154, 167, 220, 287 Total = 7 sites	53-L 60-L 275-I 278-I	0.514 0.646 0.505 0.561

Symbols L=Leucine and I=Isoleucine

al. (2001) identified 4,200 phosphorylation events affecting 1,325 proteins from the 87 yeast protein kinase assays by the use of proteome chip technology. Bodenmiller *et al.* (2010) studied two serine phosphorylation sites at 67 and 69 positions. Winter *et al.* (2017) studied phosphorylation at two other serine residues at 129 and 573 positions. Albuquerque *et al.* (2008) also reported phosphorylation at one serine residue at 573 amino acid position experimentally for RKM1 of *S. cerevisiae*.

Protein kinases are the most crucial enzymes, involve in protein phosphorylation. They transfer a phosphate group from ATP to the protein substrate and phosphorylate the protein. They constitute about 2% of eukaryotes genome and phosphorylate about 30% cellular proteins (Ubersax and Ferrell 2007). CKII, Unsp, PKC, and PKA are some important protein kinases, which are potentially involved in phosphorylation at different residues of RKM4 protein of *S. fimicola* and *S. cerevisiae* (Table 2). Arif *et al.* (2019)

reported four protein kinases (PKC, Unsp, PKA, cdc2), which are involved in the phosphorylation of Cytochrome c oxidase (COX1).

Lysine acetylation is a reversible post-translational modification of proteins and plays a key role in regulating gene expression (Choudhary *et al.* 2009). Protein lysine acetylation has emerged as a key post-translational modification in cellular regulation, particularly through the modification of histones and nuclear transcription regulators (Zhao *et al.* 2010). For the RKM4 protein of *S. cerevisiae*, acetylation predicted at 18 lysine (K) residues. In the S_1, S_2 and S_3 strains of *S. fimicola*, acetylation observed at five lysine residues and in the N_5, N_6, and N_7 strains, seven acetylation sites were reported (Table 3). The difference in the acetylation sites for RKM4 protein lies between the strains of two contrasting slopes of EC. There was no difference in between the strains of each slope. Winter *et al.* (2017) studied the two lysine sites at 403 and 488 positions in the RKM4 protein of *S. cerevisiae* and these sites are also identified in all strains of *S. fimicola*. This shows the conservation of both sites and these sites might be involved in the regulation of RKM4 protein.

Nuclear export signals (NESs) play an extremely important role in the regulation of the subcellular location of proteins. Other nuclear processes and transcription are affected by this regulation. These processes are very important in maintaining the viability of the cell. The most important properties of NESs are accessibility and flexibility allowing relevant proteins to interact with the signal (Cour *et al.* 2004). For the RKM4 protein of *S. cerevisiae,* two positions; 140 L and 359 L have been predicted. Three positions (60-L, 276-I, 79-I) of nuclear export signals for S_1, S_2, and S_3, four positions (53-L, 60-L, 275-I, 278-I) for N_5, N_6 and N_7 have been predicted. One position 60-L is common for all studied strains of *S. fimicola* (Table 3). The presence of these nuclear export signals in the RKM4 protein of *S. cerevisiae* as well as in different strains of *S. fimicola* indicates that this protein is primarily regulated through these nuclear export signals. Arif *et al.* (2017a) predicted nuclear export signals in the frequency clock protein of *S. fimicola* at 328th amino acid residue and in *Neurospora crassa* at 323rd amino acid residue. Furthermore, some recent studies have been carried out by Jamil *et al.* (2018) and Arif *et al.* (2019) on the post-translational modifications of H3/H4 histones and cytochrome c oxidase (COX1) of *S. fimicola*, respectively by using different bioinformatics tools, also used in this study. These studies and current study will help to bridge the knowledge gap related to the post-translational modifications of proteins in *S. fimicola*.

Conclusion

It is concluded that SFS strains have more genetic variations than NFS strains because SFS strains have stressful environmental conditions. These genetic variations also have their reflections upon post-translational modifications of proteins. Therefore, in this study different post-translational modified sites are reported for RKM4 methyltransferase of *S. fimicola*. RKM4 has an important role in molecular and biological processes, which are predicted by I-TASSER. However, experimental studies need to authenticate these functions and to clarify the unknown functions of phosphorylation, methylation, and acetylation.

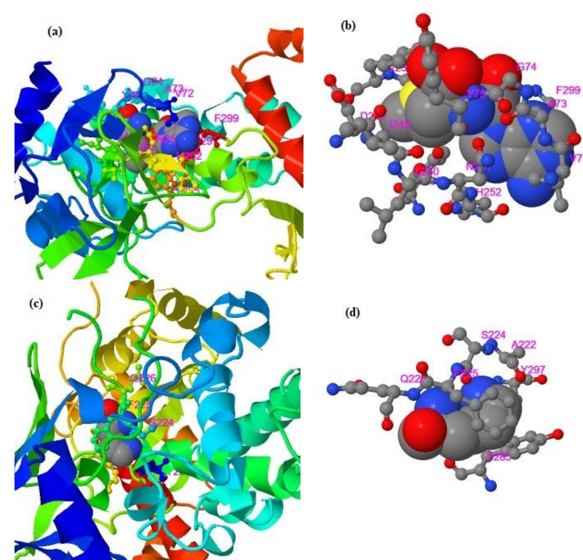

Fig. 5: Ligand-protein interaction of RKM4 protein of *S. fimicola.* RKM4 protein has two ligands; SAM and lysine generated from BioLip (ligand-protein binding database). (a). SAM ligand attachment with protein. (b). SAM ligand with its binding site residues, while H252 is catalytic site residue (c). Lysine ligand attachment with protein. (d). Lysine ligand with its binding site residues

References

Albuquerque CP, MB Smolka, SH Payne, V Bafna, J Eng, H Zhou (2008). A multidimensional chromatography technology for in-depth phosphoproteome analysis. *Mol Cell Proteomics* 7:1389–1396

Arber W (2000). Genetic variation: molecular mechanisms and impact on microbial evolution. *FEMS Microbiol Rev* 24:1–7

Arif R, SH Bukhari, M Ishfaq, MG Shahid, SF Lee, M Saleem (2019). Genetic variation and post-translational modifications of cytochrome c oxidase-1 (COX1) in different strains of *Sordaria fimicola. Intl J Agric Biol* 21:1055–1062

Arif R, F Akram, T Jamil, H Mukhtar, SF Lee, M Saleem (2017a). Genetic variation and its reflection on posttranslational modifications in frequency clock and mating type a-1 proteins in *Sordaria fimicola*. *Biol Med Res Intl* 2017; Article 1268623

Arif R, SF Lee, M Saleem (2017b). Evaluating short sequence markers in *S. fimicola* by high resolution melt analysis. *Intl J Agric Biol* 19:248–254

Black JC, CV Rechem, JR Whetstine (2012). Histone lysine methylation dynamics: establishment, regulation, and biological impact. *J Mol Cell* 48:491–507

Bodenmiller B, S Wanka, C Kraft, J Urban, D Campbell, PG Pedrioli, B Gerrits, P Picotti, H Lam, O Vitek, MY Brusniak (2010). Phosphoproteomic analysis reveals interconnected system-wide responses to perturbations of kinases and phosphatases in yeast. *Sci Signal* 3; Article 153

Boriack-Sjodin PA, KK Swinger (2016). Protein methyltransferases: A distinct, diverse, and dynamic family of enzymes. *Biochemistry* 55:1557–1569

Bremang M, A Cuomo, AM Agresta, M Stugiewicz, V Spadotto, T Bonaldi (2013). Mass spectrometry-based identification and characterization of lysine and arginine methylation in the human proteome. *Mol BioSyst* 9:2231–2247

Choudhary C, C Kumar, F Gnad, ML Nielsen, M Rehman, TC Walther, JV Olsen, M Mann (2009). Lysine acetylation targets protein complexes and co-regulates major cellular functions. *Science* 325:834–840

Clarke SG (2013). Protein methylation at the surface and buried deep: Thinking outside the histone box. *Trends Biochem Sci* 38:243–252

Cour TL, L Kiemer, A Molgaard, R Gupta, K Skriver, S Brunak (2004). Analysis and prediction of leucine-rich nuclear export signals. *Protein Eng Desig Select* 17:527–536

Dillon SC, X Zhang, RC Trievel, X Cheng (2005). The SET-domain protein superfamily: Protein lysine methyltransferases. *Genome Biol* 6; Article 227

Erce MA, CN Pang, G Hart-Smith, MR Wilkins (2012). The methylproteome and the intracellular methylation network. *Proteomics* 12:564–586

Ficarro SB, ML McCleland, PT Stukenberg, DJ Burke, MM Ross, J Shabanowitz, DF Hunt, FM White (2002). Phosphoproteome analysis by mass spectrometry and its application to *Saccharomyces cerevisiae*. *Nat Biotechnol* 20:1-6

Gardner KE, L Zhou, MA Parra, X Chen, BD Strahl (2011). Identification of lysine 37 of histone H2B as a novel site of methylation. *PLoS One* 6; Article e0016244

Guo A, H Gu, J Zhou, D Mulhern, Y Wang, KA Lee, V Yang, M Aguiar, J Kornhauser, X Jia, J Ren (2014). Immunoaffinity enrichment and mass spectrometry analysis of protein methylation. *Mol Cell Proteomics* 13:372–387

Hamey JJ, DL Winter, D Yagoub, CM Overall, G Hart-Smith, MR Wilkins (2016). Novel N-terminal and lysine methyltransferases that target translation elongation factor 1A in Yeast and human. *Mol Cell Proteomics* 15:164–176

Hoffmann AA, MJ Hercus (2000). Environmental stress as an evolutionary force. *Biol Sci* 50:217–226

Hosid E, I Grishkan, Z Frenkel, SP Wasser, E Nevo, 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

Jamil T, S Ijaz, R Arif, F Akram, M Saleem (2019). Diversity of Retrotransposons (Ty3-gypsy, LINEs and Ty1-copia) in *Sordaria fimicola. Pak J Zool* 51:1200–1202

Jamil T, N Sami, R Arif, Q Rashid, M Saleem (2018). H3/H4 Histone genes variations and its effect on post translational modifications in various strains of *Sordaria fimicola. Intl J Agric Biol* 20:1021–1026

Jenuwein T, G Laible, R Dorn, G Reuter (1998). SET domains modulate chromatin domains in eu and heterochromatin. *Cell Mol Life Sci* 54:80–93

Lipson RS, KJ Webb, SG Clarke (2010). Two novel methyltransferases acting upon eukaryotic elongation factor 1A in *Saccharomyces cerevisiae*. *Arch Biochem Biophys* 500:137–143

Lothrop AP, MP Torres, SM Fuchs (2013). Deciphering posttranslational modification codes. *FEBS Lett* 587:1247–1257

Low JK, MR Wilkins (2012). Protein arginine methylation in *Saccharomyces cerevisiae*. *FEBS J* 279:4423–4443

Manning G, GD Plowman, T Hunter, S Sudarsanam (2002). Evolution of protein kinase signaling from yeast to man. *Trends Biochem Sci* 27:514–520

Nevo E (2012). "Evolution Canyon," a potential microscale monitor of global warming across life. *Proc Natl Acad Sci* 109:2960–2965

Nishioka K, S Chuikov, K Sarma, H Erdjument-Bromage, CD Allis, P Tempst, D Reinberg (2002). Set9, a novel histone H3 methyltransferase that facilitates transcription by precluding histone tail modifications required for heterochromatin formation. *Genes Dev* 16:479–489

Petrossian TC, SG Clarke (2011). Uncovering the human methyltransferasome. *Mol Cell Proteomics* 10; Article M110.000976

Plank M, R Fischer, V Geoghegan, PD Charles, R Konietzny, O Acuto, C Pears, CJ Schofield, BM Kessler (2015). Expanding the Yeast protein arginine methylome. *Proteomics* 15:3232–3243

Porras-Yakushi TR, JP Whitelegge, S Clarke (2007). Yeast ribosomal/cytochrome c SET domain methyltransferase subfamily identification of Rpl23ab methylation sites and recognition motifs. *J Biol Chem* 282:12368–12376

Porras-Yakushi TR, JP Whitelegge, S Clarke (2006). A novel SET domain methyltransferase in Yeast Rkm2-dependent trimethylation of ribosomal protein L12ab at lysine 10. *J Biol Chem* 281:35835–35845

Rea S, F Eisenhaber, D O'carroll, BD Strahl, ZW Sun, M Schmid, S Opravil, K Mechtler, CP Ponting, CD Allis, T Jenuwein (2000). Regulation of chromatin structure by site-specific histone H3 methyltransferases. *Nature* 406:1–7

Rong C, H Chen, X Tang, Z Gu, J Zhao, H Zhang, W Chen, YQ Chen (2019). Characterization and molecular docking of new Δ17 fatty acid desaturase genes from *Rhizophagus irregularis* and *Octopus bimaculoides*. *RSC Adv* 9:6871–6880

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

Spano P, TM Fulton, J Chunwongesm, SD Tanksley (1995). Extraction of high quality DNA for genome sequencing. *Mol Biol Rep* 13:207

Ubersax JA, JE Ferrell Jr (2007). Mechanisms of specificity in protein phosphorylation. *Nat Rev Mol Cell Biol* 8:530-541

Uslupehlivan M, R Deveci, C Ün (2018). *In silico* investigation of the prion protein glycosylation profiles in relation to scrapie disease resistance in domestic sheep (*Ovis aries*). *Mol Cell Probes* 42:1–9

Webb K, JA Laganowsky, JP Whitelegge, SG Clarke (2008). Identification of two SET domain proteins required for methylation of lysine residues in Yeast ribosomal protein Rpl42ab. *J Biol Chem* 283:35561–35568

Winter DL, G Hart-Smith, MR Wilkins (2017). Characterization of protein methyltransferase Rkm1, Rkm4, Efm4, Efm7, Set5 and Hmt1 reveals extensive posttranslational modification. *J Mol Biol* 430:1102–1118

Yagoub D, G Hart-Smith, J Moecking, MA Erce, MR Wilkins (2015). Yeast Proteins Gar1p, Nop1p, Npl3p, Nsr1p, and Rps2p are natively methylated and are substrates of the arginine methyltransferase Hmt1p. *Proteomics* 15:3209–3218

Yang J, Y Zhang (2015). I-TASSER server: new development for protein structure and function predictions. *Nucl Acids Res* 43: 174–181

Yeates TO (2002). Structures of SET domain proteins: protein lysine methyltransferases make their mark. *Cell* 111:5–7

Zhang C, PL Freddolino, Y Zhang (2017). COFACTOR: improved protein function prediction by combining structure, sequence and protein–protein interaction information. *Nucl Acids Res* 45: 291–299

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

Zhu H, M Bilgin, R Bangham, D Hall, A Casamayor, P Bertone, N Lan, R Jansen, S Bidlingmaier, T Houfek, T Mitchell (2001). Global analysis of protein activities using proteome chips. *Science* 293:2101–2105