



**Full Length Article**

## **Analysis of Genetic Polymorphisms and Post Translational Modifications of Cytochrome C-1 in *Sordaria fimicola***

Syeda Hina Bukhari<sup>1\*</sup>, Iqra Mobeen<sup>1\*</sup>, Uzma Naureen<sup>1</sup>, Faiza Akram<sup>1</sup>, Rabia Arif<sup>1</sup>, Memuna Ghafoor Shahid<sup>2</sup> and Muhammad Saleem<sup>1</sup>

<sup>1</sup>Molecular Genetics Research Laboratory, Department of Botany, University of the Punjab, 54590 Lahore, Pakistan

<sup>2</sup>Department of Botany, Government College University, Lahore Pakistan

\*For correspondence: iqramobeen905@gmail.com; iqra.phd.botany@pu.edu.pk; hina17feb@gmail.com

Received 31 Jul 2019; Accepted 16 November 2019; 04 February 2020

### **Abstract**

Post-translational modifications produce multiple range of more specific chemical modifications not only between different species but also in the same species so their finding is helpful in describing remarkable protein diversity within species. The purpose of present study was to show the diversity in genetic makeup of *Sordaria fimicola* strains (South Facing Strains and North Facing Strains) by analyzing nucleotide and protein sequence variations of cytochrome C-1 gene and their subsequent proteins in comparison to reference organism to prove that stressful environment produce the changes in the genome sequence of an organism for which *CyC1* gene and protein sequence of *Neurospora crassa* was taken as model organism. This study investigated the six variations in nucleotide sequence and three variations in protein sequence of *CyC1*. To check the effect of these variations on protein, post translational modifications (PTMs) were predicted by using bioinformatics tools which revealed that variations in nucleotide sequence ultimately altered the PTMs sites of *CyC1* protein in different strains of *S. fimicola* (S<sub>1</sub>, S<sub>2</sub>, S<sub>3</sub>, N<sub>5</sub>, N<sub>6</sub> and N<sub>7</sub>) as SFS strains exhibited more changes in positions of PTMs than NFS strains. Some PTMs sites were also found missing in SFS strains which were present in NFS strains. The entire glycosylation sites (65T, 173T, 279T, 325T) were found conserved in *CyC1* of *S. fimicola* and only three phosphorylation sites (143T, 236S, 365S) were non conserved in *S. fimicola*. Except one acetylation site (151K), all sites were conserved in *S. fimicola* and *N. crassa*. There was an addition of methylation site (237K) in south strains of *S. fimicola* whereas remaining 11 sites were conserved in *S. fimicola*. All these results revealed that variation in PTMs sites of *CyC1* protein in South strains of *S. fimicola* were to modify the functions of these proteins to overcome the environmental stress faced by SFS strains (S<sub>1</sub>, S<sub>2</sub>, S<sub>3</sub>) of *S. fimicola*. © 2020 Friends Science Publishers

**Keywords:** Cytochrome C-1; Genetic diversity; Glycosylation; Phosphorylation; Post-translational modifications

### **Introduction**

The progress in the field of proteomics and bioinformatics has eased the determination of modified sites with detail about the molecules which are responsible for the attachment and detachment of functional groups (Chandramouli and Qian 2009) and this information is enough for the description of PTMs regulatory network in the cell (Shen 2013). More than 200 types of PTMs have been reported in Ascomycota (Yu *et al.* 2007). In eukaryotes, phosphorylation, methylation, glycosylation, acetylation, ubiquitylation, carboxylation and S-nitrosylation are most commonly studied modifications (Oliveira and Sauer 2011; Jimenez-Morales *et al.* 2013; Marquez *et al.* 2016). All these types of modifications consist of multiple ranges of more specific chemical modifications not only between different species but also in the same specie (Walsh *et al.* 2005; Walsh and Jefferis 2006).

Cytochrome c acts as an electron transporter, where cytochrome C-1 is the electron donor and cytochrome C

oxidase is electron acceptor and the cluster of positively charged lysine residues present on the outer side of cytochrome C is responsible for the binding of cytochrome C with membrane surface (Hough *et al.* 2014). Cytochrome C has a heme that is accountable for carrying electrons and then transporting them to cytochrome C oxidase (Yeagle 2016). Heme is covalently attached to a distinguishing CXXCH motif through protein-catalyzed post-translational modification reactions. Cytochrome C is also the initiating protein of intrinsic apoptosis pathway (Mavridou *et al.* 2013; Matsuura *et al.* 2016). Cytochrome C is among those proteins which have massive contribution in energy transduction processes like respiration and photosynthesis (Mavridou *et al.* 2013).

The current study was designed to define the post-translational modifications of cytochrome C-1(*CyC1*) in different strains of *Sordaria fimicola*- a model organism in the field of genetics (Zickler and Espagne 2016). Ascomycetes are the mostly studied organisms for

eukaryotic evolution due to their extensive phylogenetic range, excess of sequenced genomes record and advanced genome examination (Wohlbach *et al.* 2009). Sordariomycetes is one of the largest classes of Ascomycota and the dominant features of this class are perithecial ascomata and inoperculate unitunicate asci. 600 genera having over 3000 species are included in this class (Zhang *et al.* 2006). *Sordaria fimicola* is one of the members of Sordariomycetes, which is characterized by homothallic pyrenomyces having eight dark colored ascospores in each ascus, which have resemblance with *Neurospora crassa* asci (Olive 1956). Recent study will provide the genetic information of *CyC1* gene in *S. fimicola* strains which will provide content for future researches in the field of molecular genetics as *S. fimicola* is now taken as model organism in many genetic research works.

## Materials and Methods

### Sub culturing of different strains of *S. fimicola*

Fungal strains of *S. fimicola* were provided by Molecular Genetics Research Laboratory of Botany Department, PU, Lahore for sub-culturing from their pure stock culture (Stored at -20°C) which were primarily isolated from “Evolutionary Canyon” by Prof. Nevo. N strains were collected from three locations on NFS (North Facing Slope) and named N<sub>5</sub>, N<sub>6</sub>, N<sub>7</sub>, while S strains were collected from three stations on SFS (South Facing Slope) *i.e.*, S<sub>1</sub>, S<sub>2</sub> and S<sub>3</sub>. During the present research work, Parental strains of North and South sides were sub cultured for DNA extraction. Six parental strains of *S. fimicola* were sub-cultured on PDA media under aseptic conditions to avoid contamination and incubated at 18°C in thermostat incubator to get perithecia in 12–14 days for DNA extraction.

### Isolation of DNA

Modified Pietro *et al.* (1995) method was adopted for isolating DNA from six parental strains of *S. fimicola*. Method proposed by Hoisington *et al.* (1994) was applied for the verification of DNA presence after DNA extraction. Standard Agarose Gel Electrophoresis stained with 1% (w/v) ethidium bromide was carried out.

### Primer designing and preparation

Primers (forward and reverse) were designed for to amplify the selected genes with the help of NCBI Primer Blast and Primer 3 Plus tool. Made their stock concentration (100  $\mu$ M) and working stock of 10  $\mu$ M which finally stored at -20°C.

### PCR amplification of targeted genes

Touchdown PCR conditions used for the amplification of *CyC1* gene (Korbie and Mattick 2008). Touchdown PCR was run with 15  $\mu$ L reaction mixture containing 2  $\mu$ L DNA sample, 10  $\mu$ L Master Mix (GeneAll), 1  $\mu$ L Forward primer,

1  $\mu$ L Reverse primer and 1  $\mu$ L Double distilled water.

For the confirmation of the amplification of targeted genes 1% agarose gel electrophoresis was done (Hoisington *et al.* 1994). Gel was exposed to UV light and bands were visualized by Gel documentation system and comparison of amplified bands was done with the standard 1 Kb plus DNA Ladder which was loaded in the first well of agarose gel.

### Gene sequencing of PCR products

PCR products were sent to Macrogen, Korea for direct sequencing of *CyC1* gene. *N. crassa* was used as reference organism. Protein sequences of gene under study were obtained with the help of online server EMBOSS Transeq ([https://www.ebi.ac.uk/Tools/st/emboss\\_transeq/](https://www.ebi.ac.uk/Tools/st/emboss_transeq/)). Nucleotide sequence of reference organism's *CyC1* gene was obtained from NCBI (<http://blast.ncbi.nlm.nih.gov/Blast/>) and protein sequence from Uniprot database (<http://www.uniprot.org/>).

### Analysis of sequenced data

For checking the nucleotide and amino acid variations in gene and protein sequences of *S. fimicola* (all strains) and *N. crassa*, Clustal Omega (<http://ebi.ac.uk/Tools/msa/clustalo/>) and SIAS: Sequence Identity and Similarity (<http://imed.med.ucm.es/Tools/sias.html>) were used.

### Prediction and comparative evaluation of PTM's by PTM prediction servers

Online servers *i.e.*, YinOYang 1.2 ([www.cbs.dtu.dk/services/YinOYang/](http://www.cbs.dtu.dk/services/YinOYang/)), NetPhos 3.1 ([www.cbs.dtu.dk/services/NetPhos/](http://www.cbs.dtu.dk/services/NetPhos/)), PAIL ([bdmpail.biocuckoo.org/](http://bdmpail.biocuckoo.org/)) and ModPred ([montana.informatics.indiana.edu/ModPred/faq.html](http://montana.informatics.indiana.edu/ModPred/faq.html)) were used to define the types of post-translational modifications in proteins of interest as each server results show prediction for specific kind of modification.

## Results

Genetic diversity at 6 positions of *CyC1* gene in reference specie and all strains of *Sordaria fimicola* was determined by multiple sequence alignment (Fig. 1) which showed that all the 6 point mutations were in nucleotide sequence of South Facing Slope (SFS) strains of *S. fimicola*. The result of clustal multiple sequence alignment for *CyC1* protein sequence depicted the polymorphism in protein sequence at 3 positions of amino acid (Fig. 2).

The results of YinOYang 1.2 server (Table 1) showed that there is glycosylation on threonine residue at four positions *i.e.*, 65, 173, 279 and 325 in both *N. crassa* and *S. fimicola* strains. No YinOYang site was present in *CyC1*. Graphical representation of glycosylation in *CyC1* protein of *N. crassa*, S<sub>1</sub>, S<sub>2</sub>, S<sub>3</sub>, N<sub>5</sub>, N<sub>6</sub>, N<sub>7</sub> is given in Fig. 3. NetPhos

**Table 1:** Comparison of *CyC1* glycosylation sites, phosphorylation sites on serine, threonine and tyrosine residues in *N. crassa* and *S. fomicola* (S<sub>1</sub>, S<sub>2</sub>, S<sub>3</sub>, N<sub>5</sub>, N<sub>6</sub>, N<sub>7</sub>)

Organism	Residues	<i>CyC1</i>	
		Glycosylation Sites	Phosphorylation Sites
<i>N. crassa</i>	S	-	10, 21, 24, 33, 68, 89, 91, 96, 99, 101, 102, 117, 123, 127, 166, 230, 260, 271, 283, 330, 365
	T	65, 173, 279, 325	80, 106, 164, 173, 182, 187, 202, 220, 241, 253, 281, 295, 300, 350, 359
	Y	-	170, 172
South strains	S	-	10, 21, 24, 33, 68, 89, 91, 96, 99, 101, 102, 117, 123, 127, 166, 230, 236, 260, 271, 283, 330
	T	65, 173, 279, 325	80, 106, 143, 164, 173, 182, 187, 202, 220, 241, 253, 281, 295, 300, 350, 359
	Y	-	170, 172
North strains	S	-	10, 21, 24, 33, 68, 89, 91, 96, 99, 101, 102, 117, 123, 127, 166, 230, 260, 271, 283, 330, 365
	T	65, 173, 279, 325	80, 106, 164, 173, 182, 187, 202, 220, 241, 253, 281, 295, 300, 350, 359
	Y	-	170, 172

N_e_CyC1	TGTGGCCGGGATCCCCTCACTTCCACTCTGGTTTGAACCTTTTGTTCCTACTATT	240
N5_CyC1	TGTGGCCGGGATCCCCTCACTTCCACTCTGGTTTGAACCTTTTGTTCCTACTATT	240
N6_CyC1	TGTGGCCGGGATCCCCTCACTTCCACTCTGGTTTGAACCTTTTGTTCCTACTATT	240
N7_CyC1	TGTGGCCGGGATCCCCTCACTTCCACTCTGGTTTGAACCTTTTGTTCCTACTATT	240
S1_CyC1	TGTGGCCGGGATCCCCTCACTTCCACTCTGGTTTGAACCTTTTGTTCCTACTATT	240
S2_CyC1	TGTGGCCGGGATCCCCTCACTTCCACTCTGGTTTGAACCTTTTGTTCCTACTATT	240
S3_CyC1	TGTGGCCGGGATCCCCTCACTTCCACTCTGGTTTGAACCTTTTGTTCCTACTATT	240
*****		
N_e_CyC1	TGCGCTCAGTGCCACACCCCTTGAGGAGGGCGGGCAACAAGATCGGCCCGCTCTTCAC	480
N5_CyC1	TGCGCTCAGTGCCACACCCCTTGAGGAGGGCGGGCAACAAGATCGGCCCGCTCTTCAC	480
N6_CyC1	TGCGCTCAGTGCCACACCCCTTGAGGAGGGCGGGCAACAAGATCGGCCCGCTCTTCAC	480
N7_CyC1	TGCGCTCAGTGCCACACCCCTTGAGGAGGGCGGGCAACAAGATCGGCCCGCTCTTCAC	480
S1_CyC1	TGCGCTCAGTGCCACACCCCTTGAGGAGGGCGGGCAACAAGATCGGCCCGCTCTTCAC	480
S2_CyC1	TGCGCTCAGTGCCACACCCCTTGAGGAGGGCGGGCAACAAGATCGGCCCGCTCTTCAC	480
S3_CyC1	TGCGCTCAGTGCCACACCCCTTGAGGAGGGCGGGCAACAAGATCGGCCCGCTCTTCAC	480
*****		
N_e_CyC1	ATCATCACCTTCATGAAGGAGGCTACTGCTTAAATGCAATCTGTTGATGATGGGCGTTG	720
N5_CyC1	ATCATCACCTTCATGAAGGAGGCTACTGCTTAAATGCAATCTGTTGATGATGGGCGTTG	720
N6_CyC1	ATCATCACCTTCATGAAGGAGGCTACTGCTTAAATGCAATCTGTTGATGATGGGCGTTG	720
N7_CyC1	ATCATCACCTTCATGAAGGAGGCTACTGCTTAAATGCAATCTGTTGATGATGGGCGTTG	720
S1_CyC1	ATCATCACCTTCATGAAGGAGGCTACTGCTTAAATGCAATCTGTTGATGATGGGCGTTG	720
S2_CyC1	ATCATCACCTTCATGAAGGAGGCTACTGCTTAAATGCAATCTGTTGATGATGGGCGTTG	720
S3_CyC1	ATCATCACCTTCATGAAGGAGGCTACTGCTTAAATGCAATCTGTTGATGATGGGCGTTG	720
*****		
N_e_CyC1	TTCTCGAGGAGTTATGGGACTGTATTAATAAAAAGGGAGATTTTTTTTTGCATTAGACG	780
N5_CyC1	TTCTCGAGGAGTTATGGGACTGTATTAATAAAAAGGGAGATTTTTTTTTGCATTAGACG	780
N6_CyC1	TTCTCGAGGAGTTATGGGACTGTATTAATAAAAAGGGAGATTTTTTTTTGCATTAGACG	780
N7_CyC1	TTCTCGAGGAGTTATGGGACTGTATTAATAAAAAGGGAGATTTTTTTTTGCATTAGACG	780
S1_CyC1	TTCTCGAAGAGTTATGGGACTGTATTAATAAAAAGGGAGATTTTTTTTTGCATTAGACG	780
S2_CyC1	TTCTCGAAGAGTTATGGGACTGTATTAATAAAAAGGGAGATTTTTTTTTGCATTAGACG	780
S3_CyC1	TTCTCGAAGAGTTATGGGACTGTATTAATAAAAAGGGAGATTTTTTTTTGCATTAGACG	780
*****		
N_e_CyC1	AATACCCCAATCCCCCTTACACCCCTTCTTGATCCCGTCTCGCCGTGTATGTATGGT	960
N5_CyC1	AATACCCCAATCCCCCTTACACCCCTTCTTGATCCCGTCTCGCCGTGTATGTATGGT	960
N6_CyC1	AATACCCCAATCCCCCTTACACCCCTTCTTGATCCCGTCTCGCCGTGTATGTATGGT	960
N7_CyC1	AATACCCCAATCCCCCTTACACCCCTTCTTGATCCCGTCTCGCCGTGTATGTATGGT	960
S1_CyC1	AATACCCCAATCCCCCTTACACCCCTTCTTGATCCCGTCTCGCCGTGTATGTATGGT	960
S2_CyC1	AATACCCCAATCCCCCTTACACCCCTTCTTGATCCCGTCTCGCCGTGTATGTATGGT	960
S3_CyC1	AATACCCCAATCCCCCTTACACCCCTTCTTGATCCCGTCTCGCCGTGTATGTATGGT	960
*****		
N_e_CyC1	AGGACTGTCATATCTAAGTGCATTCAAAATACAATAAAGCATCTCTGCTACCGGATAAT	1140
N5_CyC1	AGGACTGTCATATCTAAGTGCATTCAAAATACAATAAAGCATCTCTGCTACCGGATAAT	1140
N6_CyC1	AGGACTGTCATATCTAAGTGCATTCAAAATACAATAAAGCATCTCTGCTACCGGATAAT	1140
N7_CyC1	AGGACTGTCATATCTAAGTGCATTCAAAATACAATAAAGCATCTCTGCTACCGGATAAT	1140
S1_CyC1	AGGACTGTCATATCTAAGTGCATTCAAAATACAATAAAGCATCTCTGCTACCGGATAAG	1140
S2_CyC1	AGGACTGTCATATCTAAGTGCATTCAAAATACAATAAAGCATCTCTGCTACCGGATAAG	1140
S3_CyC1	AGGACTGTCATATCTAAGTGCATTCAAAATACAATAAAGCATCTCTGCTACCGGATAAG	1140

**Fig. 1:** Multiple sequence alignment of *CyC1* gene in *S. fomicola* strains (S<sub>1</sub>, S<sub>2</sub>, S<sub>3</sub>, N<sub>5</sub>, N<sub>6</sub> and N<sub>7</sub>) and *N. crassa* (reference organism). Asterisk (\*) depict the conserved sites between *CyC1* of *S. fomicola* and *N. crassa* whereas space or gap depict the polymorphic sites which are also highlighted

3.1 server results revealed that there was phosphorylation on *CyC1* of *N. crassa*, S<sub>1</sub>, S<sub>2</sub>, S<sub>3</sub>, N<sub>5</sub>, N<sub>6</sub> and N<sub>7</sub> on three types of residues i.e., Serine (S), Threonine (T), and Tyrosine (Y). Phosphorylation sites (10, 21, 24, 33, 68, 89, 91, 96, 99, 101, 102, 117, 123, 127, 166, 230, 260, 271, 283, 330) were found same on serine residue (Table 1). All other strains except SFS strains also have phosphorylation at 365S. Phosphorylation sites on threonine (T) residue were 80, 106, 164, 173, 182, 187, 202, 220, 241, 253, 281, 295, 300, 350, 359 in *N. crassa* and *S. fomicola* but SFS strains of *S. fomicola* have an additional site at position 143T. Graphical

representation of phosphorylation with these positions is given in Fig. 4.

Acetylation was observed at positions 17, 18, 119, 128, 129, 135, 163, 203, 210, 211, 213, 290, 323, 361, and 368 on internal lysine (K) in *N. crassa* and *S. fomicola*. *N. crassa*, N<sub>5</sub>, N<sub>6</sub> and N<sub>7</sub> have an additional acetylation at 151 K (Table 2). Methylation prediction results showed methylation on 135, 151, 163, 177, 179, 196, 197, 210, 211, 213, 335) lysine residues in *CyC1* of *N. crassa* and NFS strains whereas SFS strains have 12 methylation sites (Table 2).

N_c_CyC1_1	CGRDPVTFHSWFATFLFLLFLTPHRSFFFLSLSPFLFSVISTSSIPYTPLR* <sup>*</sup> RFLINQS	117
N5_CyC1_1	CGRDPVTFHSWFATFLFLLFLTPHRSFFFLSLSPFLFSVISTSSIPYTPLR* <sup>*</sup> RFLINQS	117
N6_CyC1_1	CGRDPVTFHSWFATFLFLLFLTPHRSFFFLSLSPFLFSVISTSSIPYTPLR* <sup>*</sup> RFLINQS	117
N7_CyC1_1	CGRDPVTFHSWFATFLFLLFLTPHRSFFFLSLSPFLFSVISTSSIPYTPLR* <sup>*</sup> RFLINQS	117
S1_CyC1_1	CGRDPVTFHSLFATFLFLLFLTPHRSFFFLSLSPFLFSVISTSSIPYTPLR* <sup>*</sup> RFLINQS	117
S2_CyC1_1	CGRDPVTFHSLFATFLFLLFLTPHRSFFFLSLSPFLFSVISTSSIPYTPLR* <sup>*</sup> RFLINQS	117
S3_CyC1_1	CGRDPVTFHSLFATFLFLLFLTPHRSFFFLSLSPFLFSVISTSSIPYTPLR* <sup>*</sup> RFLINQS	117
*****		
N_c_CyC1_1	QKGITWDENTLFEYLENPKKYIPGTMAFGGLKKDKDRNDIITFMKEATA* <sup>*</sup> MQSV**WAL	234
N5_CyC1_1	QKGITWDENTLFEYLENPKKYIPGTMAFGGLKKDKDRNDIITFMKEATA* <sup>*</sup> MQSV**WAL	234
N6_CyC1_1	QKGITWDENTLFEYLENPKKYIPGTMAFGGLKKDKDRNDIITFMKEATA* <sup>*</sup> MQSV**WAL	234
N7_CyC1_1	QKGITWDENTLFEYLENPKKYIPGTMAFGGLKKDKDRNDIITFMKEATA* <sup>*</sup> MQSV**WAL	234
S1_CyC1_1	QKGITWDENTLFEYLENPKKYIPGTMAFGGLKKDKDRNDIITFMKEATA* <sup>*</sup> MQSV**GAL	234
S2_CyC1_1	QKGITWDENTLFEYLENPKKYIPGTMAFGGLKKDKDRNDIITFMKEATA* <sup>*</sup> MQSV**GAL	234
S3_CyC1_1	QKGITWDENTLFEYLENPKKYIPGTMAFGGLKKDKDRNDIITFMKEATA* <sup>*</sup> MQSV**GAL	234
*****		
N_c_CyC1_1	NTPPFLTPPFLIPFLPCMYGLVLDGHARL.KATE* <sup>*</sup> SLVSFF* <sup>*</sup> AFKSEAGIFQI***CYKEW	348
N5_CyC1_1	NTPPFLTPPFLIPFLPCMYGLVLDGHARL.KATE* <sup>*</sup> SLVSFF* <sup>*</sup> AFKSEAGIFQI***CYKEW	348
N6_CyC1_1	NTPPFLTPPFLIPFLPCMYGLVLDGHARL.KATE* <sup>*</sup> SLVSFF* <sup>*</sup> AFKSEAGIFQI***CYKEW	348
N7_CyC1_1	NTPPFLTPPFLIPFLPCMYGLVLDGHARL.KATE* <sup>*</sup> SLVSFF* <sup>*</sup> AFKSEAGIFQI***CYKEW	348
S1_CyC1_1	NTPPFLTPRFLIPFLPCMYGLVLDGHARL.KATE* <sup>*</sup> SLVSFF* <sup>*</sup> AFKSEAGIFQI***CYKEW	348
S2_CyC1_1	NTPPFLTPRFLIPFLPCMYGLVLDGHARL.KATE* <sup>*</sup> SLVSFF* <sup>*</sup> AFKSEAGIFQI***CYKEW	348
S3_CyC1_1	NTPPFLTPRFLIPFLPCMYGLVLDGHARL.KATE* <sup>*</sup> SLVSFF* <sup>*</sup> AFKSEAGIFQI***CYKEW	348
*****		

**Fig. 2:** Multiple sequence alignment of *CyC1* protein sequence of *S. fomicola* strains and *N. crassa*. Space/gap shows the polymorphic sites, asterisk (\*) shows the position having single or full conserved residue, colon (:) shows conserve sites between groups of strongly similar properties whereas dot (.) shows conserve sites between groups of weakly similar properties

## Discussion

In present study genetic polymorphism (Fig. 1–2) and variation in PTMs position in *CyC1* within specie of *Sordaria fomicola* and with reference specie (Table 1–2) were predicted by using bioinformatics tools as Genomic study of an organism is mostly done by using molecular techniques in combination with advance technologies of proteomics and metabolomics specifically bioinformatics tools (to save time and money by analyzing a bulk of data through computational techniques) in order to get more precise results (Vega *et al.* 2012). Jamil *et al.* (2018) and Rana *et al.* (2018) also used these strategies to predict PTMs of Histone 3 and 4 protein, Manganese oxide super dismutase (MnSOD) protein respectively in different strains of *S. fomicola*.

The specific regions of gene may have polymorphism in nucleotide sequences among same species and closely related species (Park *et al.* 2007). The current results depicted genetic variation on different nucleotides of *S. fomicola* when compared with reference organism (*N. crassa*) for *CyC1* gene. Presently, specific variations in the nucleotide sequences at 6 positions out of total 1154 nucleotides of *CyC1* gene have been found (Fig. 1). All of these 6 variations are observed in SFS strains of *S. fomicola*.

As nucleotide variations of gene eventually effect the composition of protein (Amino acid sequence) translated from that gene, so, the prediction of possible effects of nucleotide variation on protein stability, folding and its function, the detection of amino acid variations in protein is one of the promising approaches (Cardoso *et al.* 2015). The present results of protein sequences alignment of *CyC1* of *N. crassa*, S<sub>1</sub>, S<sub>2</sub>, S<sub>3</sub>, N<sub>5</sub>, N<sub>6</sub>, N<sub>7</sub> revealed that amino acid sequences of *CyC1* have 3 variations in south strains (Fig. 2) along with three sites of conserved amino acids. Similarly, Nevo (2001) investigated the genetic diversity (both of allozyme and DNA) of different model organisms (*Nostoc linckia*, *Sordaria fomicola*, *Triticum dicoccoides*, *Hordeum spontaneum*, *Drosophila melanogaster*, *Spalax ehrenbergi*

superspecies, *Acomys cahirinus* and *Apodemus mystacinus*) belong to different taxa, collected from evolution canyon (EC) and obtained results similar to our present work that south facing slope organisms have much diversity in DNA than north facing slope and heritable mutation rates were found three folds higher in *S. fomicola*. Lamb *et al.* (2008) also compared polymorphic sites of allozymes and DNA markers. Their results illustrated that genetic variability (Polymorphism) was prominent (64%) in South Facing Slope model organisms (*Penicillium lanosum* and *Aspergillus niger*) due to stressful environment in Evolution Canyon 1, which cleared that stressful environment is directly proportional to genetic diversity.

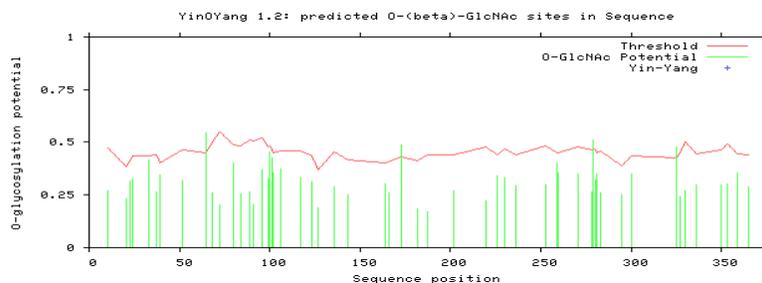
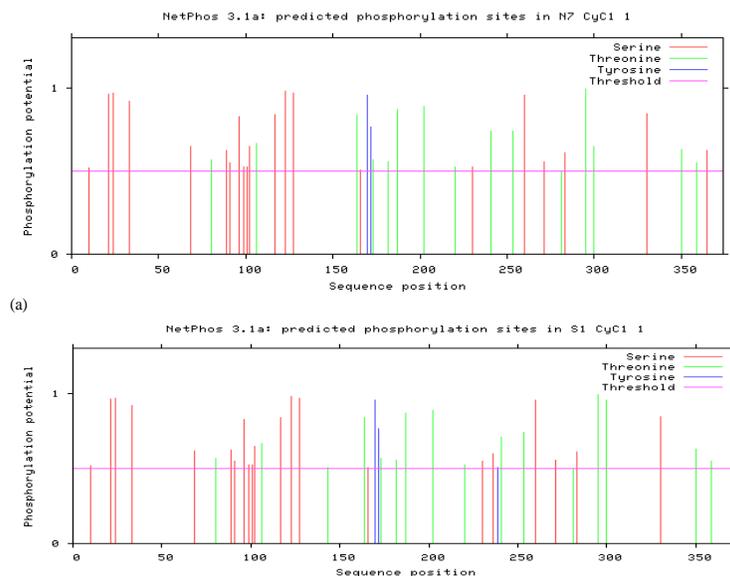
Modification of proteins by glycosylation is functionally also very important for cells as chemical glycosylation of *CyC1* causes increase in stability and its application (Delgado *et al.* 2014). YinOYang results predicted glycosylation at positions 65, 173, 279, 325 on (T) threonine residues in *CyC1* of *N. crassa* and *S. fomicola*. Results showed that there was not a single glycosylation modification on serine residue in reference strain and *S. fomicola* strains and all glycosylation sites were found conserved on *CyC1* in *S. fomicola* (Table 1).

Kinase mediated phosphorylation is included in the most essential post-translational modifications in eukaryotes (Dou *et al.* 2014). Its function is the rearrangement of all intracellular processes and it is responsible for the addition of phosphate groups on the side chain of amino acids in proteins (Beene and Scott 2007; Patwardhan and Miller 2007). Sanderson *et al.* (2013) mentioned the sites of Phosphorylation of *CyC1* in liver and heart tissue. Phosphorylation modification occurred at position 48 on tyrosine (Y) residue in liver and at position 97 on (Y) tyrosine residue in heart tissue. In present work *N. crassa* and *S. fomicola* have phosphorylation at position 96 on serine residue which is close to the position reported by Sanderson *et al.* (2013) in liver and heart tissues.

Presently *CyC1* has 38 phosphorylation sites in *N. crassa* and N<sub>5</sub>, N<sub>6</sub>, N<sub>7</sub> strains of *S. fomicola* in which 21 were

**Table 2:** Comparison of *CyC1* acetylation sites, methylation sites on lysine and arginine residues in *N. crassa* and *S. fimicola* (S<sub>1</sub>, S<sub>2</sub>, S<sub>3</sub>, N<sub>5</sub>, N<sub>6</sub>, N<sub>7</sub>)

<i>CyC1</i>			
Organism	Residue	Acetylation Sites	Methylation Sites
<i>N. crassa</i>	K	17, 18, 119, 128, 129, 135, 151, 163, 203, 210, 211, 213, 290, 323, 361	135, 151, 163, 177, 179, 196, 197, 210, 211, 213, 335
South strains	K	17, 18, 119, 128, 129, 135, 163, 203, 210, 211, 213, 290, 323, 361, 368	135, 151, 163, 177, 179, 196, 197, 210, 211, 213, 237, 335
North strains	K	17, 18, 119, 128, 129, 135, 151, 163, 203, 210, 211, 213, 290, 323, 361	135, 151, 163, 177, 179, 196, 197, 210, 211, 213, 335

**Fig. 3:** Graphical representation of glycosylation potential of *CyC1* protein in *N. crassa* and *S. fimicola* (S<sub>1</sub>, S<sub>2</sub>, S<sub>3</sub>, N<sub>5</sub>, N<sub>6</sub>, N<sub>7</sub>). Red zigzag line shows the threshold (0.5), green lines shows the O-glycosylation potential. Green lines exceeding threshold shows the positions of glycosylation**Fig. 4:** Graphical representation of phosphorylation potential of *CyC1* protein (a) In *N. crassa*, N<sub>5</sub>, N<sub>6</sub> and N<sub>7</sub> (b) In S<sub>1</sub>, S<sub>2</sub> and S<sub>3</sub>. Red Lines= Serine, Green Lines= Threonine, Blue Lines= Tyrosine and Pink Line= Threshold. Lines exceeding pink line (threshold= 0.5) shows the positions of phosphorylation in *CyC1* and red, green, blue color of lines shows the residue present at that position

on serine residue, 15 were on threonine residue and 2 on tyrosine residue (Table 1). About 39 phosphorylation sites were predicted in SFS strains out of which 21 were on serine residue, 16 on threonine and 2 on tyrosine residue, while 20 sites on serine residues, 15 sites on threonine residues and 2 sites on tyrosine residues were conserved in *N. crassa* and *S. fimicola*.

Lower eukaryotes like *S. cerevisiae* have about 50% acetylated proteins (Polevoda and Sherman 2003). Similarly in current study *CyC1* protein has acetylation sites. The results of *CyC1* in *N. crassa* and *S. fimicola* strains indicated 15 acetylated sites out of which 14 sites on lysine (K)

residues were conserved in all of these but one modification position was different in south strains (S<sub>1</sub>, S<sub>2</sub>, S<sub>3</sub>) which is 368K and position of acetylation in *N. crassa*, N<sub>5</sub>, N<sub>6</sub> and N<sub>7</sub> was found to be 151K which is absent in SFS strains of *S. fimicola* (Table 2).

A number of researchers have tested out the methylation of cytochrome C at specific position 77 on lysine residue of lower eukaryotes as a model for methylation studies (Paik *et al.* 1989). Farooqui *et al.* (1981) worked on the *in vivo* stability of methylated and unmethylated cytochrome C in *S. cerevisiae*. Their results showed constant rate of degradation of unmethylated

cytochrome C where as in case of methylated cytochrome C the rate of degradation was three times lower than unmethylated species which confirmed the importance of methylation in cytochrome C protection against the intracellular proteolytic enzyme attack. In present results 11 methylation sites were found in *N. crassa*, N<sub>5</sub>, N<sub>6</sub> and N<sub>7</sub>, 12 sites in S<sub>1</sub>, S<sub>2</sub> and S<sub>3</sub> in *CyC1*. The additional methylation site in SFS strains of *S. fomicola* was 237K which was missing in other strains (Table 2). Alike Paik et al. (1989) reported that *N. crassa* cytochrome C molecule was fully methylated however; in case of *S. cerevisiae* 21 positions were methylated and unmethylated sites were also present. DiMaria et al. (1979) also mentioned position of cytochrome C methylation in horse heart which was 72 on lysine residue. This position of a single trimethylated residue at position 72 was also seen in cytochrome C of *S. cerevisiae*. Similarly, Winter et al. (2015) reported methylation site of *CyC1* at position 78 on lysine residue which has trimethylation. Molecular analysis and PTMs prediction results of *CyC1* in present research indicate that south strains of *S. fomicola* have more variations in their genetic makeup and PTMs are also inducing significant changes in mature proteins, and stressful environmental conditions might be an ultimate factor to induce such variations as reported by Arif et al. (2017).

## References

- Arif R, F Akram, T Jamil, H Mukhtar, SF Lee, M Saleem (2017). Genetic variation and its reflection on post translational modifications in frequency clock and mating type a-1 proteins in *Sordaria fomicola*. *BioMed Res Intl* 3:1–10
- Beene DL, JD Scott (2007). A kinase anchoring proteins take shape. *Curr Opin Cell Biol* 19:192–198
- Cardoso JG, RMR Andersen, MJ Herrgård, N Sonnenschein (2015). Analysis of genetic variation and potential applications in genome-scale metabolic modeling. *Bioeng Biotechnol* 3:1–12
- Chandramouli K, P Qian (2009). Proteomics: challenges, techniques and possibilities to overcome biological sample complexity. *Hum Genom Proteom* 2009:1–22
- Delgado Y, M Morales-Cruz, J Hernandez-Roman, Y Martinez, K Griebenow (2014). Chemical glycosylation of cytochrome c improves physical and chemical protein stability. *BMC Biochem* 15:1–11
- DiMaria P, E Polastro, RJ DeLange, S Kim, WK Paik (1979). Studies on cytochrome c methylation in yeast. *J Biol Chem* 254:4645–4652
- Dou Y, B Yao, C Zhang (2014). Phospho SVM: Prediction of phosphorylation sites by integrating various protein sequences attributes with a support vector machine. *Amino Acids* 46:1459–1469
- Farooqui J, P DiMaria, S Kim, WK Paik (1981). Effect of methylation on the stability of cytochrome c of *Saccharomyces cerevisiae* in vivo. *J Biol Chem* 256:5041–5045
- Hoisington R, M Khairallah, D Gonzalez (1994). *Laboratory Protocols: CIMMYT Applied Biotechnology Center*, El Batán, Mexico
- Hough MA, G Silkstone, JA Worrall, MT Wilson (2014). No binding to the proapoptotic cytochrome c-cardiolipin complex. *Vitam Horm* 96:193–209
- 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 fomicola*. *Intl J Agric Biol* 20:1021–1026
- Jimenez-Morales D, L Adamian, D Shi, J Liang (2013). Lysine carboxylation: unveiling a spontaneous post-translational modification. *Acta Crystallogr D Biol Crystallogr* 70:48–57
- Korbie DJ, JS Mattick (2008). Touchdown PCR for increased specificity and sensitivity in PCR amplification. *Nat Prot* 3:1452–1456
- Lamb BC, S Mandaokar, B Bahoun, I Grishkan, E Nevo (2008). Differences in spontaneous mutation frequencies as a function of environmental stress in soil fungi at Evolution Canyon, Israel. *Proc Natl Acad Sci USA* 105:5792–5796
- Marquez J, SR Lee, N Kim, J Han (2016). Post-translational modifications of cardiac mitochondrial proteins in cardiovascular disease: not lost in translation. *Kor Circ J* 46:1–12
- Matsuura K, K Canfield, W Feng, M Kurokawa (2016). Metabolic regulation of apoptosis in cancer. *Intl Rev Cell Mol Biol* 327:43–87
- Mavridou DAI, SJ Ferguson, JM Stevens (2013). Cytochrome C assembly. *IUBMB Life* 65:209–216
- Nevo E (2001). Evolution of genome-phenome diversity under environmental stress. *Proc Natl Acad Sci USA* 98:6233–6240
- Olive LS (1956). Genetics of *Sordaria fomicola* L. ascospore color mutants. *Amer J Bot* 43:97–107
- Oliveira AP, U Sauer (2011). The importance of post-translational modifications in regulating *Saccharomyces cerevisiae* metabolism. *FEMS Yeast Res* 12:104–117
- Park EJ, S Fukuda, H Endo, Y Kitade, N Saga (2007). Genetic polymorphism within *Porphyra yezoensis* (Bangiales, Rhodophyta) and related species from Japan and Korea detected by cleaved amplified polymorphic sequence analysis. *Eur J Phycol* 42:29–40
- Paik WK, YB Cho, B Frost, S Kim (1989). Cytochrome C methylation. *Biochem Cell Biol* 67:602–611
- Patwardhan P, WT Miller (2007). Processive phosphorylation: mechanism and biological importance. *Cell Signal* 19:2218–2226
- Pietro S, TM Fulton, J Chunwongesem, SD Tanksley (1995). Extraction of high quality DNA for genome sequencing. *Mol Biol Rep* 13:207
- Polevoda B, F Sherman (2003). N-terminal acetyltransferases and sequence requirements for N-terminal acetylation of eukaryotic proteins. *J Mol Biol* 325:595–622
- Rana J, R Arif, M Ishfaq, SF Lee, M Saleem (2018). Amplification and posttranslational modification study of manganese superoxide dismutase gene in *Sordaria fomicola*. *Intl J Agric Biol* 20:2059–2063
- Sanderson TH, G Mahapatra, P Pecina, Q Ji, K Yu, C Sinkler, A Varughese, R Kumar, MJ Bukowski, RN Tousignant, AR Salomon, L Lee, M Huttemann (2013). Cytochrome C is tyrosine 97 phosphorylated by neuroprotective insulin treatment. *PLoS One* 8:91–9
- Shen B (2013). *Bioinformatics for Diagnosis, Prognosis and Treatment of Complex Diseases*, Vol. 4. Springer Science + Business Media, Berlin, Germany
- Vega IE, D Rittschof, GH Dickinson, I Musgrave (2012). Evolutionary proteomics: empowering tandem mass spectrometry and bioinformatics tools for the study of evolution tandem mass spectrometry. In: *Tandem Mass Spectrophotometry Applications and Principle*. Prasain JK (eds.). Intech Open, London
- Walsh CT, SG Tsodikova, GJJ Gatto (2005). Protein post translational modifications: the chemistry of proteome diversifications. *Angew Chem Intl* 44:7342–7372
- Walsh G, R Jefferis (2006). Post-translational modifications in the context of therapeutic proteins. *Nat Biotechnol* 24:1241–1252
- Winter DL, D Abeygunawardena, G Hart-Smith, MA Erce, MR Wilkins (2015). Lysine methylation modulates the protein-protein interactions of yeast cytochrome c CyC1p. *Proteomics* 15:2166–2176
- Wohlbach DJ, DA Thompson, AP Gasch, A Regev (2009). From elements to modules: regulatory evolution in ascomycota fungi. *Curr Opin Genet Dev* 19:571–578
- Yeagle PL (2016). Membrane proteins. In: *The Membranes of Cells*, 3<sup>rd</sup> edn, pp: 219–268. Academic Press, Cambridge, Massachusetts, USA
- Yu LR, HJ Issaq, TD Veenstra (2007). Phosphoproteomics for the discovery of kinases as cancer biomarkers and drug targets. *Proteomics* 1:1042–1057
- Zhang N, LA Castlebury, AN Miller, SM Huhndorf, CL Schoch, KA Seifert, AY Rossman, JD Rogers, J Kohlmeyer, BV Kohlmeyer, G Sung (2006). An overview of the systematics of the Sordariomycetes based on a four-gene phylogeny. *Mycologia* 98:1076–1087
- Zickler D, E Espagne (2016). *Sordaria*, a model system to uncover links between meiotic pairing and recombination. *Semin Cell Dev Biol* 54:149–157