



**Full Length Article**

## Genetic Variation and Post-Translational Modifications of Cytochrome C Oxidase-1 (*COX1*) in different Strains of *Sordaria fimicola*

Rabia Arif<sup>1\*</sup>, Syeda Hina Bukhari<sup>1</sup>, Muhammad Ishfaq<sup>4</sup>, Memuna Ghafoor Shahid<sup>2</sup>, Siu Fai Lee<sup>3</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, 54000 Lahore, Pakistan

<sup>3</sup>Department of Biological Sciences, Macquarie University and CSIRO Black Mountain Laboratories, Building 101, Clunies Ross Street, Black Mountain, ACT 2601, Australia

<sup>4</sup>Institute of Molecular Biology and Biotechnology, The University of Lahore, Defence Road Campus, Lahore, Pakistan

\*For correspondence: phdgenetics@gmail.com

### Abstract

Post translational modifications, not only alter the protein structure but also its dynamics due to which there is a definite change in the performance of that protein resulting in remarkable diversity and complexity. The main purpose behind this study was to point out genetic variations in Cytochrome c oxidase-1 at gene level, protein level and modifications of protein in six parental strains and subsequent generations (F3 and F4) of *Sordaria fimicola*. All genetic variations and protein modifications were compared with reference organism (*Neurospora crassa*). Overall, point mutations on 10 different positions were observed in all strains when compared with reference organism, 7 of these were exhibited by the strains isolated from harsh environmental conditions while variations on 3 positions were common between strains of two opposite environments. Various post-translational modification predictor tools were used to predict post-translational modifications. The accuracy of results of PTMs can be enhanced by combining some bioinformatics techniques with mass spectrophotometry, which fascinate the researchers who are working in the field of proteomics. © 2019 Friends Science Publishers

**Keywords:** Evolution canyon; Glycosylation; Nucleotide; Polymorphisms; Phosphorylation

### Introduction

In eukaryotes, proteins after translation need some chemical alterations to modify their functions. Post translational modifications (PTMs) enhance the role of proteins by linking up with other biochemical functional groups *i.e.*, acetate, phosphate, various lipids and carbohydrates and by altering the chemical composition and structure of proteins such as the formation of disulfide bridges etc. These modifications, not only alter the protein structure but also its dynamics due to which there is a definite change in the performance of that protein (Walsh and Jefferis, 2006; Bhadauria *et al.*, 2007; Li *et al.*, 2010; Arif *et al.*, 2017a). PTMs are of more than 200 types on the basis of functional group attached or type of bond formation (Yu *et al.*, 2007). The proteins transitions after their modifications, work as regulatory factor for maintaining many physiological and cellular processes such as differentiation of cells, degradation, signaling and regulation of proteins, gene expression and interaction of proteins with other proteins (Minguez *et al.*, 2012).

Recently mass spectrophotometry technique has been designed for the identification, purification and quantification of modified sites on proteins. The accuracy of results of PTMs can be enhanced by combining some

techniques like, Affinity-based improvement by combining extraction methods and; multidimensional separation techniques with mass spectrophotometry, Which fascinate the researchers who are working in the field of proteomics (Jensen, 2004; Seo and Lee, 2004; Chandramouli and Qian, 2009).

The present study describes different types of PTMs and their positions on Cytochrome c oxidase-1 (*COX1*) protein of *Sordaria fimicola* (Six parental strains) and *Neurospora crassa*. The above mentioned protein is functionally important in respiratory chain reaction of Ascomycetes (Shoubridge, 2001), as it is responsible for accepting electrons from Cytochrome c and completing oxidation reaction (Hough *et al.*, 2014; Soto and Barrientos, 2016). In eukaryotes, *COX* protein is positioned in the inner mitochondrial membrane and has multiple subunits. Subunits 1 and 2 (*COX1*, *COX2*) are the largest constituent and essential to form a complex enzyme, Cytochrome c oxidase. *COX1* (Cytochrome c oxidase subunit-1) consists of heme a, heme a<sub>3</sub> and CuB for the reduction of oxygen (Diaz *et al.*, 2006; Shingu-VaZquez *et al.*, 2010). It is predicted through Hydropathy plots that *COX1* sequences have minimum 12  $\alpha$ -helical transmembrane regions which have redox centers of the enzyme (Fontanesi *et al.*, 2008;

Khalimonchuk et al., 2010; Soto and Barrientos, 2016).

To the best of our understanding, no study on post translational modifications of *COX1* in *S. fimicola* has been reported so far. Therefore, current study was conducted to determine some covalent modifications of *COX1* protein such as phosphorylation, glycosylation, acetylation and methylation in *S. fimicola* and *N. crassa*. Regulation of *COX1* activity by the post-translational modifications could characterize a new field of interest regarding the cell signal process.

## Materials and Methods

### Culture Retrieving of *S. fimicola*

Six pure culture strains (collected by Prof. Nevo from Evolutionary Canyon) of *S. fimicola* were made available by Molecular Genetics Research Laboratory of Botany Department, Punjab University Lahore. Three of these strains (S1, S2 and S3) were collected from South Facing Slope (SFS) at 60 m, 90 m and 120 m height above sea level respectively and remaining three strains (N<sub>5</sub>, N<sub>6</sub> and N<sub>7</sub>) were collected from North Facing Slope (NFS) at the same heights. These six parental strains cultures were retrieved on PDA media under aseptic conditions and incubated for 12–14 days at 18°C for subsequent analyses.

### DNA Extraction

DNA of six strains of *S. fimicola* was extracted by modified Pietro's method of DNA extraction (Pietro et al., 1995). For more detail please see Rana et al. (2018).

### Cytochrome c Oxidase-1 PCR Amplification and Sequencing

Forward and reverse primers for the amplification of *COX1* gene were designed by using Primer3 software. Touchdown PCR cycling conditions (Korbie and Mattick, 2008) were used for *COX1* amplification. PCR reaction mixture of 15 µL (2 µL DNA sample, 10 µL Master Mix (GeneAll), 1 µL Forward primer, 1 µL Reverse primer and 1 µL Double distilled water) was used. For the verification of amplification of desired PCR product 1% agarose gel electrophoresis was done with 1 Kb DNA ladder (Invitrogen) and PCR amplicons were sequenced and analyzed by sequence analysis software (Chromas). Nucleotide sequence of *COX1* of *N. crassa* was obtained from NCBI under the accession number KY\_498478.1.

### Prediction of PTMs on *COX1* Protein

For the predictions of glycosylation, phosphorylation, acetylation and methylation on *COX1* following online servers:

YinOYang	1.2
(www.cbs.dtu.dk/services/YinOYang/),	Netphos
	3.1

(www.cbs.dtu.dk/services/NetPhos/), PAIL (bdmpail.biocuckoo.org/) and ModPred (montana.informatics.indiana.edu/ModPred/faq.html) were used respectively. All of these servers provided detail of PTMs positions on *COX1*.

## Results

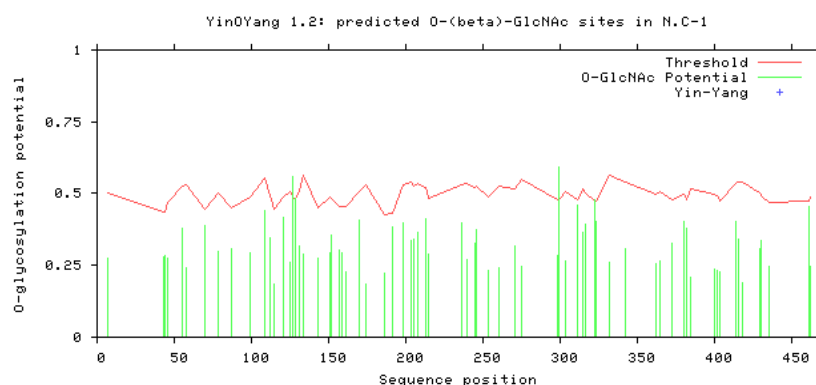
All the nucleotide sequences of *COX1* gene in reference specie and all strains of *Sordaria fimicola* were aligned by clustal omega to calculate polymorphic sites. This alignment showed polymorphisms at 10 different positions (Table 1 and 2). Glycosylation predicted sites by YinOYang 1.2 server (Table 3) showed glycosylation at positions 127, 128 on threonine residue and at positions 299, 322 on serine residue in *N. crassa*. Further, 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> have same positions of glycosylation i.e., 126 and 127 on threonine residue but no Ser modifications were observed in any strain. One conserved position of glycosylation (127T) was found in *N. crassa* and *S. fimicola*. Graphical representation of glycosylation in *COX1* protein of *N. crassa*, and parental strains of *S. fimicola* are shown in Fig. 1a and b.

NetPhos 3.1 server results (Table 3) showed that there was phosphorylation on *COX1* 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 that are Ser (Serine), T (Threonine) and Y (Tyrosine), respectively. In total, 59 residues of *N. crassa* and 57 residues of *S. fimicola* exhibited phosphorylation while only 8 phosphorylation sites were found to be conserved in all strains of *S. fimicola* and *N. crassa*. Furthermore, all threonine residues were found to be conserved among all strains of *S. fimicola*, however, variations were observed in case of serine and tyrosine residues i.e., 370 Ser and 318 Y in S<sub>1</sub> and S<sub>2</sub> strains while 313 Ser and 315 Y in S<sub>3</sub>, N<sub>5</sub>, N<sub>6</sub> and N<sub>7</sub> strains, respectively (Fig. 2).

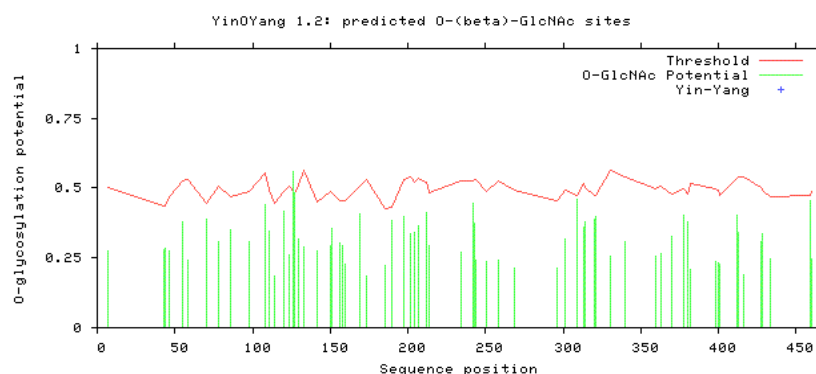
Prediction of Acetylation on Internal Lysines (PAIL) server is shown in Table 4. Acetylation sites differ between *N. crassa* and S<sub>1</sub>, S<sub>2</sub>, S<sub>3</sub>, N<sub>5</sub>, N<sub>6</sub> and N<sub>7</sub> at 3 positions i.e., 297 K, 310 K and 439 K in *N. crassa* while 308 K, 437 K and 438 K in all strains of *S. fimicola*. One acetylation at lysine residue 440 present in *N. crassa* was found to be absent in *S. fimicola*. All others positions (1, 16, 19, 64 and 68) were conserved positions of acetylation on internal lysine residue in *N. crassa* and *S. fimicola*.

In addition, methylation (Table 4) was conserved; at position 1 on lysine residue methylation was conserved in *N. crassa* and *S. fimicola*. *N. crassa* has 7 methylation sites whereas S<sub>1</sub>, S<sub>2</sub>, S<sub>3</sub>, have methylation on arginine residue at position 112 R only and on lysine residues at positions 1 K, 186 K, 308 K, 406 K. N<sub>5</sub>, N<sub>6</sub> and N<sub>7</sub> have methylation on arginine residues at positions 112 R, 347 R and on lysine residues at positions 1 K, 186 K, 308 K, 406 K. All the positions of methylation are same in all strains of *S. fimicola* except at position 347 on arginine residue which is only observed in N<sub>5</sub>, N<sub>6</sub> and N<sub>7</sub> strains of *S. fimicola*.

**a**



**b**



**Fig. 1:** Graphical Representation of Glycosylation Potential in *COXI* Protein (a) In *N. crassa* (b) In 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>. Red Zigzag Line Shows the Threshold (0.5), Green Lines Shows the O-Glycosylation Potential. Green Lines Exceeding Threshold Shows the Positions of Glycosylation

## Discussion

For broad scale assessment of genetic diversity, nucleotide variation is used (Wright, 2005). In Clark *et al.* (2007) examined the pattern of nucleotide sequence variation in *Arabidopsis thaliana* and reported that essential single nucleotide polymorphisms (SNPs) were more than 1 million and 4% of the genome sequence was found to be divergent from reference genome sequence. *COXI* gene Clustal analysis also revealed 10 variations (Table 1) in nucleotide sequence out of which 7 variations were within specie and these variations were also found in all the SFS (South Facing Slope) strains but not in any strain of the N-slope.

Hebert *et al.* (2003) also had worked on polymorphism of *COXI* gene for the discrimination of closely allied species because *COXI* have high rate of sequence change in most animal groups. Mogadam and Yousefi (2013) inspected the *COXI* gene polymorphism in Malaysia population. They found 13 nucleotide variations within 1548 bp. Nucleotide diversity in *COXI* of *Cooperia oncophora* was not more than 2% between species (Vaulin

and Novikov, 2012). *COXI* sequence alignment results have shown 10 variations of protein residues along with conservation of residues in similar groups at five positions and conservation of residues in weakly similar groups at 2 positions. All these molecular analysis results showed the genetic diversity between different strains of *S. fimicola*.

Baer and Millar (2016) reported that DNA has crucial role in evolution due to genetic variability in it and genetic diversity of DNA directly affects the expression of proteins (which form from translation of its genes). More over post translational modifications influence significant variations in mature proteins of same gene leading to evolutionary changes and adaptations as PTMs are actually responsible for the complexity of the proteome (Audagnotto and Peraro, 2017). Tahir *et al.* (2016) found genetic diversity in Cytochrome c oxidase I among the 44 different species of Anopheles mosquitoes.

In eukaryotes, protein glycosylation is vital post translational modification that is of generally three types *i.e.*, N-linked, O-linked and glycosyl-phosphatidylinositol-anchored. During this research work, we have found O-

**Table 1:** Pairwise sequence alignment of *COXI* gene to determine genetic diversity between different *S. fimicola* strains and *N. crassa*

N.c	CATTGCAGATAATCAATTATAACAATGCTATAATAACTGCACATGCGATCTTAATGATTTT	300
S <sub>1</sub>	CATTGCAGATAATCAATTATAACAATGCTATAATAACTGCACATGCGATCTTAATGATATT	300
S <sub>2</sub>	CATTGCAGATAATCAATTATAACAATGCTATAATAACTGCACATGCGATCTTAATGATATT	300
S <sub>3</sub>	CATTGCAGATAATCAATTATAACAATGCTATAATAACTGCACATGCGATCTTAATGATATT	300
N <sub>5</sub>	CATTGCAGATAATCAATTATAACAATGCTATAATAACTGCACATGCGATCTTAATGATATT	300
N <sub>6</sub>	CATTGCAGATAATCAATTATAACAATGCTATAATAACTGCACATGCGATCTTAATGATATT	300
N <sub>7</sub>	CATTGCAGATAATCAATTATAACAATGCTATAATAACTGCACATGCGATCTTAATGATATT	300
	*****	
N.c	TATGCTATTAACAGATAGAAATTTAATACATCATTCTTTGAAACAGCTGGTGGTGGTGA	780
S <sub>1</sub>	TATGCTATTAACAGATAGAAATTTAATACATCATTCTTTGAAACAGCTGGTGGTGGTGA	780
S <sub>2</sub>	TATGCTATTAACAGATAGAAATTTAATACATCATTCTTTGAAACAGCTGGTGGTGGTGA	780
S <sub>3</sub>	TATGCTATTAACAGATAGAAATTTAATACATCATTCTTTGAAACAGCTGGTGGTGGTGA	780
N <sub>5</sub>	TATGCTATTAACAGATAGAAATTTAATACATCATTCTTTGAAACAGCTGGTGGTGGTGA	780
N <sub>6</sub>	TATGCTATTAACAGATAGAAATTTAATACATCATTCTTTGAAACAGCTGGTGGTGGTGA	780
N <sub>7</sub>	TATGCTATTAACAGATAGAAATTTAATACATCATTCTTTGAAACAGCTGGTGGTGGTGA	780
	*****	
N.c	TCCTATTTTATTCCAACATCTTTTCTGATTCTTCGGGCATCCTGAGGTTTACATTTTAAT	840
S <sub>1</sub>	CCCTATTTTATTCCAACATCTTTTCTGATTCTTCGGGCATCCTGAGGTTTACATTTTAAT	840
S <sub>2</sub>	CCCTATTTTATTCCAACATCTTTTCTGATTCTTCGGGCATCCTGAGGTTTACATTTTAAT	840
S <sub>3</sub>	CCCTATTTTATTCCAACATCTTTTCTGATTCTTCGGGCATCCTGAGGTTTACATTTTAAT	840
N <sub>5</sub>	CCCTATTTTATTCCAACATCTTTTCTGATTCTTCGGGCATCCTGAGGTTTACATTTTAAT	840
N <sub>6</sub>	CCCTATTTTATTCCAACATCTTTTCTGATTCTTCGGGCATCCTGAGGTTTACATTTTAAT	840
N <sub>7</sub>	CCCTATTTTATTCCAACATCTTTTCTGATTCTTCGGGCATCCTGAGGTTTACATTTTAAT	840
	*****	
N.c	AAGTCATCAGTGTATACAGTTGGTTTAGACGTGGATACAAGAGCGTATTTACAGCAGC	1020
S <sub>1</sub>	AAGTCACCACATGTGTACAGTTGGTTTAGACGTGGATACAAGAGCGTATTTACAGCAGC	1020
S <sub>2</sub>	AAGTCACCACATGTGTACAGTTGGTTTAGACGTGGATACAAGAGCGTATTTACAGCAGC	1020
S <sub>3</sub>	AAGTCACCACATGTGTACAGTTGGTTTAGACGTGGATACAAGAGCGTATTTACAGCAGC	1020
N <sub>5</sub>	AAGTCATCAGTGTATACAGTTGGTTTAGACGTGGATACAAGAGCGTATTTACAGCAGC	1020
N <sub>6</sub>	AAGTCATCAGTGTATACAGTTGGTTTAGACGTGGATACAAGAGCGTATTTACAGCAGC	1020
N <sub>7</sub>	AAGTCATCAGTGTATACAGTTGGTTTAGACGTGGATACAAGAGCGTATTTACAGCAGC	1020
	*****	
N.c	TACATTAATTATTGCAGTTCCTACAGGAATTTAAATATTCTCATGATTAGCTACATGTTA	1080
S <sub>1</sub>	AACATTAATTATTGCAGTTCCTACTGGAATTTAAATATTCTCATGATTAGCTACATGTTA	1080
S <sub>2</sub>	AACATTAATTATTGCAGTTCCTACTGGAATTTAAATATTCTCATGATTAGCTACATGTTA	1080
S <sub>3</sub>	TACATTAATTATTGCAGTTCCTACTGGAATTTAAATATTCTCATGATTAGCTACATGTTA	1080
N <sub>5</sub>	TACATTAATTATTGCAGTTCCTACGGAATTTAAATATTCTCATGATTAGCTACATGTTA	1080
N <sub>6</sub>	TACATTAATTATTGCAGTTCCTACGGAATTTAAATATTCTCATGATTAGCTACATGTTA	1080
N <sub>7</sub>	TACATTAATTATTGCAGTTCCTACGGAATTTAAATATTCTCATGATTAGCTACATGTTA	1080
	*****	
N.c	TGGAGGTTCTATTAGATTAACCTCTTCTATGTTATTTGCTTTAGGTTTGTATTTATGTT	1140
S <sub>1</sub>	TGGAGGTTGTATTAGATTAACCTCTTCTATGTTATTTGCTTTAGGTTTGTATTTATGTT	1140
S <sub>2</sub>	TGGAGGTTGTATTAGATTAACCTCTTCTATGTTATTTGCTTTAGGTTTGTATTTATGTT	1140
S <sub>3</sub>	TGGAGGTTGTATTAGATTAACCTCTTCTATGTTATTTGCTTTAGGTTTGTATTTATGTT	1140
N <sub>5</sub>	TGGAGGTTCTATTAGATTAACCTCTTCTATGTTATTTGCTTTAGGTTTGTATTTATGTT	1140
N <sub>6</sub>	TGGAGGTTCTATTAGATTAACCTCTTCTATGTTATTTGCTTTAGGTTTGTATTTATGTT	1140
N <sub>7</sub>	TGGAGGTTCTATTAGATTAACCTCTTCTATGTTATTTGCTTTAGGTTTGTATTTATGTT	1140
	*****	

Highlighted region shows mutation on that point

linked glycosylation of *COXI* protein in *S. fimicola* and *N. crassa* because these *O*-linked modifications occurred only on serine or threonine residues and not on tyrosine residues and these modifications are also called mucin-type *O*-linked glycosylation.

In glycosylation, the hydroxyl groups of serine and threonine residues (*O*-glycosylation) get transformed (Bektasa et al., 2011; Jamil et al., 2018) but present results show glycosylation only on threonine residue in *COXI* of *S. fimicola*. *COXI* have two glycosylation modification positions on threonine (T) residues and two on serine (S) residues in *N. crassa* whereas all strains of *S. fimicola* have glycosylation only on two threonine (T) residues. In conclusion one position was found to be conserved in *N.*

*crassa* and *S. fimicola* (Table 3).

Phosphorylation modification of proteins have gained much consideration in regulatory process of living organisms in which protein kinases and their corresponding phosphatases have central role (Knorre et al., 2009) and most of the protein kinases studied along with protein phosphorylation belongs to eukaryotes (Trost and Kusalik, 2011). Present results have four major protein kinases involved in phosphorylation of Cytochrome c oxidase where PKC, unsp, PKA, cdc2 have their contribution in phosphorylation respectively. In Liko et al. (2016) reported 8 phosphorylated sites (only on serine and threonine residues) in bovine Cytochrome c oxidase with no phosphorylation on tyrosine residue while in contrast

**Table 2:** Pairwise amino acid sequence alignment of *COX1* protein of *S. fimicola* strains and *N. crassa*

N.c	TWNSIFCFNKGKRTWGA V HCR* <b>SI</b> IQCYNNCTCDLNDFLYGYASINRWIW*FLV <b>TI</b> IS	112
S <sub>1</sub>	TWNSIFCFNKGKRTWGA V HCR* <b>SI</b> IQ*YNNCTCDLNDFLYGYASINRWIW*FLV <b>TI</b> IS	111
S <sub>2</sub>	TWNSIFCFNKGKRTWGA V HCR* <b>SI</b> IQ*YNNCTCDLNDFLYGYASINRWIW*FLV <b>TI</b> IS	111
S <sub>3</sub>	TWNSIFCFNKGKRTWGA V HCR* <b>SI</b> IQ*YNNCTCDLNDFLYGYASINRWIW*FLV <b>TI</b> IS	111
N <sub>5</sub>	TWNSIFCFNKGKRTWGA V HCR* <b>SI</b> IQ*YNNCTCDLNDFLYGYASINRWIW*FLV <b>TI</b> IS	111
N <sub>6</sub>	TWNSIFCFNKGKRTWGA V HCR* <b>SI</b> IQ*YNNCTCDLNDFLYGYASINRWIW*FLV <b>TI</b> IS	111
N <sub>7</sub>	TWNSIFCFNKGKRTWGA V HCR* <b>SI</b> IQ*YNNCTCDLNDFLYGYASINRWIW*FLV <b>TI</b> IS	111
*****		
N.c	YAI <b>NR</b> * <b>KF</b> *YIIL*NSWWW* <b>S</b> YFIPT <b>S</b> FLILRAS*GLHFNYTWFWY <b>N</b> KYNNISLF** <b>SI</b>	276
S <sub>1</sub>	YGEN** <b>KF</b> *YIIL*NSWWW* <b>P</b> YFIST <b>S</b> FLILRAS*GLHFNYTWFWY <b>N</b> KYNNISLF** <b>ER</b> I	274
S <sub>2</sub>	YGEN** <b>KF</b> *YIIL*NSWWW* <b>P</b> YFIST <b>S</b> FLILRAS*GLHFNYTWFWY <b>N</b> KYNNISLF** <b>ER</b> I	274
S <sub>3</sub>	YGEN** <b>KF</b> *YIIL*NSWWW* <b>P</b> YFIST <b>S</b> FLILRAS*GLHFNYTWFWY <b>N</b> KYNNISLF** <b>ER</b> I	274
N <sub>5</sub>	YGEN** <b>KF</b> *YIIL*NSWWW* <b>P</b> YFIST <b>S</b> FLILRAS*GLHFNYTWFWY <b>N</b> KYNNISLF** <b>ER</b> I	274
N <sub>6</sub>	YGEN** <b>KF</b> *YIIL*NSWWW* <b>P</b> YFIST <b>S</b> FLILRAS*GLHFNYTWFWY <b>N</b> KYNNISLF** <b>ER</b> I	274
N <sub>7</sub>	YGEN** <b>KF</b> *YIIL*NSWWW* <b>P</b> YFIST <b>S</b> FLILRAS*GLHFNYTWFWY <b>N</b> KYNNISLF** <b>ER</b> I	274
*.* *****		
N.c	RLYWYGLRHDVYWNIRIYCLKSP <b>H</b> VYSW <b>RR</b> GYKSVFHSSYIN <b>Y</b> CSSY <b>R</b> N*NILMISYML	335
S <sub>1</sub>	RLYWYGLRHDVYWNIRIYCLKSP <b>H</b> VYSW <b>RR</b> GYKSVFHSSNIN <b>Y</b> CSSY <b>W</b> N*NILMISYML	333
S <sub>2</sub>	RLYWYGLRHDVYWNIRIYCLKSP <b>H</b> VYSW <b>RR</b> GYKSVFHSSNIN <b>Y</b> CSSY <b>W</b> N*NILMISYML	333
S <sub>3</sub>	RLYWYGLRHDVYWNIRIYCLKSP <b>H</b> VYSW <b>RR</b> GYKSVFHSSYIN <b>Y</b> CSSY <b>W</b> N*NILMISYML	333
N <sub>5</sub>	RLYWYGLRHDVYWNIRIYCLKSP <b>H</b> VYSW <b>RR</b> GYKSVFHSSYIN <b>Y</b> CSSY <b>W</b> N*NILMISYML	333
N <sub>6</sub>	RLYWYGLRHDVYWNIRIYCLKSP <b>H</b> VYSW <b>RR</b> GYKSVFHSSYIN <b>Y</b> CSSY <b>W</b> N*NILMISYML	333
N <sub>7</sub>	RLYWYGLRHDVYWNIRIYCLKSP <b>H</b> VYSW <b>RR</b> GYKSVFHSSYIN <b>Y</b> CSSY <b>W</b> N*NILMISYML	333
*****		
N.c	WR <b>F</b> Y*INSFYVICFR <b>E</b> CIYVHNWGIK <b>W</b> SCFSECFRYSI <b>R</b> YLLRSCSFSLC <b>I</b> KYGCCIC	394
S <sub>1</sub>	WR <b>L</b> Y*INSFYVICFR <b>E</b> CIYVHNWGIK <b>W</b> SCFSECFRYSI <b>R</b> YLLRSCSFSLC <b>I</b> KYGCCIC	392
S <sub>2</sub>	WR <b>L</b> Y*INSFYVICFR <b>E</b> CIYVHNWGIK <b>W</b> SCFSECFRYSI <b>R</b> YLLRSCSFSLC <b>I</b> KYGCCIC	392
S <sub>3</sub>	WR <b>L</b> Y*INSFYVICFR <b>E</b> CIYVHNWGIK <b>W</b> SCFSECFRYSI <b>R</b> YLLRSCSFSLC <b>I</b> KYGCCIC	392
N <sub>5</sub>	WR <b>F</b> Y*INSFYVICFR <b>E</b> CIYVHNWGIK <b>W</b> SCFSECFRYSI <b>R</b> YLLRSCSFSLC <b>I</b> KYGCCIC	392
N <sub>6</sub>	WR <b>F</b> Y*INSFYVICFR <b>E</b> CIYVHNWGIK <b>W</b> SCFSECFRYSI <b>R</b> YLLRSCSFSLC <b>I</b> KYGCCIC	392
N <sub>7</sub>	WR <b>F</b> Y*INSFYVICFR <b>E</b> CIYVHNWGIK <b>W</b> SCFSECFRYSI <b>R</b> YLLRSCSFSLC <b>I</b> KYGCCIC	392
**.* *****		

Highlighted region shows mutation on that point

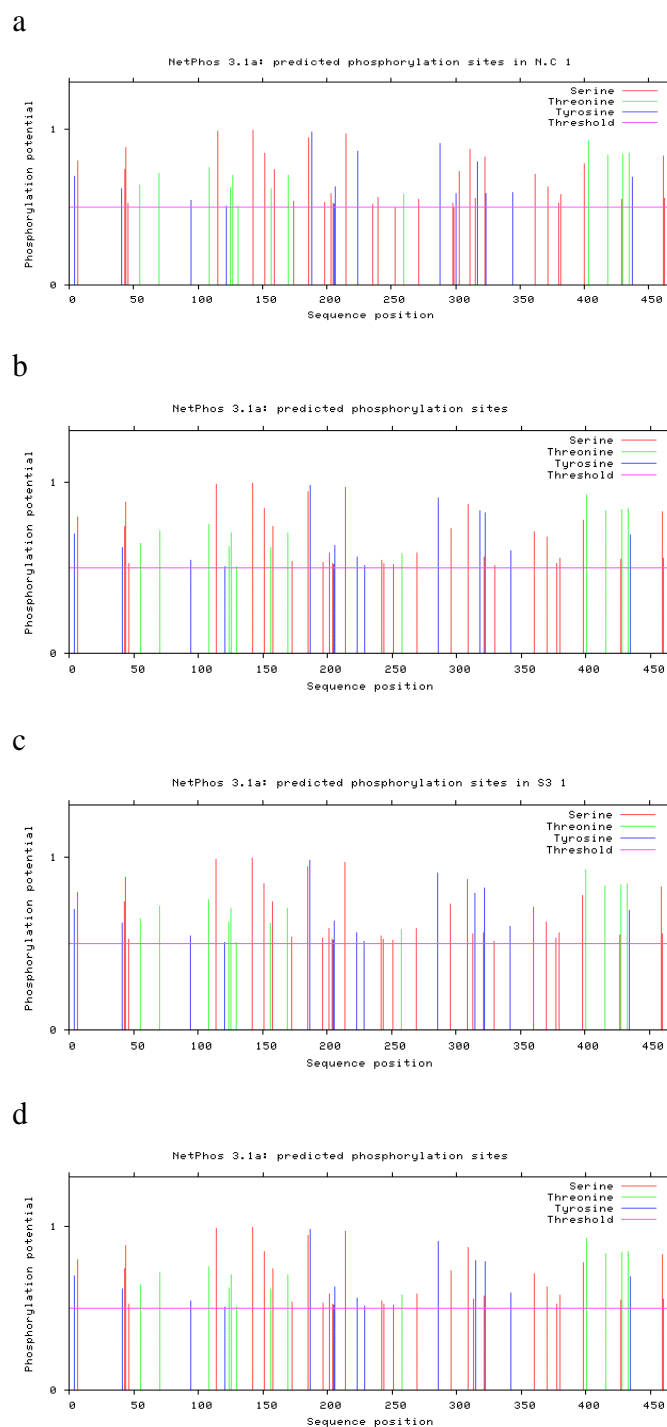
**Table 3:** Prediction of glycosylation and phosphorylation on serine, threonine and tyrosine residues in *N. crassa* and in different strains of *S. fimicola*

Organism	Residue	Glycosylation	COX1			
			Phosphorylation	unsp	PKA	PKC
<i>N. crassa</i>	Serine	299, 322	7, 43, 44, 46, 115, 143, 152, 159, 174, 186, 198, 203, 205, 215, 236, 240, 253, 271, 298, 299, 303, 311, 315, 323, 362, 372, 380, 382, 400, 429, 461, 462	43, 115, 143, 159, 186, 215, 311, 323	43, 44, 115, 143, 203, 240, 253, 271, 315, 362, 372, 382, 400, 429, 461, 462	7, 115, 152, 159, 186, 23, 205, 298, 303, 311, 380, 382, 400, 429, 461, 462
	Threonine	127, 128	55, 70, 109, 125, 127, 131, 157, 170, 260, 403, 418, 430, 435	157	Nil	55, 70, 109, 125, 127, 131, 157, 170, 260, 403, 418, 430, 435
	Tyrosine	Nil	4, 41, 95, 122, 188, 206, 207, 224, 288, 300, 317, 324, 344, 437	4, 41, 188, 207, 224, 288, 300, 317, 324, 437	Nil	Nil
<b>S<sub>1</sub>, S<sub>2</sub>, S<sub>3</sub></b> <b>Strains</b>	Serine	Nil	7, 43, 44, 46, 114, 142, 151, 158, 173, 185, 197, 202, 204, 214, 242, 244, 251, 269, 296, 309, 321, 330, 360, 370, 378, 380, 398, 427, 459, 460	43, 114, 142, 158, 185, 214, 296, 309	43, 44, 114, 142, 202, 242, 251, 269, 330, 360, 370, 380, 398, 427, 459	7, 114, 151, 158, 185, 309, 321, 460
	Threonine	126, 127	55, 70, 108, 124, 126, 130, 156, 169, 258, 401, 416, 428, 433	156	Nil	55, 70, 108, 124, 126, 130, 156, 169, 258, 401, 416, 428, 433
	Tyrosine	-	4, 41, 94, 121, 187, 205, 206, 223, 229, 286, 318, 322, 342, 435	4, 41, 206, 223, 229, 286, 318, 322, 435	187	-
<b>N<sub>5</sub>, N<sub>6</sub>, N<sub>7</sub></b>	Serine	-	7, 43, 44, 46, 114, 142, 151, 158, 173, 185, 197, 202, 204, 214, 242, 244, 251, 269, 296, 309, 313, 321, 330, 360, 378, 380, 398, 427, 459, 460	43, 114, 142, 158, 185, 214, 296, 309	43, 44, 114, 142, 202, 242, 251, 269, 313, 330, 360, 370, 380, 398, 427, 459	7, 114, 151, 158, 185, 309, 321, 460
	Threonine	126, 127	55, 70, 108, 124, 126, 130, 156, 169, 258, 401, 416, 428, 433	156	-	55, 70, 108, 124, 126, 130, 156, 169, 258, 401, 416, 428, 433
	Tyrosine	-	4, 41, 94, 121, 187, 205, 206, 223, 229, 286, 315, 322, 342, 435	4, 41, 206, 223, 229, 286, 315, 322, 435	187	-

Highlighted area = conserved regions

present research results showed that tyrosine residue of *COX1* have phosphorylation modification in both *N. crassa* and *S. fimicola*. Trost and Kusalik (2011) confirmed that

computational prediction of potential sites of phosphorylation is more significant than time taking experimental procedures. Liko *et al.* (2016) have



**Fig. 2:** Graphical Representation of Phosphorylation Potential in *COX1* Protein (a) In *N. crassa* (b) In  $S_1$  and  $S_2$  (c) In  $S_3$  (d) In  $N_5$ ,  $N_6$  and  $N_7$ . 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 *COX1* and Red, Green, Blue Color of Lines Shows the Residue Present at that Position

reported 14 acetylation sites in *COX* of bovine. Rieder and Bosshard (1980) compared attachment sites on Cytochrome c for *COX* and *CyC1* where they also found 6 positions of acetylated lysine residues. Vassilev *et al.* (1995) had also reported acetylation at 324 on internal

lysine of Cytochrome c oxidase subunit 1 for formation of myristic acid. These computational analysis results are quite satisfactory due their improved prediction accuracy and these predictions of methylation on covalently modified lysine and arginine residues are important information for

**Table 4:** Comparison of *COX1* Acetylation Sites, Methylation Sites on Lysine (K) and Arginine (R) Residues in *N. crassa* and *S. fimicola*

		COX1	
Organism	Residue	Acetylation Sites	Methylation Sites
<i>N. crassa</i>	K	1, 16, 19, 64, 68, 297, 310, 439, 440 Total = 9sites	1, 187, 229, 310, 408 Total = 5sites
	R	-	113, 349 Total = 2sites
S <sub>1</sub> , S <sub>2</sub> , S <sub>3</sub>	K	1, 16, 19, 64, 68, 308, 437, 438 Total = 8sites	1, 186, 308, 406 Total = 4sites
	R	-	112 Total = 1site
N <sub>5</sub> , N <sub>6</sub> , N <sub>7</sub>	K	1, 16, 19, 64, 68, 308, 437, 438 Total = 8sites	1, 186, 308, 406 Total = 4sites
	R	-	112, 347 Total = 2sites

future experimental research (Deng *et al.*, 2017). Various workers have applied the same protein modification predictor tools to determine the post translational modifications of mating type and frequency clock proteins (Arif *et al.*, 2017b); Histone 3 and 4 proteins (Jamil *et al.*, 2018); Manganese oxide super dismutase (MnSOD) protein (Rana *et al.*, 2018) in different strains of *S. fimicola* as used in present study.

Hoffmann and Hercus (2000) also mentioned environmental stress as an evolutionary force which starts with variation in genetic makeup. It is concluded from above results that genetic variations help to overcome the harsh conditions confronted by an organism (Arif *et al.*, 2017a). These genetic variations within same species due to any force (point mutation) initially takes place in the nucleotide sequence of genes leading to translation of these genes with diverse amino acid sequence of the same protein in same species. Finally, the protein modifications are also effected by genetic variations in basic unit of genome due to which dynamics and functions of protein are enhanced.

## References

- Arif, R., F. Akram, T. Jamil, H. Mukhtar and M. Saleem, 2017a. Genetic variation and its reflection on post translational modifications in frequency clock and mating type a-1 proteins in *Sordaria fimicola*. *Biol. Med. Res. Intl.*, 2017: 1–10
- Arif, R., S.F. Lee and M. Saleem, 2017b. Evaluating short sequence markers in *S. fimicola* by high resolution melt analysis. *Intl. J. Agric. Biol.*, 19: 248–254
- Audagnotto, M. and M.D. Peraro, 2017. Protein post translational modifications: *In silico* prediction tools and molecular modeling. *Comput. Struct. Biotechnol. J.*, 15: 307–319
- Baer, B. and H. Millar, 2016. Proteomics in evolutionary ecology. *J. Proteom.*, 135: 4–11
- Bektasa, M., S. David and B. Rubenstein, 2011. The role of intracellular protein o-glycosylation in cell adhesion and disease. *J. Biomed. Res.*, 25: 227–236
- Bhadauria, V., W. Zhao, L. Wang, Y. Zhang, J. Liu, J. Yang, L. Kong and Y. Peng, 2007. Advances in fungal proteomics. *Microbiol. Res.*, 162: 193–200
- Chandramouli, K. and P. Qian, 2009. Proteomics: challenges, techniques and possibilities to overcome biological sample complexity. *Hum. Genom. Proteom.*, 20: 1–22
- Clark, R.M., G. Schweikert, C. Toomajian, S. Ossowski, G. Zeller, P. Shinn, N. Warthmann, T.T. Hu, G. Fu, D.A. Hinds, H. Chen, K.A. Frazer, D.H. Huson, B. Scholkopf, M. Nordborg, G. Ratsch, J.R. Ecker and D. Weigel, 2007. Common sequence polymorphisms shaping genetic diversity in *Arabidopsis thaliana*. *Science*, 317: 338–342
- 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: 647–658
- Diaz, F., H. Fukui, S. Garcia and C.T. Morae, 2006. Cytochrome c oxidase is required for the assembly/stability of respiratory complex 1 in mouse fibroblasts. *Mol. Cell. Biol.*, 26: 4872–4881
- Fontanesi, F., I.C. Soto and A. Barrientos, 2008. Cytochrome c oxidase biogenesis: new levels of regulation. *Intl. Union Biochem. Mol. Biol. Life*, 60: 557–568
- Hebert, P.D.N., S. Ratnasingham and J.R.D. Waard, 2003. Barcoding animal life: cytochrome c oxidase subunit I divergences among closely related species. *Proc. Roy. Soc. Lond. B Biol. Sci.*, 270: 96–99
- Hoffmann, A.A. and M.J. Hercus, 2000. Environmental stress as an evolutionary force. *Biol. Sci.*, 50: 217–226
- Hough, M.A., G. Silkstone, J.A. Worrall and M.T. Wilson, 2014. No binding to the proapoptotic cytochrome c-cardiolipin complex. *Vitam. Horm.*, 96: 193–209
- Jamil, T., N. Sami, R. Arif, Q. Rashid and 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
- Jensen, O.N., 2004. Modification-specific proteomics: characterization of post-translational modifications by mass spectrometry. *Curr. Opin. Chem. Biol.*, 8: 33–41
- Khalimonchuk, O., M. Bestwick, B. Meunier, T.C. Watts and D.R. Winge, 2010. Formation of the redox cofactor centers during *COX1* maturation in yeast cytochrome oxidase. *Mol. Cell. Biol.*, 30: 1004–1017
- Knorre, D.G., N.V. Kudryashova and T.S. Godovikova, 2009. Chemical and functional aspects of post translational modification of proteins. *Acta Nat.*, 1: 29–51
- Korbie, D.J. and J.S. Mattick, 2008. Touchdown PCR for increased specificity and sensitivity in pcr amplification. *Nat. Protoc.*, 3: 1452–1456
- Li, S., L.M. Iakoucheva, S.D. Mooney and P. Radivojac, 2010. Loss of post-translational modification sites in disease. *Pac. Symp. Biol. Comput.*, 15: 337–347
- Liko, I., M.T. Degiacomia, S. Mohammeda, S. Yoshikawa, C. Schmidta and C.V. Robinsona, 2016. Dimer interface of bovine cytochrome c oxidase is influenced by local posttranslational modifications and lipid binding. *Proc. Natl. Acad. Sci.*, 113: 8230–8235
- Minguez, P., L. Parca, F. Diella, D.R. Mende, R. Kumar, M. Helmer-Citterich, A.C. Gavin, V.V. Noort and P. Bork, 2012. Deciphering a global network of functionally associated post-translational modifications. *Mol. Syst. Biol.*, 8: 599

- Mogadam, R.G. and P. Yousefi, 2013. Genetic polymorphisms in cytochrome c oxidase subunit 1 of the Malaysian population. *Ann. Biol. Res.*, 4: 56–60
- Pietro, S., T.M. Fulton, J. Chunwonges and S.D. Tanksley, 1995. Extraction of high quality DNA for genome sequencing. *Mol. Biol. Rep.*, 13: 207
- Rana, J., R. Arif, M. Ishfaq, S.F. Lee and M. Saleem, 2018. Amplification and posttranslational modification study of manganese superoxide dismutase gene in *Sordaria fimicola*. *Intl. J. Agric. Biol.*, 20: 2059–2063
- Rieder, R. and H.R. Bosshard, 1980. Comparison of the binding sites on cytochrome C for cytochrome C oxidase, cytochrome bc1, and cytochrome c1. Differential acetylation of lysyl residues in free and complexed cytochrome C. *J. Biol. Chem.*, 255: 4732–4739
- Seo, J. and K.J. Lee, 2004. Post translational modifications and their biological functions: proteomic analysis and systematic analysis approaches. *J. Biochem. Mol. Biol.*, 37: 35–44
- Shingu-VaZquez, M., Y. Camacho-Villasana, L. Sandoval-Romero, C.A. Butler, T.D. Fox and X. Pe'Rez-Martínez, 2010. The carboxyl-terminal end of *COX1* is required for feedback assembly regulation of *COX1* synthesis in *Saccharomyces cerevisiae* mitochondria. *J. Biol. Chem.*, 285: 34382–34389
- Shoubridge, E.A., 2001. Cytochrome c oxidase deficiency. *Amer. J. Med. Genet.*, 106: 46–52
- Soto, I.C. and A. Barrientos, 2016. Mitochondrial cytochrome c oxidase biogenesis is regulated by the redox state of a heme-binding translational activator. *Antioxid. Redox Signal.*, 24: 281–298
- Tahir, H.M., Mehwish, N. Kanwal, A. Butt, S.Y. Khan and A. Yaqub, 2016. Genetic diversity in cytochrome c oxidase 1 gene of anopheles mosquitoes. *Mitochondr. DNA A*, 27: 4298–4301
- Trost, B. and A. Kusalik, 2011. Computational prediction of eukaryotic phosphorylation sites. *Bioinformatics*, 27: 2927–2935
- Vassilev, A.O., N. Plesofsky-Vig and R. Brambl, 1995. Cytochrome c oxidase in *Neurospora crassa* contains myristic acid covalently linked to subunit 1. *Proc. Natl. Acad. Sci.*, 92: 8680–8684
- Vaulin, O.V. and Y.M. Novikov, 2012. Polymorphism and interspecific variability of cytochrome oxidase subunit I (COI) gene nucleotide sequence in sibling species of A and B *Anopheles messeae* and *An. beklemishevi* (Diptera: Culicidae). *Appl. Res.*, 2: 421–429
- Walsh, G. and R. Jefferis, 2006. Post-translational modifications in the context of therapeutic proteins. *Nat. Biotechnol.*, 24: 1241–1252
- Wright, A.F., 2005. Genetic variation: polymorphisms and mutations. *eLS*, 1: 1–10
- Yu, L.R., H.J. Issaq and T.D. Veenstra, 2007. Phosphoproteomics for the discovery of kinases as cancer biomarkers and drug targets. *Proteomics*, 1: 1042–1057

(Received 01 August 2018; Accepted 06 November 2018)