

OPEN

# Differential protein profiling of soil diazotroph *Rhodococcus qingshengii* S10107 towards low-temperature and nitrogen deficiency

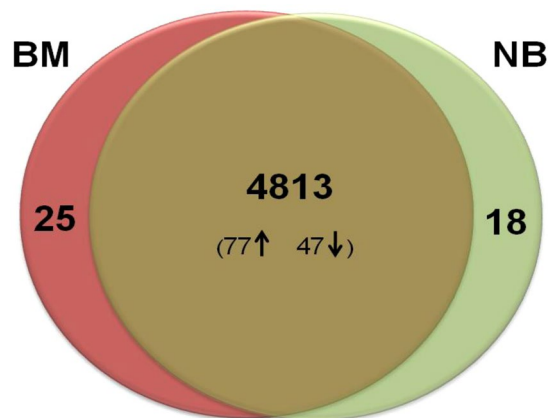
Deep Chandra Suyal<sup>1,2</sup>, Divya Joshi<sup>2</sup>, Saurabh Kumar<sup>2</sup>, Ravindra Soni<sup>3</sup> & Reeta Goel<sup>2\*</sup>

Protein-based biomarkers can be a promising approach for identification and real-time monitoring of the bio-inoculants employed under sustainable agricultural plans. In this perspective, differential proteomics of psychrophilic diazotroph *Rhodococcus qingshengii* S10107 (JX173283) was performed to unravel its adaptive responses towards low-temperature nitrogen deficiency and identification of a biomarker for respective physiological conditions. LC-MS/MS-based proteome analysis mapped more than 4830 proteins including 77 up-regulated and 47 down-regulated proteins ( $p \leq 0.05$ ). Differential expression of the structural genes of *nif* regulon viz. *nifH*, *nifD*, and *nifK* along with their response regulators i.e. *nifA*, *nifL*, and *nifB* indicated that the nitrogenase complex was activated successfully. Besides up-regulating the biosynthesis of certain amino acids viz. Leucine, Lysine, and Alanine; the expression of the peptidoglycan synthesis proteins were also increased; while, the enzymes involved in Lipid biosynthesis were found to decrease. Furthermore, two important enzymes of the pentose phosphate pathway viz. *Transketolase* and *Transaldolase* along with *Ribose import ATP-binding protein RbsA* were also found to induce significantly under low temperature a nitrogen deficient condition, which suggests the cellular need for ample ribose sugar instantly. Additionally, comparative protein profiling of S10107 strain with our previous studies revealed that *CowN* protein was significantly up-regulated in all the cases under low-temperature nitrogen deficient conditions and therefore, can be developed as a biomarker. Conclusively, present study for the first time provides an in-depth proteome profiling of *R. qingshengii* S10107 and proclaims *CowN* as a potential protein biomarker for monitoring BNF under cold niches.

Nitrogen is an essential nutrient for crop growth and development. In spite of having a largest atmospheric reservoir, its soil availability is very low and thus, mostly requires external inputs. Biofertilizers are among the effective and eco-friendly alternative available in the respective sector. Global Industry Report revealed that their market size was USD 787.8 million in 2016 including diazotrophs as the largest segment with 75.0% of global revenue share<sup>1</sup>.

Recent trends indicating that microbial biofertilizers are unable to confer expected benefits to the farmers. However, they should not be blamed at first glance, because, it might be the result of their adaptive failure towards natural field conditions as well as local edaphic factors. Instead, efforts should be made to develop such technologies which can monitor the real-time performance of the bioinoculant(s) so that they can be replaced on time and loss of the farmer's can be minimized. Under *in situ* conditions, biological nitrogen fixation (BNF) process can be monitored by using acetylene reduction assay (ARA)<sup>2</sup>, <sup>15</sup>N<sub>2</sub> assimilation technique and gene probes<sup>2,3</sup>. Besides having their own limitations, these techniques require sophisticated instrumentations and skilled personals which restrict their implementation. An alternative can be the microbial proteins for their identification and tracking functional interactions within a niche<sup>4,5</sup>. Such protein-based biomarker technologies are already being commercialized for clinical purpose<sup>6</sup>, bioremediation<sup>7</sup> and heavy metal detection<sup>8</sup>; but, in case of agriculture, they are still in infancy.

<sup>1</sup>Department of Microbiology, Akal College of Basic Sciences, Eternal University, Baru Sahib, Sirmaur, 173101, Himachal Pradesh, India. <sup>2</sup>Department of Microbiology, College of Basic Sciences and Humanities, G.B.Pant University of Agriculture and Technology, Pantnagar, 263145, Uttarakhand, India. <sup>3</sup>Department of Agricultural Microbiology, College of Agriculture, Indira Gandhi Krishi Viswavidyalaya, Raipur, C.G., India. \*email: [rg55@rediffmail.com](mailto:rg55@rediffmail.com)



**Figure 1.** Venn diagram depicting the total number of expressed proteins of *Rhodococcus qingshengii* S10107 strain under Nitrogen deficient (BM) and Nitrogen sufficient (NB) conditions. Cells were harvested at mid-logarithmic phase. A total of 4813 proteins are shared by both BM and NB, whereas the BM (25) showed more exclusive proteins than the NB (18). The analysis was performed in triplicates. The diagram shows the average value of proteins obtained from three replicates of the experiment.

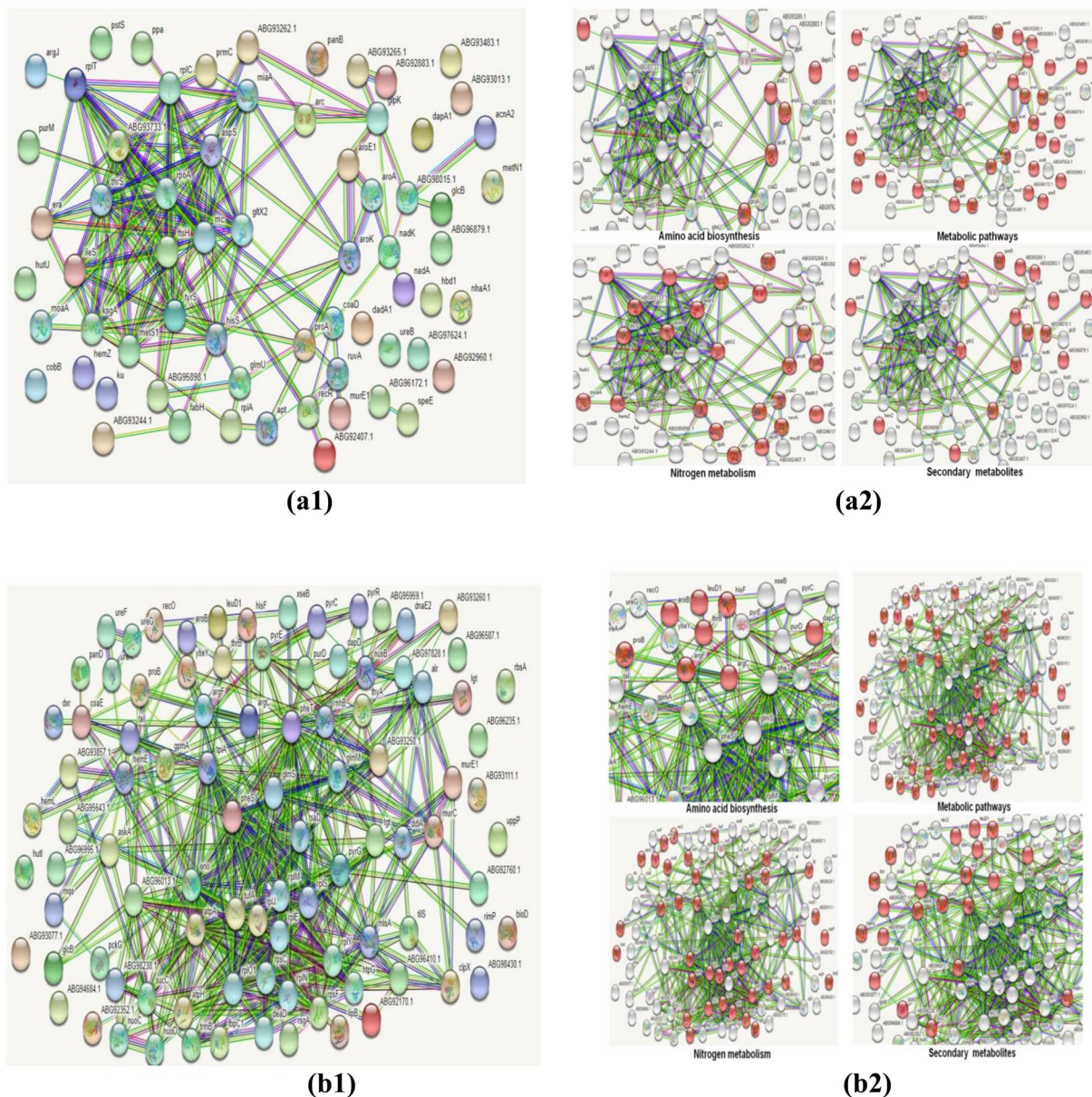
A series of proteomic investigations have been performed by the author group on different diazotrophs viz. *Dyadobacter psychrophilus* B2 and *Pseudomonas jessenii* MP1<sup>9</sup>; *P. palleroniana* N26-GL<sup>10</sup>; *P. palleroniana* N26-GB<sup>11</sup> and *P. migulae* S10724<sup>12</sup> under low temperature nitrogen deficient conditions to identify the associated protein biomarker(s). In this series, for the first time, the proteome of psychrophilic and diazotrophic actinobacterium *R. qingshengii* S10107 is being analyzed. Besides exploring the adaptive mechanisms of diazotrophic actinomycete towards cold diazotrophy, present study provides a concluding remark to the previous outcomes and proclaims *CowN* as a potential protein biomarker for monitoring BNF under cold niches. Further, on the contrary to other cold adapted diazotrophs, *R. qingshengii* S10107 has displayed wide ranges of metabolic responses which could be the active area of investigation in future. Being a potential plant growth promoting strain<sup>13</sup>, *R. qingshengii* S10107 can also be explored for hill agricultural sustainability and zero budget farming.

## Results

The proteome of *R. qingshengii* S10107 was analyzed under two different physiological conditions i.e. low-temperature nitrogen deficient condition (without nitrogen source; BM) and low-temperature nitrogen sufficient condition (with nitrogen source, NB). The analysis was performed with three biological and three technical replicates.

**Identification of the differentially expressed proteins.** The whole proteomic dataset mapped to a total of 4838 proteins under BM and 4831 proteins under NB with a 1% false detection rate (FDR) (Fig. 1). There were 4813 proteins shared between the BM and NB along with 25 exclusive proteins under BM, and 18 under NB ( $p \leq 0.05$ ). Moreover, 77 proteins were found significantly up-regulated, while, 47 were down-regulated with  $\geq 2$ -fold change under BM in reference to its counterpart.

**Protein-protein interaction (PPI) network analysis.** Protein-protein interactions (PPIs) are considered very critical for cell survival. They are essential for the understanding of cellular physiology under different conditions. Therefore, PPI network analysis was performed to emphasize the crucial proteins involved in low-temperature nitrogen fixation (Fig. 2). Among BM proteins, *ArgF* (*Ornithine carbamoyltransferase*, Q8YMM6); *HisF* (*Imidazole glycerol phosphate synthase subunit*, A8HYT7); *ProB* (*Glutamate 5-kinase*, B2JHD6); *Eno* (*Enolase*, B2JIX0); *PyrG* (*CTP synthase*, B7KF08); *PheT* (*Phenylalanine—tRNA ligase beta subunit*, Q8YMH5); *PheS* (*Phenylalanine—tRNA ligase alpha subunit*, A0A1D8TAJ1); *GlmS* (*Glutamine—fructose-6-phosphate aminotransferase*, P59362) and *EcaA* (*Carbonic anhydrase*, P94170) were identified as hub nodes along with higher BC values (Table 1). Besides them, *AtpA* (*ATP synthase subunit alpha*, Q98EV6); *AtpH* (*ATP synthase subunit delta*, A8HS18); *UppP* (*Undecaprenyl-diphosphatase*, Q89WH1); *DeaD* (*ATP-dependent RNA helicase*, V9XLR7); *TpiA* (*Triosephosphate isomerase*, Q8YP17); *TufA* (*Elongation factor Tu*, C0ZVT7); *PckG* (*Phosphoenolpyruvate carboxykinase*, B2JJT8); *Tkt* (*Transketolase*, Q8YRU9) and *SucC* (*Succinate-CoA ligase*, A1KAU3) were identified as hub nodes only while, *LeuD1* (*3-isopropyl malate dehydratase*, Q98E51); *DnaE* (*Error-prone DNA polymerase*, Q98E34); *DdlA* (*D-alanine-D-alanine ligase*, B2JHF8); *Tal* (*Transaldolase*, P58561); *UreF* (*Urease accessory protein*, B5XU25); *UreG* (*Urease accessory protein*, A1KBB1) and *UreA* (*Urease subunit gamma*, Q8YQZ3) had large BC values. However, nine proteins viz. *NifH* (*Nitrogenase iron protein*, C7SI80), *NifA* (*nif-specific regulatory protein*, P56266), *NifL* (*Nitrogen fixation regulatory protein*, P06772), *NifB* (*FeMo cofactor biosynthesis protein*, P09825), *NifD* (*Nitrogenase molybdenum-iron protein*, P06120), *NifK* (*Nitrogenase molybdenum-iron protein*, P07329), *NirS* (*Nitrite reductase*, P24474), *CowN* (*N<sub>2</sub>-fixation sustaining protein*, C1DIY8) and *CooA* (*Carbon monoxide oxidation transcription regulator*, C1DIY7) did not appear in a match with *R. qingshengii* proteins, thereby, indicating towards the unavailability of respective proteins in the database (Supplementary material).

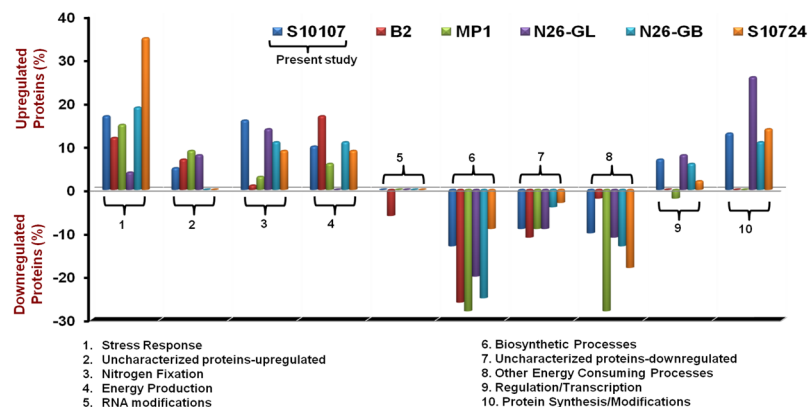


**Figure 2.** Protein-Protein Interaction (PPI) networks of differentially expressed proteins (with 2-fold or greater change) in NB (a) and BM (b). The networks a1 & b1 represent the total proteins expressed while a2 & b2 highlight the proteins involves in various processes. The protein-protein interactions were analyzed using the String database (version 9.1, Available online: <http://string-db.org/>) with the confidence score > 0.7. The analysis was performed in triplicates on the average value of proteins obtained from three replicates of the experiment.

In case of NB, *ProA* (*Gamma-glutamyl phosphate reductase*, Q89X85); *AroA* (*3-phosphoshikimate 1-carboxyvinyltransferase*, C1B1F1); *AroK* (*Shikimate kinase* Q98FY0); *FtsH* (*ATP-dependent zinc metalloprotease*, A0A059MKS5); *ThrS* (*Threonine- tRNA ligase*, Q0S1E2); *RpoA* (*DNA-directed RNA polymerase subunit alpha*, Q0S3E7); *KsgA* (*rRNA small subunit methyltransferase A*, A0A379MN38); *AspS* (*Aspartate- tRNA ligase*, Q0S0P3); *HisS* (*Histidine- tRNA ligase*; Q0S1B6) and *MiaA* (*tRNA dimethylallyltransferase*, B2JGD1) were recognized as hub as well as higher BC nodes. Proteins *IleS* (*Isoleucine— tRNA ligase*, U0FH2); *RplC* (*50S ribosomal protein L3*, X0QHH3); *GlcB* (*Acyl-[acyl-carrier-protein]—UDP-N-acetylglucosamine O-acyltransferase*, A8I491); *GltX2* (*Glutamate- tRNA ligase*, A0A1F2PR92); *Rnc* (*Ribonuclease 3*, Q0S2E1); *MetS1* (*Methionine— tRNA ligase*, Q0S4U4), and *RplT* (*50S ribosomal protein L20*, Q0S147) were only hubs; while, *FusA* (*Elongation factor G*, Q0SFF3); *CobB* (*Hydrogenobyrinate a,c-diamide synthase*, Q98KP1); *MurE1* (*UDP-N-acetylmuramoyl-L-alanyl-D-glutamate— 2,6-diaminopimelate ligase*, Q89FU2); *HbdA* (*3-hydroxybutyryl-CoA dehydrogenase*, Q45223); *ureB* (*Urease subunit beta*, Q98CY6); *hemZ* (*Ferrocyclase*, Q0S0F7) and *rpoA* (*DNA-directed RNA polymerase subunit alpha*, Q0S3E7) were found to have higher BC under NB.

Parameters	Gene Names	
	Nitrogen Stress (BM)	Nitrogen Sufficient (NB)
(Hub + Large BC) nodes	<i>argF, hisF, ecaA, eno, pyrG, pheT, pheS, glmS, proB</i>	<i>proA, aroA, aroK, ftsH, thrS, rpoA, ksgA, aspS, hisS, miaA</i>
Hub nodes	<i>argF, eno, hisF, pyrG, pheT, pheS, glmS, uppP, ecaA, deaD, tpiA, tufA, atpA, atpH, tkt, pckG, sucC, proB</i>	<i>ileS, rplC, glcB, gltX2, rnc, metS1, rplT, proA, aroA, aroK, ftsH, thrS, rpoA, ksgA, aspS, hisS, miaA</i>
Large BC Nodes	<i>argF, hisF, leuD1, dnaE, glmS, ddlA, tal, ureF, ureG, ureA, eno, pyrG, pheT, pheS, proB, ecaA</i>	<i>cobB, murE1, hbdA, ureB, hemZ, ABG93733.1 (fusA), rpoA, proA, aroA, aroK, ftsH, thrS, rpoA, ksgA, aspS, hisS, miaA</i>

**Table 1.** Significant key nodes of the PPI Network\*. \*As revealed by PPI network analysis (SM).



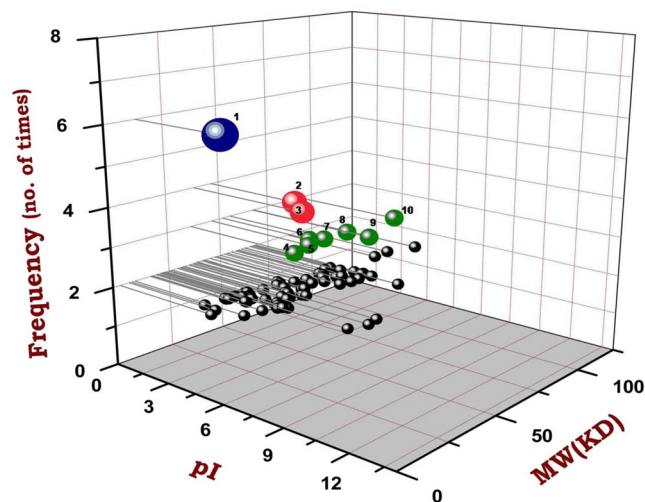
**Figure 3.** Comparative protein profiling among *Rhodococcus qingshengii* S10107 (present study); *Dyadobacter psychrophilus* B2 and *P. jessenii* MP1<sup>9</sup>; *P. palleroniana* N26-GL<sup>10</sup>; *P. palleroniana* N26-GB<sup>11</sup> and *P. migulae* S10724<sup>12</sup> under low temperature N<sub>2</sub> fixing conditions, according to their biological functions.

**GO-based functional characterization of the differentially expressed proteins.** All the differentially expressed proteins were characterized on the basis of their biological functions and compared with the earlier studies (Fig. 3). Proteins associated with stress response, nitrogen fixation and energy production were up-regulated while that of biosynthetic processes and energy consuming processes were observed to down-regulate by all the diazotrophs studied under low-temperature nitrogen deficient conditions. Besides them, a good fraction of the up-regulated and down-regulated proteins were found uncharacterized which need a detailed investigation in the future. In case of the present study, 17% of the proteins were related to stress response, followed by nitrogen fixation (16%), protein synthesis/modifications (13%), energy production (10%), gene regulation/transcription (7%) and uncharacterized (5%). Among down-regulated proteins, the majority of the proteins were involved in biosynthetic processes (13%) and energy consuming processes (10%) besides uncharacterized proteins (9%).

Additionally, functional enrichments in the PPI network were explored to provide a more specific description of the differentially expressed proteins in terms of their involvement in amino acid biosynthesis (AAB); metabolic pathways (MP); nitrogen metabolism (NM) and production of secondary metabolites (Fig. 2).

**Amino acid biosynthesis (AAB).** Proteins *LeuD1*, *ArgF*, *ArgC* (*N*-acetyl- $\gamma$ -glutamyl-phosphate reductase, B7JY20), *HisF*, *Putative glutamate—cysteine ligase* (AZC2303,A8I5N7), *ProB*, *ThrB* (*Homoserine kinase*, A9HS91), *DapD* (*2,3,4,5-tetrahydropyridine-2,6-dicarboxylate N-succinyltransferase*,B2JID7), *Alr* (*Alanine racemase*, W8HAH8) and *AroB* (*3-dehydroquinate synthase*, Q0S0N0) were involved in biosynthesis of different amino acids under BM, while, *DapA1* (*4-hydroxy-tetrahydrodipicolinate synthase*, A0A1H4N6J7), *TyrS* (*Tyrosine—tRNA ligase*, Q98NS5), *IleS*, *ProA*, *AroE1* (*Shikimate dehydrogenase*, Q98DY3) *AroA*, *ThrS*, *HisS*, *AspS*, and *AroK* were responsible for AAB under NB.

**Metabolic pathways (MP).** Proteins involved in MP were *Bphy\_3511* (*Acetaldehyde dehydrogenase1*, B2JLM7), *LeuD1*, *UppP*, *PyrG*, *ArgF*, *CoaE* (*Dephospho-CoA kinase*, Q98DY2), *Lgt* (*Prolipoprotein diacylglycerol transferase*, A9HBY2), *HisF*, *BioD* (*ATP-dependent dethiobiotin synthetase*, B7K5E6), *MurC* (*UDP-N-acetylmuramate—L-alanine ligase*, B2JHF9), *MloA* (*Protein MloA*, Q8RN11), *LipB* (*Octanoyltransferase*, Q89JM6), *YbeY* (*Endoribonuclease YbeY*, Q98BK1), *ProB*, *atpA*, *AtpH*, *ThrB*, *TufA*, *GpmA* (*2,3-bisphosphoglycerate-dependent phosphoglycerate mutase*, B2JC95), *AskA* (*Acetate kinase*, B8HVB6), *ThyA* (*Thymidylate synthase*, Q0SEI1), *RpsF* (*30S ribosomal protein S6*), *Tkt*, *AccC* (*Biotin carboxylase*, Q06862), *PanD* (*Aspartate 1-decarboxylase*, Q8YR79), *HutI* (*Imidazolonepropionase*, B2JCI1), *PckG*, *SucC*, *Eno*, *Urease accessory proteins*, *glmM* (*Phosphoglucosamine mutase*, Q98F91), *GlmS*, *DdlA*, *Dxr* (*1-deoxy-D-xylulose 5-phosphate reductoisomerase*, B7K5G6), *Mqo* (*Probable malate:uinine oxidoreductase*, Q89XM4), *AroB*, *TpiA*, *Tal* etc. under BM and *DapA1*, *GlcB*, *NadK* (*NAD kinase*, Q0S170), *ArgJ* (*Arginine biosynthesis protein*, Q8YPF9), *PanB* (*3-methyl-2-oxobutanoate hydroxymethyltransferase*, Q0SHJ0), *ProA*, *AroE1*, *AcpS* (*Holo-[acyl-carrier-protein]*



**Figure 4.** 3D Scatter plot illustrating the expressional pattern of the proteins associated with low-temperature  $N_2$  fixation. OriginPro software was used to perform the analysis based on the Isoelectric point of the proteins (pI), molecular weight (KD) and their frequency to occur as observed during previous studies. Proteins with  $\geq 3$  time of occurrence are numbered viz. 1, *CowN*; 2, *Eno*; 3, *PyrG*; 4, *IlvC*; 5, *ClpX*; 6, *MnmE*; 7, *AtpA*; 8, *HtpG*; 9, *MutL*; 10, *IleS*.

*synthase*, Q0SGU2), *HbdA*, *RpiA* (Ribose-5-phosphate isomerase A, B7K6D3), *SpeE* (Polyamine aminopropyltransferase), *HutU* (Urocanate hydratase, Q89GV4), *FabH* (3-oxoacyl-[acyl-carrier-protein] synthase, Q982Z8), *GlmU* (Bifunctional protein *GlmU*, Q0S4N3), *Rha1\_ro06238* (Isocitrate dehydrogenase, Q0S371), *GlpK* (Glycerol kinase, A1K962), *AroA*, *MoaA* (GTP 3',8-cyclase, A0A143Q879), *CobB*, *AroK*, *HemZ* etc. under NB.

**Nitrogen metabolism (NM).** In case of Nitrogen metabolism (NM), *R. qingshengii* S10107 has up-regulated (BM) *UppP*, *ArgF*, *Lgt*, *HisF*, *BioD*, *MurC*, *MloA*, *LipB*, *YbeY*, *AZC2303*, *ProB*, *AtpA*, *AtpH*, *TufA*, *HemL* (*Glutamate-1-semialdehyde 2,1-aminomutase*, B8HYK1), *GpmA*, *ThyA*, *PanD*, *Eno*, *Ribosomal proteins* and *Urease* accessory proteins while, *Arc* (*Proteasome-associated ATPase*, Q0SIF4), *NadK*, *PanB*, *ProA*, *FusA*, *KsgA*, *GlmU*, *AroA*, *MoaA*, *GltX2*, *CoaD* (*Phosphopantetheine adenylyltransferase*, B5XTG9), *ThrS*, *Apt* (*Adenine phosphoribosyltransferase*, Q0S1C1), *HisS*, *AroK*, *AspS* etc. were down-regulated (NB).

**Secondary metabolites (SM).** The proteins involved in secondary metabolite formation (SM) under BM were *LeuD1*, *ArgF*, *ArgC*, *MloA*, *YbeY*, *ThrB*, *HemL*, *GpmA*, *PanD*, *PurD*, *PckG*, *SucC*, *Eno*, *Dxr*, *Mqo*, *HemE*, *AroB*, *TpiA* and *Tal* while, *DapA1*, *argJ*, *Rha1\_ro01056* (*Probable glycogen debranching enzyme*, Q0SHV3), *PanB*, *ProA*, *AroE1*, *Rha1\_ro01447* (*Glycogen phosphorylase*, Q0SGS1), *RpiA*, *PurM* (*Phosphoribosylformylglycinamide cyclo-ligase*, Q0S760), *Rha1\_ro05098* (*3-demethylubiquinone-9 3-O-methyltransferase*, Q0S6F7), *Rha1\_ro06238*, *AroA*, *GltX2*, *MiaA*, *AroK*, *AcnA2* (*Aconitate hydratase*, Q0S0G5) and *HemZ* were under NB.

**CowN as a protein biomarker.** In order to identify the protein biomarker(s) associated with the nitrogen fixation process at lower temperatures, the author group has performed a series of experiments on six different cold-adapted microorganisms viz. *P. palleroniana* N26-GL<sup>10</sup>; *Dyadobacter psychrophilus* B2 and *P. jessenii* MP1<sup>9</sup>; *P. palleroniana* N26-GB<sup>11</sup> and *P. migulae* S10724<sup>12</sup> including present study (*R. qingshengii* S10107). Moreover, two distinct proteomic approaches i.e. gel-based<sup>9,11,12</sup> and gel-less approach<sup>10</sup> were also employed to rectify the experimental biases.

In this perspective, comparative protein profiling of all the diazotrophs revealed that *CowN* was significantly up-regulated in all the six experimentation followed by *Eno* which was up-regulated in four cases only (Fig. 4). *PyrG* was also found to express in four instances, however, in three cases it was a down-regulated and up-regulated only once (SM). Similarly, proteins *IlvC*, *ClpX*, *MnmE*, *AtpA*, *HtpG*, *MutL*, and *IleS* were expressed in three cases only. Among them, *AtpA*, *HtpG*, and *MutL* were up-regulated in all the three cases while; *MnmE* and *IleS* were down-regulated, subsequently.

## Discussion

Genus *Rhodococcus* is well known for its broad metabolic versatility, however, the diazotrophic potential of this actinobacterium is least studied. All the previous proteomic investigations, which were performed on *Rhodococcus*, only highlighted its catabolic behavior towards different organic compounds viz. Fluoranthene<sup>14</sup>, toluene and phenol<sup>15</sup>, propane<sup>16</sup>, triacylglycerol<sup>17</sup>, 4,4'-Dithiodibutyric Acid<sup>18</sup>, benzoate<sup>19</sup> and polyaromatic compounds<sup>20</sup>. In this perspective, the present study for the first time unravels its proteomic response towards the abiotic stress and identified a protein biomarker for low-temperature nitrogen deficient conditions.

The study revealed that *R. qingshengii* S10107 expressed all the three structural *nif* genes viz. *nifD*, *nifH* and *nifK* under BM along with their regulators i.e. *nifA*, *nifL*, and *nifB*; thereby, confirming its nitrogen fixation potential<sup>21</sup>. However, complete *nif* regulon could not be observed, which might be due to the fact that diazotrophs do not express all the *nif* genes at a time in order to minimize their energy consumption<sup>22</sup>. Additionally, the accumulated *Nif* proteins may get cleaved by the proteases i.e. *ClpX* (N1M9I2) as also evident from earlier studies<sup>10</sup>. Therefore, the kinetics of *nif* regulon must be studied over time to scan the complete network. Besides the *Nif* proteins, two associated proteins were also observed viz. *CO Weal-Nitrogenase (CowN)* and *CO-responsive regulator (CooA)*. *CowN* is known to protect the nitrogenase from Carbon monoxide (CO) stress which is supposed to induce under low-temperature nitrogen deficient conditions<sup>10,23,24</sup>.

*R. qingshengii* S10107 was found to employ unique strategies for performing important cellular processes under low-temperature nitrogen deficient condition. Under BM, proteins *3-isopropylmalate dehydratase*, *2,3,4,5-tetrahydropyridine-2,6-dicarboxylate N-succinyltransferase* and *Alanine racemase* were up-regulated which are responsible for the biosynthesis of the amino acids Leucine, Lysine, and Alanine, respectively. Lysine and Alanine in a peptide help it to get  $\alpha$ -helical conformation<sup>25</sup> and thus, contributed towards low-temperature adaptation<sup>26</sup>. Similarly, *Tyrosine-tRNA ligase*, *Isoleucine-tRNA ligase* and *Aspartate-tRNA(Asp/Asn) ligase* were down-regulated which are responsible for the biosynthesis of Tyrosine, Iso-leucine, and Aspartate. These amino acids tend to be decreased when the organism exposed to the cold<sup>27</sup>.

Further, among the up-regulated proteins, *Ornithine carbamoyltransferase (ArgF)* and *N-acetyl-gamma-glutamyl-phosphate reductase (ArgC)* are known for the biosynthesis of Arginine, which mostly serves as a key source of carbon and nitrogen for ornithine, proline, and pyrimidine<sup>28,29</sup>. Up-regulation of *ArgC* during low-temperature nitrogen fixation needs to be investigated in detail because of the two main reasons - first - it catalyzes high energy consuming reaction which is otherwise very rare event during such conditions<sup>10,30</sup> and second - it causes growth delay as well as inefficient nodule formation in diazotrophs<sup>30</sup>. In the case of Proline biosynthesis, the genes encode for protein *Gamma-glutamyl phosphate reductase (ProA)* and *Glutamate 5-kinase (ProB)* are organized in a single operon i.e. *proBA*. In the present study, S10107 strain increases the expression of *ProB* while, decreases *ProA*, thereby, regulated the intracellular concentration of Proline. *P. putida*, *K. aerogenes*, and *E. coli*, are known to enhance the catabolism of Proline under nitrogen-deficient conditions<sup>29,31</sup>. A similar condition was also observed in case of *Histidine* which showed up-regulation of *Imidazole glycerol phosphate synthase subunit (HisF)* and down-regulation of *Histidine-tRNA ligase (HisS)*. These expressional patterns can be supported with the fact that nitrogen fixation is an energy consuming process and therefore, down-regulation and/or shutting down the nonessential proteins is pre-requisite to the cell survival.

Cold along with nitrogen deficiency represents dual-stress condition for the microorganisms involving environmental as well as nutritional stress. Therefore, S10107 strain up-regulated stress associated proteins viz. *UvrABC system protein A (Q0SI39)*; *DNA repair protein RecO (Q985A3)*, *Chaperone protein HtpG (Q0S467)*, and *Error-prone DNA polymerase (Q98E34)*. *UvrABC* and *RecO* are the multienzyme systems which aid in DNA repair; while *Chaperone HtpG* protects the cellular proteins from the stress conditions<sup>9,10</sup>. Also, up-regulation of *Error-prone DNA polymerase*, *UvrABC*, and *RecO* revealed the cellular need for DNA maintenance under given physiological conditions<sup>32,33</sup>. Conclusively, these proteins are responsible for tight regulation of the housekeeping genes along with their protection to sustain the life under multi-stress conditions.

Under BM conditions, S10107 strain showed a distinct metabolic behavior by up-regulating the proteins of peptidoglycan synthesis (*Undecaprenyl-diphosphatase*, Q89WH1 and *D-alanine-D-alanine ligase*, B2JHF8); Glutamine synthesis (*CTP synthase*, B7KF08; *Glutamine-fructose-6-phosphate aminotransferase*, P59362); Biotin metabolism (*ATP-dependent dethiobiotin synthetase*, B7K5E6 and *Biotin carboxylase*, Q06862); Nitrogen metabolism (*Urease enzyme complex*) and Nucleotide metabolism (*dITP/XTP pyrophosphatase*, Q98DN4 and *Thymidylate synthase*, Q0SEI1) while down-regulating the enzymes for Lipid biosynthesis (*Acyl-[acyl-carrier-protein]-UDP-N-acetylglucosamine O-acyltransferase*, A8I491 and *3-hydroxybutyryl-CoA dehydrogenase*, Q45223) and Pantothenate biosynthesis (*3-methyl-2-oxobutanoate hydroxymethyltransferase*, Q0SHJ0). Up-regulation of the peptidoglycan synthesis proteins while down-regulation of Lipid biosynthesis-related proteins revealed that *R. qingshengii* S10107 tend to thicken their cell wall to respond towards external stress. Further, enhancement of Glutamine synthesis proteins can be justified by the general tendency of the microbial cell to prefer ammonium and glutamine as a nitrogen source<sup>29</sup>. Furthermore, *CTP synthase* is responsible to regulate cytidine triphosphate (CTP) concentration within a cell. Killer *et al.*, (2018) has developed a *CTP synthase* based tool to identify the family Bifidobacteriaceae. *Glutamine-fructose-6-phosphate aminotransferase* has glutaminase activity that links hexosamine biosynthetic pathway to the hexose pathway and thereby, plays a regulatory role for the nutrient-sensing system<sup>34</sup>.

The up-regulation of the two enzymes of the Pentose phosphate pathway viz. *Transketolase (Q8YRU9)* and *Transaldolase (P58561)* along with *Ribose import ATP-binding protein RbsA (Q0S9A4)* under BM suggests the cellular need of excessive ribose sugar instantly. It can again be supported with the down-regulation of *Ribose-5-phosphate isomerase (B7K6D3)* which is responsible for the conversion of ribose-5-phosphate to ribulose-5-phosphate. In addition to these, three enzymes of Gluconeogenesis (*2,3-bisphosphoglycerate-dependent phosphoglycerate mutase*, B2JC95; *Phosphoenolpyruvate carboxykinase*, B2JJT8, and *Triosephosphate isomerase*, Q8YP17) and two of glucose catabolism (*Enolase*, B2JIX0 and *Succinate-CoA ligase*, A1KAU3) were also found to up-regulate under BM. Contrary to these, proteins related to the glycogen metabolism (*glycogen debranching enzyme*, Q0SHV3 and *Glycogen phosphorylase*, Q0SGS1) and glycerol metabolism (*Glycerol kinase*, A1K962) were found down-regulated along with *Isocitrate dehydrogenase (Q0S371)*, an important enzyme of TCA cycle. It might be the part of the adaptation strategies of the microorganism so that it could utilize carbohydrates as well as non-carbohydrates precursors for fulfilling its energy demands. The similar results were also observed by Ma *et al.*<sup>35</sup>. This group have used SILAC method for analysis the proteome of *Edwardsiella tarda* ATCC 15947 under prolonged cold stress and reported that the enzymes of Gluconeogenesis were significantly enhanced under cold.

Furthermore, Tullio *et al.*<sup>36</sup> have analysed the proteome of the diazotroph *Rhizobium freirei* and observed the up-regulation of Gluconeogenesis associated enzymes under metabolic stress conditions.

The expression of CowN was reported in the proteome of all the six previously studied cold adapted diazotrophs *viz.* *P. palleroniana* N26-GL<sup>10</sup>; *Dyadobacter psychrophilus* B2 and *P. jessenii* MP1<sup>9</sup>; *P. palleroniana* N26-GB<sup>11</sup> and *P. migulae* S10724<sup>12</sup> including this study (*R. qingshengii* S10107) under cold and nitrogen deficient conditions. Therefore, this protein is being identified as a protein biomarker for monitoring BNF under cold niches. Several protein-based biomarkers are already being used for the clinical purpose<sup>6</sup>, veterinary science<sup>37</sup>, bioremediation<sup>7</sup> and heavy metal detection<sup>8</sup>. Saito *et al.*<sup>38</sup> detected the multiple nutrient stresses at Pacific Ocean biomes by using the protein biomarkers. Recently, Andrade-Herrera *et al.*<sup>39</sup> have developed an earthworm based biomarker for pesticides and toxicity assessment in agricultural soils. However, information is not available about such biomarkers which can monitor any biogeochemical processes under abiotic stress that is going to be imperative for precision farming in the future.

In conclusion, the present study provides a detailed investigation on the adaptive responses of *R. qingshengii* S10107 towards cold diazotrophy which can be explored for further advance research. Further, CowN can be proclaimed as a protein biomarker for diazotrophic identification and monitoring of BNF under cold and nitrogen deficient conditions. Moreover, the role of carbon monoxide and ribose sugar in low-temperature diazotrophy should be explored for detailed investigation.

## Methods

**Bacterial strain and growth conditions.** *R. qingshengii* S10107 (JX173283) was originally isolated from the rhizospheric soil of *Phaseolus vulgaris* L. (Red Kidney Bean, RKB) from Chhiplakot region (3,290 m, 30.06 °N, 79.01 °E) of Kumaun Himalayas on the Burk's Medium (nitrogen deficient medium) at 10 °C. This strain has already been characterized as diazotroph<sup>12</sup>.

**Proteome extraction and LC-MS/MS analysis.** The proteome of S10107 strain was extracted at the mid-log phase in triplicates as per the earlier reports<sup>9,11</sup> (Supplementary material). ACQUITY UPLC system (Waters, UK) having ACQUITY UPLC BEH C18 column (Waters, UK)(150mm X 2.1mm X 1.7µm) was used to perform liquid chromatography. A gradient of two mobile phases A (0.1% Formic Acid in WATER), and B (0.1% Formic Acid in Acetonitrile) were used for the chromatographic separation. Further, mass spectrometric detection was performed by using SYNAPT G2 QTOF (Waters, UK) having an electrospray ionization (ESI) source.

**Database searching and analysis.** The observed proteins were matched to *Pseudomonas* protein database which was downloaded from Swiss-Prot through PLGS software 3.0.2. The data analysis was performed in triplicates using the following parameters: peptide tolerance = 50 ppm, fragment tolerance = 100 ppm, minimum no. of fragment matches (for Proteins) = 5, minimum no. of fragment matches (for peptides) = 2, minimum no. of peptide matches (for proteins) = 2, No. of missed cleavages = 2 and modifications include oxidation-m and carbamidomethyl-c. The FDR (false discovery rate) was set as <1% on both protein and peptide levels. The data were normalized across the conditions including the replicates using spectral counts method<sup>40,41</sup>. For identifying differentially expressed proteins, the normalized peptide matches of the treatment were divided by that of control conditions. Statistical analysis was performed through pair-wise Student's t-test ( $p \leq 0.05$ ).

**Construction of the PPI network.** Among the total expressed proteins, only two types of proteins were selected for the construction of the PPI (protein-protein interactions) Networks: (1) unique and (2) which showed 2-fold or higher change in their expression. PPI information of proteins was obtained from the String database (version 9.1; <http://string-db.org/>) having the confidence score > 0.7. It was further imported in Cytoscape (version 3.6.0; <http://www.cytoscape.org/>) and the union calculation was performed, followed by removal of the duplicated edges by using Advanced Network Merge<sup>40,42</sup>.

**Gene ontology (GO) enrichment analysis.** The ontologies and their genes from the PPI network of selected proteins were identified using BiNGO plugin (version 3.0.3; <http://apps.cytoscape.org/apps/bingo>) for the Cytoscape which are: molecular functions, biological processes, and cellular components. For a detailed description of the biological process, the ClueGo plugin (version 2.5.0; <http://apps.cytoscape.org/apps/cluego>) of Cytoscape was used to integrate several ontology sources by extracting the non-redundant biological information from several databases *viz.* KEGG, REACTOME, GO, Wiki Pathways and BioCarta<sup>40,43</sup>.

**Statistical analysis.** Significance was determined using unpaired two-tailed t-test or linear regression analysis (GraphPad Prism software, version 6.01). Differences were noted as significant \* $p \leq 0.05$  for t-test or linear regression analysis. The scatter plot analysis was performed by using the software PERMANOVA<sup>44</sup>.

## Data availability

The datasets used and/or analyzed during the current study are available from the corresponding authors on reasonable request.

Received: 29 August 2019; Accepted: 3 December 2019;

Published online: 30 December 2019

## References

- Global Industry Report. Biofertilizers Market Size, Share & Trends Analysis Report By Product (Nitrogen Fixing, Phosphate Solubilizing), By Application (Seed Treatment, Soil Treatment), And Segment Forecasts, 2012–2022. <https://www.grandviewresearch.com/industry-analysis/biofertilizers-industry>. Accessed on 11/04/2019 (2018).
- Saiz, E., Sgouridis, F., Drijfhout, F. P. & Ullah, S. Biological nitrogen fixation in peatlands: Comparison between acetylene reduction assay and  $^{15}\text{N}_2$  assimilation methods. *Soil. Biol. Biochem.* **131**, 157–165, <https://doi.org/10.1016/j.soilbio.2019.01.011> (2019).
- Roley, S. S., Xue, C., Hamilton, S. K., Tiedje, J. M. & Robertson, G. P. Isotopic evidence for episodic nitrogen fixation in switchgrass (*Panicum virgatum* L.). *Soil. Biol. Biochem.* **129**, 90–98, <https://doi.org/10.1016/j.soilbio.2018.11.006> (2019).
- Blagodatskaya, E. & Kuzyakov, Y. Active microorganisms in soil: Critical review of estimation criteria and approaches. *Soil. Biol. Biochem.* **67**, 192–211, <https://doi.org/10.1016/j.soilbio.2013.08.024> (2013).
- Chen, W., Zhang, X., Xiong, D., Jin, J. M. & Tang, S. Y. Engineering the effector specificity of regulatory proteins for the in vitro detection of biomarkers and pesticide residues. *Appl. Microbiol. Biotechnol.* **103**, 3205–3213, <https://doi.org/10.1007/s00253-019-09679-1> (2019).
- Wan, Y. *et al.* Electrical signal reporter, pore-forming protein, for rapid, miniaturized, and universal identification of microorganisms. *Anal. Chem.* **90**, 9853–9858, <https://doi.org/10.1021/acs.analchem.8b01933> (2018).
- Liu, Z. *et al.* Highly sensitive microbial biosensor based on recombinant *Escherichia coli* overexpressing catechol 2,3-dioxygenase for reliable detection of catechol. *Biosens. Bioelectron.* **126**, 51–58, <https://doi.org/10.1016/j.bios.2018.10.050> (2019).
- Kim, H. J., Jeong, H. & Lee, S. J. Synthetic biology for microbial heavy metal biosensors. *Anal. Bioanal. Chem.* **410**, 1191–1203, <https://doi.org/10.1007/s00216-017-0751-6> (2018).
- Suyal, D. C., Kumar, S., Yadav, A., Shouche, Y. & Goel, R. Cold stress and nitrogen deficiency affected protein expression of psychrotrophic *Dyadobacter psychrophilus* B2 and *Pseudomonas jessenii* MP1. *Front. Microbiol.* **8**, 1–6, <https://doi.org/10.3389/fmicb.2017.00430> (2017).
- Suyal, D. C., Kumar, S., Joshi, D., Soni, R. & Goel, R. Quantitative proteomics of psychrotrophic diazotroph in response to nitrogen deficiency and cold stress. *J. Proteomics* **187**, 235–242, <https://doi.org/10.1016/j.jpro.2018.08.005> (2018).
- Soni, R. *et al.* Differential proteomic analysis of Himalayan psychrotolerant diazotroph *Pseudomonas palleroniana* N26 Strain under low temperature diazotrophic conditions. *CryoLetters* **36**, 74–82 (2015).
- Suyal, D. C., Yadav, A., Shouche, Y. & Goel, R. Differential proteomics in response to low temperature diazotrophy of Himalayan psychrophilic nitrogen fixing *Pseudomonas migulae* S10724 strain. *Curr. Microbiol.* **68**, 543–550, <https://doi.org/10.1007/s00284-013-0508-1> (2014).
- Joshi, D., Chandra, R., Suyal, D. C., Kumar, S. & Goel, R. Impact of bioinoculants *Pseudomonas jessenii* MP1 and *Rhodococcus qingshengii* S10107 on *Cicer arietinum* yield and soil nitrogen status. *Pedosphere* **29**, 388–399, [https://doi.org/10.1016/S1002-0160\(19\)60807-6](https://doi.org/10.1016/S1002-0160(19)60807-6) (2019).
- Xu, J., Wang, H. & Kong, D. 2-DE Compared with iTRAQ-based proteomic analysis of the functional regulation of proteins in *Rhodococcus* sp. BAP-1 response to fluoranthene. *IOP Conf. Ser. Earth. Environ. Sci.* **111**, 1–7, <https://doi.org/10.1088/1755-1315/111/1/012032> (2018).
- Kuang, S., Fan, X. & Peng, R. Quantitative proteomic analysis of *Rhodococcus ruber* responsive to organic solvents. *Biotech. Biotechnol. Equip.* **32**, 1418–1430, <https://doi.org/10.1080/13102818.2018.1533432> (2018).
- Tupa, P. R. & Masuda, H. Comparative proteomic analysis of propane metabolism in *Mycobacterium* sp. strain ENV421 and *Rhodococcus* sp. strain ENV425. *J. Mol. Microbiol. Biotechnol.* **28**, 107–115, <https://doi.org/10.1159/000490494> (2018).
- Costa, J. S. D., Silva, R. A., Leichert, L. & Alvarez, H. M. Proteome analysis reveals differential expression of proteins involved in triacylglycerol accumulation by *Rhodococcus jostii* RHA1 after addition of methyl viologen. *Microbiol. Soc* **163**, 343–354, <https://doi.org/10.1099/mic.0.000424> (2017).
- Khairy, H. *et al.* Genome and Proteome Analysis of *Rhodococcus erythropolis* MI2: Elucidation of the 4,4'-Dithiodibutyric Acid Catabolism. *Plos. One.* **11**, e0167539, <https://doi.org/10.1371/journal.pone.0167539> (2016).
- Patrauchan, M. A. *et al.* Proteomic Analysis of Survival of *Rhodococcus jostii* RHA1 during Carbon Starvation. *Appl. Environ. Microbiol.* **78**, 1–12, <https://doi.org/10.1128/AEM.01293-12> (2012).
- Tomas-Gallardo, L. *et al.* Proteomic and transcriptional characterization of aromatic degradation pathways in *Rhodococcus* sp. strain TFB. *Proteomics* **1**, S119–32, <https://doi.org/10.1002/pmic.200500422> (2006).
- Kuypers, M. M. M., Marchant, H. K. & Kartal, B. The microbial nitrogen-cycling network. *Nat. rev. Microbiol.* **16**, 263–276, <https://doi.org/10.1038/nrmicro.2018.9> (2018).
- Poza-Carrion, C., Jiménez-Vicente, E., Navarro-Rodríguez, M., Echavarrri-Erasun, C. & Rubio, L. M. Kinetics of *nif* gene expression in a nitrogen-fixing bacterium. *J. Bacteriol.* **196**, 595–603, <https://doi.org/10.1128/JB.00942-13> (2014).
- Kerby, R. L. & Roberts, G. P. Sustaining  $\text{N}_2$ -Dependent Growth in the Presence of  $\text{CO}_2$ . Sustaining  $\text{N}_2$ -dependent growth in the presence of  $\text{CO}_2$ . *J. Bac.* **193**, 774–777, <https://doi.org/10.1128/JB.00794-10> (2011).
- Kourtelesis, M., Panagiotopoulou, S. & Verekios, X. E. Influence of the Support on the Reaction Network of Ethanol Steam Reforming at Low Temperatures Over Pt Catalysts. *Top. Catal.* **58**, 1202–1217, <https://doi.org/10.1007/s11244-015-0485-7> (2015).
- Johnson, N. P., Lindstrom, J., Baase, W. A., Von Hippel, P. H. & Double-stranded, D. N. A. templates can induce a-helical conformation in peptides containing lysine and alanine: Functional implications for leucine zipper and helix-loop-helix transcription factors. *Proc. Natl. Acad. Sci. USA* **91**, 4840–4844 (1994).
- Siddiqui, K. S. *et al.* Psychrophiles. *Annu. Rev. Earth. Planet. Sci.* **41**, 87–115, <https://doi.org/10.1146/annurev-earth-040610-133514> (2006).
- Gianese, G., Argos, P. & Pascarella, S. Structural adaptation of enzymes to low temperatures. *Protein. Eng.* **14**, 141–148, <https://doi.org/10.1093/protein/14.3.141> (2001).
- Caldara, M., Charlier, D. & Cunin, R. The arginine regulon of *Escherichia coli*: whole-system transcriptome analysis discovers new genes and provides an integrated view of arginine regulation. *Microbiology.* **152**, 3343–54, <https://doi.org/10.1099/mic.0.29088-0> (2006).
- Christgen, S. L. & Becker, D. F. Role of Proline in Pathogen and Host Interactions. *Antioxid. Redox. Signal.* **30**, 683–709, <https://doi.org/10.1089/ars.2017.7335> (2019).
- Vargas-Lagunas, C. *et al.* Overproduction of Sinorhizobium meliloti ArgC (N-acetyl-gamma-glutamyl phosphate reductase) promotes growth delay and inefficient nodules. *FEMS Microbiol. Lett.* **364**, fnx233, <https://doi.org/10.1093/femsle/fnx233> (2017).
- Vilchez, S., Manzanera, M. & Ramos, J. L. Control of expression of divergent *Pseudomonas putida* put promoters for proline catabolism. *Appl. Environ. Microbiol.* **66**, 5221–5225 (2000).
- Rinaldi, G., Rossi, M. & Fendt, S. M. Metabolic interactions in cancer cellular metabolism at the interface between the microenvironment, the cancer cell phenotype and the epigenetic landscape. *Wiley Interdiscip. Rev. Syst. Biol. Med.* **10**, e1397, <https://doi.org/10.1002/wsbm.1397> (2018).
- Rinaldo, S., Giardina, G., Mantoni, F., Paone, A. & Cutruzzola, F. Beyond nitrogen metabolism: nitric oxide, cyclic-di-GMP and bacterial biofilms. *FEMS Microbiol. Lett.* **365**, 1–9, <https://doi.org/10.1093/femsle/fny029> (2018).
- Foley, S. *et al.* Characterisation of glutamine fructose-6-phosphate amidotransferase (EC 2.6.1.16) and N-acetylglucosamine metabolism in *Bifidobacterium*. *Arch. Microbiol.* **189**, 157–167, <https://doi.org/10.1007/s00203-007-0307-9> (2008).
- Ma, W. *et al.* Stable isotope labelling by amino acids in cell culture (SILAC) applied to quantitative proteomics of *Edwardsiella tarda* ATCC 15947 under prolonged cold stress. *Microb Pathogen* **125**, 12–19, <https://doi.org/10.1016/j.micpath.2018.09.006> (2018).



36. Tullio, L. D., Gomes, D. F., Silva, L. P., Hungria, M. & Batista, J. S. D. Proteomic analysis of *Rhizobium freirei* PRF 81T reveals the key role of central metabolic pathways in acid tolerance. *Appl. Soil Ecol.* **135**, 98–103, <https://doi.org/10.1016/j.apsoil.2018.11.014> (2019).
37. Gagaoua, M., Bonnet, M., De Koning, L. & Picard, B. Reverse Phase Protein array for the quantification and validation of protein biomarkers of beef qualities: The case of meat color from Charolais breed. *Meat Science*. **145**, 308–319, <https://doi.org/10.1016/j.meatsci.2018.06.039> (2018).
38. Saito, M. A. *et al.* Multiple nutrient stresses at intersecting Pacific Ocean biomes detected by protein biomarkers. *Science*. **345**, 1173–1177, <https://doi.org/10.1126/science.1256450> (2014).
39. Andrade-Herrera, M. *et al.* Presence of Pesticides and Toxicity Assessment of Agricultural Soils in the Quintana Roo Mayan Zone, Mexico Using Biomarkers in Earthworms (*Eisenia fetida*). *Water Air Soil Pollut* **230**, 1–13, <https://doi.org/10.1007/s11270-019-4105-1> (2019).
40. Sabetian, S. & Shamsir, M. S. Systematic analysis of protein interaction network associated with Azoospermia. *Int. J. Mol. Sci.* **17**, E1857, <https://doi.org/10.3390/ijms17111857> (2016).
41. Oh, Y., Robertson, S. L., Parker, J., Muddiman, D. C. & Dean, R. A. Comparative proteomic analysis between nitrogen supplemented and starved conditions in *Magnaporthe oryzae*. *Proteome Science*. **15**, 1–12, <https://doi.org/10.1186/s12953-017-0128-y> (2017).
42. Huo, M., Wang, Z., Wu, D., Zhang, Y. & Qiao, Y. Using co-expression protein interaction network analysis to identify mechanisms of danshensu affecting patients with coronary heart disease. *Int. J. Mol. Sci.* **18**, E1298, <https://doi.org/10.3390/ijms18061298> (2017).
43. Pedron, N. *et al.* Proteomic responses of European flounder to temperature and hypoxia as interacting stressors: Differential sensitivities of populations. *Sci. Total. Environ.* **15**, 890–899, <https://doi.org/10.1016/j.scitotenv.2017.02.068> (2017).
44. Anderson, M. J. Permutational multivariate analysis of variance (PERMANOVA). *Wiley Stats. Ref Stat. Ref.* <https://doi.org/10.1002/9781118445112.stat07841>, Online 1–15 (2017).

## Acknowledgements

First author (DCS) acknowledge Science and Engineering Research Board (SERB) young scientist fellowship grant no. YSS/2015/001214 during the course of this study.

## Author contributions

D.C.S.: Conducted the experiments, Proteomic analysis, preparation of the first draft. D.J.: Analysis of the results. S.K.: Proteomic analysis. R.S.: Review & Editing. R.G.: Conceptualization, review & editing.

## Competing interests

The authors declare no competing interests.

## Additional information

**Supplementary information** Supplementary material is available for this paper at <https://doi.org/10.1038/s41598-019-56592-8>.

**Correspondence** and requests for materials should be addressed to R.G.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2019