



**HAL**  
open science

## Multi-metal tolerance of von Willebrand factor type D domain isolated from metal contaminated site by metatranscriptomics approach

Bharti Thakur, Rajiv Yadav, Laurent Vallon, Roland Marmeisse, Laurence Fraissinet-Tachet, M Sudhakara Reddy

### ► To cite this version:

Bharti Thakur, Rajiv Yadav, Laurent Vallon, Roland Marmeisse, Laurence Fraissinet-Tachet, et al.. Multi-metal tolerance of von Willebrand factor type D domain isolated from metal contaminated site by metatranscriptomics approach. *Science of the Total Environment*, 2019, 661, pp.432 - 440. 10.1016/j.scitotenv.2019.01.201 . hal-04804794

**HAL Id: hal-04804794**

**<https://hal.science/hal-04804794v1>**

Submitted on 4 Dec 2024

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

1 **Multi-metal tolerance of von Willebrand factor type D domain isolated from**  
2 **metal contaminated site by metatranscriptomics approach**

3 **Bharti Thakur<sup>a</sup>, Rajiv Yadav<sup>b</sup>, Laurent Vallon<sup>b</sup>, Roland Marmeisse<sup>b</sup>, Laurence**  
4 **Fraissinet-Tachet<sup>b</sup>, and M. Sudhakara Reddy<sup>a\*</sup>**

5 <sup>a</sup>Department of Biotechnology, Thapar Institute of Engineering & Technology, Patiala, Punjab,  
6 India, 147004.

7 <sup>b</sup>Ecologie Microbienne, UMR CNRS, UMR INRA, Université Claude Bernard Lyon 1  
8 Université de Lyon, F-69622 Villeurbanne, France.

9

10

11

12

13

14

15

16

17

18

19 \*Address for Correspondence

20 Dr. M. Sudhakara Reddy  
21 Department of Biotechnology  
22 Thapar Institute of Engineering & Technology  
23 Patiala 147 004, Punjab, India  
24 Email: [msreddy@thapar.edu](mailto:msreddy@thapar.edu)  
25 Tel: 91-175-2393350  
26 Fax: 91-175-2393005

27 **Abstract**

28 Environmental pollution through heavy metals is an upcoming universal problem that  
29 relentlessly endangers human health, biodiversity and ecosystems. Hence remediating these  
30 heavy metal pollutants from the environment by engineering soil microbiome through  
31 metatranscriptomics is befitting reply. In the present investigation, we have constructed size  
32 fractionated cDNA libraries from eukaryotic mRNA of cadmium (Cd) contaminated soil and  
33 screened for Cd tolerant genes by yeast complementation system by using Cd sensitive *ycf1<sup>Δ</sup>*  
34 mutant. We are reporting one of the transformants PLCe10 (from library C, 1-4 kb) with  
35 potential tolerance towards Cd toxicity (40  $\mu$ M-80  $\mu$ M). Sequence analysis of PLCe10  
36 transcript showed homology to von Willebrand factor type D domain (VWD) of vitellogenin-  
37 6 of *Ascaris suum* encoding 338 amino acids peptide. qPCR analysis revealed that PLCe10  
38 induced in presence of Cd (32 fold) and also accumulated maximum amount of Cd at 60  
39  $\mu$ M Cd. This cDNA was further tested for its tolerance against other heavy metals like copper  
40 (Cu), zinc (Zn) and cobalt (Co). Heterologous complementation assays of cDNA PLCe10  
41 showed a range of tolerance to Cu (150  $\mu$ M-500  $\mu$ M), Zn (10 mM-12 mM) and Co (2-4 mM).  
42 Results of the present study suggest that cDNA PLCe10 is one of the functional eukaryotic  
43 heavy metal tolerant genes present among the soil microbial community and could be exploited  
44 to rehabilitate metal contaminated sites.

45 **Key words:** Metatranscriptomics, Heavy metals, VWD, vitellogenin, heavy metal tolerance

## 46 **1. Introduction**

47 Soils are a large dump for various hazardous elements such as heavy metals which are  
48 released into the soil through numerous anthropogenic activities such as industrial and  
49 agricultural activities (Fernández et al., 2018). Heavy metals not only cause pollution by  
50 changing the physical and biochemical properties of soil but also represent severe  
51 environmental effects and are considered as health hazards (Mitchell et al., 2011; Tchounwou  
52 et al., 2012; Esteban-Vasallo et al., 2012). Conditions which affect the soil properties also  
53 affect the diversity and functioning of microorganisms living in that environment. Majority of  
54 these organisms possess a rich reservoir of genes which are involved in detoxification  
55 mechanisms such as chelation, excretion of metals or cellular compartmentalization (Cobbett  
56 and Goldsbrough, 2002; Clemens, 2006; Blindauer and Leszczyszyn, 2010). Discovering these  
57 genes and the mechanisms in which they are involved will not only elaborate our understanding  
58 towards the interaction between microorganisms and their environment but also be implicated  
59 in bio-industry or for bioremediation.

60 Eukaryotic microorganisms present in the soil carry out numerous biological processes  
61 with ecological significance and many of them were unknown to the science as they  
62 withstand the laboratory cultivation (Marmeisse et al., 2017). To explore true functional  
63 diversity and activities expressed by eukaryotic strata of soil *in situ* in response to various  
64 adverse factors, a metatranscriptomics approach has been practiced (Carvalhais et al., 2012;  
65 Marmeisse et al., 2017). Metatranscriptomics involves the construction of eukaryotic-specific  
66 environmental cDNA libraries through polyadenylated mRNA and screening them by  
67 heterologous complementation in an appropriate microbial host such as *Saccharomyces*  
68 *cerevisiae* (Grant et al., 2006; Bailly et al., 2007; Thakur et al., 2018). The systematic  
69 sequencing of these transcriptomes has led to the recognition of full-length genes involved in  
70 different biological processes from soils (Damon et al., 2012; Hesse et al., 2015; Žifčáková et

71 al., 2016; Mauchline et al., 2018), animal digestive tracts (Tartar et al., 2009; Qi et al., 2011;  
72 Dai et al., 2015), marine waters (Pearson et al., 2015; Zielinski et al., 2016) and extreme aquatic  
73 environments (Edgcomb et al., 2016; Puente-Sánchez et al., 2016). Metatranscriptomics  
74 approach has also been established to discover various ecological significant genes from  
75 different environments involved in numerous phenomena such as metal tolerance (Lehembre  
76 et al., 2013), organic matter degradation (Damon et al., 2011), cyanobacterial blooming in  
77 marine bodies (Berg et al., 2018), plant-microbe interaction (Crump et al., 2018), root  
78 microbiome (Gonzalez et al., 2018), saline tolerance in deserts microbiome (Ren et al., 2018).  
79 Studies have also been reported where metatranscriptomics is explored to remediate the  
80 contaminated sites such as degradation of polycyclic aromatic compounds like phenanthrene  
81 (de menezes et al., 2012; Suenaga et al., 2007) and in neutralizing the toxic heavy metals  
82 present in soil environment (Ziller et al., 2017).

83 In this study, eukaryotic cDNA libraries constructed from metal contaminated soil was  
84 screened to identify the genes involved in metal tolerance by yeast functional complementation  
85 assay. Screening of one of the libraries (library C) yielded several genes involved in metal  
86 tolerance including von Willebrand factor type D domain (VWD) of vitellogenin (Vg) protein  
87 (PLCe10). Vitellogenin is a multimeric egg yolk protein encoded by multiple Vg genes in a  
88 number of species that include insects, fish and frog (Tufail and Takeda, 2008) and provide a  
89 potential indicator for monitoring and assessing the toxic effect of heavy metals like Cd, Cu,  
90 and As (Hwang et al., 2009). Present study aims to characterize the cDNA PLCe10 and test its  
91 multi-metal tolerance characteristics by heterologous complementation.

## 92 **2. Materials and methods**

### 93 *2.1. Collection of soil samples and analysis*

94 Soil samples were collected from Agro-forestry land in Pierrelaye (49°1'45"N, 2°10'32"E)  
95 in northwest of Paris, France. This site was previously maize growing crop field which now

96 has been converted into an Agro-forestry site where Poplars are being cultivated. These sites  
97 became polluted with heavy metal due to irrigation with untreated wastewater for many years.  
98 Twenty soil cores were collected and equal volumes of each core was mixed together to form  
99 composite sample after sieving them through 2.0 mm mesh. The soil sample was analyzed for  
100 its pH (1:2 soil: water ratio, v/v), organic carbon (Walkley, 1947); total Phosphorus (Kitson  
101 and Mellon, 1944); available Phosphorous (Olsen et al., 1954) and total nitrogen (Piper, 1966).  
102 Soil samples were digested with HNO<sub>3</sub> and HClO<sub>4</sub> (3:1 v/v) and the metal content was  
103 determined by using inductively coupled plasma atomic emission spectroscopy (ICP-AES)  
104 (ARCOS, Simultaneous ICP Spectrometer, SPECTRO Analytical Instruments GmbH,  
105 Germany) (Moor et al., 2001).

## 106 2.2. Biological materials

107 In the present study, Cd-sensitive *ycf1<sup>Δ</sup>*, copper-sensitive *cup1<sup>Δ</sup>* (DTY4), Zn-sensitive  
108 *zrc1<sup>Δ</sup>* and cobalt-sensitive *cot1<sup>Δ</sup>* mutant strains of *S. cerevisiae* were used. The yeast mutant  
109 *ycf1<sup>Δ</sup>* lacks *YCF1* gene which is an ABC transporter encoding gene responsible for conferring  
110 Cd tolerance through the vacuolar transport of Cd conjugates (Li et al., 1997). Similarly, copper  
111 sensitive *cup1<sup>Δ</sup>* lacks the both the copies of copperthionein gene while its isogenic wild type  
112 strain DTY3 (*cup1<sup>S</sup>*) (*MAT $\alpha$* , *leu2-3*, *112his3<sup>Δ</sup>1*, *trp1-1*, *ura3-50*, *gall1*, *CUP1<sup>S</sup>*) carries single  
113 copy of copperthionein gene *cup1* (Longo et al., 1996). In case of *zrc1<sup>Δ</sup>*, *ZRC1* gene encoding  
114 a transporter protein that helps in sequestering of Zn into vacuoles is knocked out (Li and  
115 Kaplan, 1998). The hypersensitivity to Co is due to deletion of *COT1* gene that mediates the  
116 efflux of cobalt ions into the vacuole (Conklin et al., 1992). *ycf1<sup>Δ</sup>*, *cot1<sup>Δ</sup>* and *zrc1<sup>Δ</sup>* strains  
117 derived from the wild-type strain BY4741 (*MAT $\alpha$* , *his3<sup>Δ</sup>1*, *leu2 $\Delta$ 0*, *met15<sup>Δ</sup>0*, *ura3 $\Delta$ 0*) were  
118 used as a negative control for metal tolerance. All yeast cultures were maintained on Yeast  
119 Peptone Dextrose media (YPD) (Hi-media laboratories, India). For all bacterial  
120 transformations, One Shot® TOP10 Electrocomp™ *E. coli* (Thermofisher, USA) was used.

121 2.3. Screening of cDNA library C for cadmium tolerance genes

122 Total RNA was extracted from soil sample using the RNA PowerSoil<sup>®</sup> Total RNA  
123 Isolation Kit (Mo Bio laboratories, Carlsbad, CA) according to the manufacturer's instructions.  
124 cDNA synthesis and size fractionation of cDNA libraries were performed as described earlier  
125 (Yadav et al., 2014). Briefly, cDNA libraries were prepared by ligation of cDNAs between  
126 SfiIA and SfiIB sites downstream of the PGK1 promoter of *S. cerevisiae* present in yeast  
127 expression vector modified pFL61 (Minet et al., 1992). The largest fraction of cDNA library  
128 (Lib C) was screened for cadmium tolerant genes by functional complementation assay using  
129 *ycf1<sup>Δ</sup>* according to the method described in Thakur et al. (2018). For positive transformants,  
130 heterologous complementation assay was performed by drop out assay and growth assay. To  
131 perform drop out assay, yeast cells were grown in SD-ura broth until OD<sub>600</sub> = 1.0 and serial  
132 dilutions were made (0.1-0.0001). Five µl of each dilution was subsequently used to perform  
133 drop out assay on SD-ura with and without Cd (40 µM CdSO<sub>4</sub>) plates. The growth assay was  
134 performed to know the range of tolerance of tolerant cDNA by inoculating the respective yeast  
135 transformant in 20 ml medium with log phase culture in range of 40-80 µM CdSO<sub>4</sub> and growth  
136 was recorded after 48hrs. One of the transformants, PLCe10 was chosen for further studies,  
137 based on its tolerance towards higher Cd concentrations.

138 2.4. Sequence analysis

139 Plasmid DNA was extracted from *ycf1<sup>Δ</sup>/PLCe10* cells by using Zymoprep<sup>™</sup> Yeast  
140 Plasmid Miniprep II (Zymo Research, USA) and electro-transformed (3-4 µl) into One  
141 Shot<sup>®</sup> TOP10 Electrocomp<sup>™</sup> *E. coli* cells (Thermofisher, USA). The insert was sequenced  
142 with vector specific primers NF and NR (Table 1). BLASTX analysis was performed to find  
143 homologous sequences for PLCe10 via NCBI web server ([http://blast.ncbi.nlm.nih.gov/Blast.](http://blast.ncbi.nlm.nih.gov/Blast.cgi)  
144 [cgi](http://blast.ncbi.nlm.nih.gov/Blast.cgi)). The resulting polypeptide sequences were aligned by multiple sequence alignment using  
145 Clustal Omega (<http://www.ebi.ac.uk/Tools/msa/clustalo/>). Potential O- and N-linked

146 glycosylation sites were determined using NetOglyc 3.1 ([http://www.cbs.dtu.dk/services/](http://www.cbs.dtu.dk/services/NetOglyc/)  
147 [NetOglyc/](http://www.cbs.dtu.dk/services/NetOglyc/)) and NetNGlyc 1.0 server (<http://www.cbs.dtu.dk/services/NetNGlyc/>) and  
148 phosphorylation sites by NetPhos 2.0 server (<http://www.cbs.dtu.dk/services/NetPhos/>). The  
149 sequence of PLCe10 cDNA was deposited at NCBI database under the accession number  
150 MF770726.

### 151 2.5. qPCR analysis and metal uptake

152 The induction level of PLCe10 in response to Cd stress was quantified by qPCR analysis.  
153 The *ycf1<sup>Δ</sup>* cells carrying PLCe10 were grown in SD-ura broth supplemented with different  
154 concentrations of Cd (0, 40, 60 and 80  $\mu$ M) for 48 hrs at 30 °C. Cells were centrifuged at 3000  
155 rpm for 5 mins and the pellet was ground with liquid nitrogen. Total RNA was extracted by  
156 using Trizol reagent followed by use of RNase-free DNase I (Thermo Scientific, USA) to  
157 remove genomic DNA contamination. The purity and integrity of RNA was determined by  
158 NanoDrop<sup>™</sup> 1000 (Thermo Scientific, USA) and formaldehyde agarose gel electrophoresis.  
159 First strand cDNA was synthesized from total RNA by using PrimeScript<sup>™</sup> first strand cDNA  
160 synthesis kit (Takara, Japan) as per the instruction manual. Gene expression analysis of Cd  
161 tolerant cDNA PLCe10 was performed by using SYBR<sup>®</sup> Premix Ex Taq<sup>™</sup> (Tli RNase H Plus)  
162 (Takara, Japan) as directed in user's manual in a Mastercycler<sup>®</sup> ep realplex system (Eppendorf  
163 AG, Hamburg, Germany) with a final volume of 20  $\mu$ l (100 times diluted RT cDNA 6.8  $\mu$ l, 0.2  
164  $\mu$ M forward primer, 0.2  $\mu$ M reverse primer (Table 1), 2x SYBR Green mix 10 $\mu$ l and H<sub>2</sub>O final  
165 volume make up). Amplification programme used was: 95 °C for 2 mins (1 cycle), 95 °C for  
166 15 sec, 53 °C for 15 sec and 68 °C for 20 sec (40 cycles). The amplification efficiency (*E*) was  
167 calculated in accordance to the equation  $E = [10^{(-1/\text{slope})}] - 1$  and efficiency values lies between  
168 0.94-1.27. These efficiency values were further used to compute Ct<sub>1</sub> values by using equation  
169  $Ct_1 = C_{tex} [\log(1+E)/\log 2]$ . The primers (Table 1) were designed for reference genes from *ACT*  
170 (Actin; SGD:S000001855), *TAF10* (TATA binding protein-associated factor;



171 SGD:S000002574) and *TFC1* (Transcription factor class C; SGD:S000000327) of *S. cerevisiae*  
172 (Teste et al., 2009). Analysis via Normfinder suggested that most stable expression and  
173 minimum stability value under Cd stress recorded was for *TFC1*. Hence, *TFC1* was used as  
174 reference genes to calculate comparative expression level of genes by using the formula  $2^{-\Delta\Delta C_T}$   
175 (Livak and Schmittgen, 2001). The  $-\Delta\Delta C_T$  is calculated by subtracting the normalised Ct value  
176 for test data from normalised Ct value for corresponding control. For every measurement, RNA  
177 was extracted from three independent biological samples (experimental replicates) and for each  
178 experimental replicate, three technical replicates were also performed.

179 The Cd accumulation by *ycf1<sup>Δ</sup>/PLCe10* cells were determined by ICP-MS (Element XR,  
180 Thermo Fisher Scientific, Germany) analysis by growing these cells at different concentrations  
181 of Cd (40  $\mu$ M to 80  $\mu$ M CdSO<sub>4</sub>). Cells were harvested and acid digested 1 gm of cell pellet  
182 with HNO<sub>3</sub> and HClO<sub>4</sub> (3:1 v/v). Acid digested samples (final volume of 50 ml with H<sub>2</sub>O) were  
183 analysed by ICP-MS. The wild type strain BY4741 transformed with empty vector was used  
184 as positive control.

#### 185 2.6. Tolerance of *PLCe10* for different metals

186 The plasmid pFL61 containing cDNA *PLCe10* was transformed into yeast mutants  
187 hypersensitive to specific heavy metal like Cu (*cup1<sup>Δ</sup>*), Zn (*zrc1<sup>Δ</sup>*) and Co (*cot1<sup>Δ</sup>*) and their  
188 respective wild type strains with empty vector. The transformants were regenerated on SD-ura  
189 media containing specific metal i.e. Cu 150  $\mu$ M, Zn 10 mM and Co 2 mM. In parallel  
190 experiments for studying growth under the influence of metal, each flasks containing 20 ml of  
191 SD-ura medium with different range of metal concentrations of Cu (150  $\mu$ M-500  $\mu$ M), Zn (10  
192 mM-12 mM) and Co (2 mM-4 mM) were inoculated with transformed *cup1<sup>Δ</sup>*, *cot1<sup>Δ</sup>*, *zrc1<sup>Δ</sup>* cells  
193 and incubated at 30°C at 220 rpm for 48 hrs. In a similar way, yeast transformant cultures  
194 (*cup1<sup>Δ</sup>*, *zrc1<sup>Δ</sup>* and *cot1<sup>Δ</sup>*) were grown in 50 ml of SD-ura supplemented with range of different

195 metals i.e. Cu (150-500  $\mu$ M), Zn (10-12 mM) and Co (2-4 mM) and the uptake of metals were  
196 determined by ICP-MS as described above.

### 197 **3. Results and Discussion**

#### 198 *3.1. Soil analysis*

199 The soil collected from Pierrelaye region is sandy luvisol with a pH of 7.1. The other  
200 chemical properties of the soil were: organic carbon 1.6%; total P 291 mg/kg; available P 14.2  
201 mg/kg and total nitrogen 0.12%. The Cd content was 2.5 mg/kg, Cu 64 mg/kg and Zn 385  
202 mg/kg of the soil.

#### 203 *3.2. Cadmium tolerance genes*

204 One of the main challenges in cDNA library construction is to obtain a high number of  
205 clones having inserts more than 1.0 kb size (Yadav et al., 2014). In the present study, the cDNA  
206 quantities were slightly increased by amplifying through three PCR cycles without affecting  
207 significantly their average length. Then the cDNAs were separated by bi-dimensional agarose  
208 gel electrophoresis to separate them and minimize cross-contamination of size fractions. Three  
209 different size fractions such as fraction A (0.1–0.5 kb), fraction B (0.5–1 kb) and fraction C (1–  
210 4 kb) were recovered and the respective cDNA libraries were constructed as described in Yadav  
211 et al. (2014). Prior to library construction, the fractions were tested for the absence of cross  
212 contaminations by performing PCR amplification of four eukaryotic genes of different length  
213 classes (Yadav et al., 2014). In the present study, library C (1-4 kb) was screened for presence  
214 of metal tolerant gene present in soil environment. The transformants from cDNA library C  
215 were selected based on their ability to rescue the Cd-hypersensitive mutant strain *ycf1<sup>A</sup>*. One of  
216 the transformants, *ycf1<sup>A</sup>PLCe10*, which showed high tolerance to Cd while performing drop  
217 test, was selected for further studies (Fig. 1a). The *ycf1<sup>A</sup>PLCe10* showed significantly higher  
218 growth in all the concentrations compared to the wild type strain transformed with empty  
219 vector. With increase in Cd concentration, reduction in the growth was observed both in

220 *ycf1*<sup>4</sup>PLCe10 and wild type strain (Fig. 1b). The metabolic and ecological roles of  
221 microorganisms present in various adverse environments including metal contaminated site  
222 could be studied by metatranscriptomics approach. Lehembre et al. (2013) reported different  
223 categories of genes such as metallothioneins (MTs), cysteine rich protein (CRPs) responsible  
224 for Cd tolerance and *BOLA* genes responsible for Zn, Mn and Co tolerance by using soil  
225 metatranscriptomics. While exploring metatranscriptomics for novel genes, Ziller et al. (2017)  
226 reported metal binding properties of five new CRPs (environmental MTs) conferring tolerance  
227 to heavy metals like Zn and Cd. Similarly, Thakur et al. (2018) isolated environmental  
228 transcript PLBe1 encoding an ubiquitin fusion protein responsible for Cd tolerance.

### 229 3.3. Sequence analysis

230 The BLASTX analysis with the deduced nucleotide sequences of PLCe10 revealed 30%  
231 identity (50% positives) to von Willebrand factor type D (VWD) of vitellogenin-6 of nematode  
232 *Ascaris suum* with the open reading frame of 1017 bp, encoding 338 amino acids. The predicted  
233 molecular mass of peptide was 38.96 KDa (average mass) with pI of 4.77. The deduced peptide  
234 showed homology with amino acid sequence of *A. suum* from 1428-1719 amino acids denoted  
235 C-terminal of the Vg-6 which is an egg yolk protein. It had been reported that in *Paracyclops*  
236 *nana*, C-terminal region of the Vg1 from amino acid sequence 1571-1739 and the Vg2 i.e. from  
237 1481-1658 amino acids codes for von Willebrand factor type D domain (VWD). Hwang et al.  
238 (2010) reported that the BLAST analysis with the inferred amino acid sequences of  
239 *Paracyclops nana* Vg1 and Vg2 had shown that in amino acid residues strongly matched  
240 with the intertidal copepod Vg1 (35% identities/56% positives) and *Lepeophtheirus salmonis*  
241 Vg2 (31% identities/52% positives), respectively. Multiple sequence alignment of homologous  
242 sequences showed that the 1017 bp gene belongs to protein family smart00216 which is also  
243 named as VWD. The deduced polypeptide identified as VWD like protein, contained  
244 characteristic phosphorylation sites at 32 amino acid sequences and 14 potential motifs for O-

245 linked glycosylation sites were also found while no N-linked glycosylation sites (Fig. 2). Finn  
246 (2007) observed that before exporting Vg from hepatocytes, phosphorylation of precursor  
247 protein takes place in the Golgi bodies and the presence of phosphorylated serine residues in  
248 the Vg polypeptide are assumed to avert premature degradation and assist the solubility of Vg  
249 molecules in the blood. Various phosphorylation sites within the predicted VWD like  
250 polypeptide sequence indicate that the wide majority of serine residues present in the  
251 polypeptide of VWD like domain may be phosphorylated along with threonine and tyrosine.  
252 In addition to this, comparison of VWD like polypeptide to other vitellogenin showed a high  
253 degree of conservation; contain two RXXR consensus cleavage sites each at C-terminal and  
254 N-terminal of the polypeptide, a well characterised the GL/ICG motif and conserved  
255 polycysteine residues at C-terminal (Fig. 3).

256 Tufail et al. (2010) characterised Vg from insect *Nilaparvata lugens* and found high degree  
257 of conservation for the motif GL/ICG. In a study, Shu et al. (2009) aligned Vg of *Spodoptera*  
258 *litura* with Vg of other species and characterised it by presence of highly conserved cleavage  
259 signal RXXR at N terminal, C-terminal the GL/ICG motif followed by cysteine residues.  
260 Similarly in a previous research, it was concluded that in C-terminal of Vg, VWD contain a  
261 signature amino acid sequence GL/ICG followed by nine cysteine residues (Tufail and Takeda,  
262 2008). Salmela et al. (2016) characterised three Vg like proteins and found that all of them  
263 share structural similarities with Vg but each of them also has unique gene/ protein architecture  
264 and distinct pattern of positive selection, which are indicative of divergence of functions.  
265 Presence of hexameric polyadenylation signal (PAS) AATAAA at 12 bp upstream the poly-A  
266 tail in ORF PLCe10, also characterised this cDNA as ORF with a potential for coding a protein  
267 as described by Beaudoin et al. (2000). From all the findings of bioinformatic analysis, we  
268 assumed that the cDNA sequence PLCe10 conferring tolerance to Cd metal obtained in this  
269 study encode for VWD like domain of putative protein Vg. Jubeaux et al. (2012) clearly

270 demonstrated the benefit of moderate phylogenetic conservation of Vg like protein such as Vg  
271 proteotypic peptides, to directly exploit them in developing biomarkers to identify more or less  
272 close phylogenetically related un-sequenced species.

### 273 3.4. qPCR analysis metal uptake

274 Induction of cDNA PLCe10 encoding VWD like protein in the presence of Cd metal was  
275 studied. Induction levels of PLCe10 increased significantly due to Cd stress and maximum  
276 mRNA accumulation was observed at 60  $\mu$ M of CdSO<sub>4</sub>, which was 32 fold higher than control  
277 (Fig. 4a). Similar findings were reported for induction of metallothioneins in fungi on exposure  
278 to heavy metals and observed that expression of different metallothioneins helps in providing  
279 tolerance towards different heavy metals (Ramesh et al., 2009; Reddy et al., 2016; Kalsotra et  
280 al., 2018). The *ycf1<sup>Δ</sup>*PLCe10 cells accumulated different levels of Cd when grown in presence  
281 of various Cd concentrations. It has been observed that the amount of Cd metal accumulated  
282 increased up to 60  $\mu$ M of Cd concentration and decreased at 80  $\mu$ M Cd (Fig. 4b).

### 283 3.5. Multi-metal tolerance of PLCe10

284 The cDNA PLCe10 was tested for its tolerance to other heavy metals such as Cu, Zn and  
285 Co. As shown in Fig. 5a, Cu sensitive mutant *cup1<sup>Δ</sup>* with empty pFL61 was sensitive to 150  
286  $\mu$ M CuSO<sub>4</sub> while *cup1<sup>Δ</sup>* cells carrying plasmid PLCe10 (*cup1<sup>Δ</sup>*PLCe10) were showing  
287 tolerance to a recognisable level in comparison to *cup1<sup>s</sup>*pFL61 which carries single copy of  
288 *CUP1* gene only. Likewise, *zrc1<sup>Δ</sup>* cells carrying PLCe10 plasmid were tolerant at 10 mM ZnCl<sub>2</sub>  
289 whereas this concentration was toxic for *zrc1<sup>Δ</sup>*pFL61 as shown in Fig. 5b. Similarly, 2 mM  
290 CoCl<sub>2</sub> inhibited the growth of *cot1<sup>Δ</sup>*pFL61 where as *cot1<sup>Δ</sup>* cells carrying plasmid PLCe10  
291 (*cot1<sup>Δ</sup>*PLCe10) were showing tolerance to Co as in case of wild strain BY4741pFL61 (Fig. 5c).

292 The growth assays of different transformants were showed positive results on different  
293 range of concentration of metals. When screened for Cu, it was observed that *cup1<sup>Δ</sup>*PLCe10  
294 cells were able to grow up to 500  $\mu$ M compared to mutant strain *cup1<sup>Δ</sup>*pFL61 and wild type

295 strain *cup1<sup>s</sup>*pFL61 (Fig. 6a). In context to Cu, it was reported that estrogen induced vitellogenin  
296 was tolerant to oxidation reactions induced by Cu and chelate low concentrations of Cu ion.  
297 Antioxidant property of vitellogenin protected the copper-stimulated oxidation of very low  
298 density lipoprotein (VLDL). Vitellogenin believed to be served as transition metals-binding  
299 lipoprotein which results into extensive depressing of free-radical reactions in the oocytes  
300 (Ando and Yanagida, 1999). While studying tolerance for Zn metal tolerance, it was found that  
301 transformant has shown more tolerance towards Zn (10 mM to 12 mM ZnSO<sub>4</sub>) in comparison  
302 to both wild strain BY4741pFL61 and mutant strain *zrc1<sup>A</sup>*pFL61 (Fig. 6b). Vitellogenin  
303 mediates the transfer of Zn from the liver to the maturing oocytes in the turkey hen, ultimately  
304 resulting in Zn deposition into yolk of the newly formed egg (Richards, 1989). It was also  
305 reported that Vg is a macromolecule phospholipid glycoprotein containing Ca and Zn ligands  
306 (Wallace, 1985; Montorzi et al., 1994, Denslow et al., 1999). Furthermore it was reported that  
307 among honeybee workers and queens, Vg plays role as a Zn carrier that protects them from  
308 oxidative stress (Amdam et al., 2004; Seehuus et al., 2006) and contributes to queen longevity  
309 (Corona et al., 2007). Similarly in cobalt, the gene provided tolerance to mutant yeast from  
310 range of 2 mM to 4 mM. Increase in tolerance was observed with the increase in range of Co  
311 concentrations (Fig. 6c). Henceforth, this study showed that gene PLCe10 was responsible for  
312 providing tolerance to all the four metals. Concurrently, while estimating the accumulated  
313 metal by yeast mutants transformed with cDNA PLCe10, it has been observed that cDNA  
314 PLCe10 is capable of accumulating a higher range of metal (Cu, Zn and Co) in comparison to  
315 wild strain BY4741 transformed with empty vector pFL61. For Cu, the level of metal  
316 accumulation increased with increase in concentration as shown in Fig. 6d, while for Zn and  
317 Co metal, the level of accumulation increases up to 11 mM of ZnCl<sub>2</sub> and 3 mM CoCl<sub>2</sub>  
318 respectively and decreases with further increase in Zn or Co concentration (Fig. 6e, 6f). Such  
319 variation in order of tolerance and accumulation of heavy metals could potentially be an

320 explanation for distinct metal-binding abilities of the deduced VWD like protein. In general,  
321 increased growth and metal accumulation was observed in mutant yeast cells carrying PLCe10  
322 compared to the respective wild type strains. The increase in growth might be due to protection  
323 of PLCe10 to the yeast mutants under metal stress and the accumulation of metal in the yeast  
324 strains also might to due increased growth.

325 In heavy metal contaminated aquatic environment, induction of Vg among aquatic animals  
326 on exposure to heavy metals and endocrine disruption chemicals has been used as a biomarker  
327 (Hwang et al., 2010, Lee-Pow et al., 2016). As on bioinformatics analysis of VWD like protein,  
328 it was found that the cDNA screened has shown GL/ICG conserved domain (Fig. 2). The  
329 presence of GL/ICG motif followed by number of cysteine residues at conserved locations was  
330 entitled as a general trait among various invertebrate and vertebrate species (Lee, 2000).  
331 Consequently, there could be a possibility to gain benefit of this peptide motif conservation to  
332 develop biomarkers applicable for groups of phylogenetically related species as these days  
333 mass spectrophotometry-based procedures necessitate small peptides i.e. from 5-20 amino  
334 acids. Hence, these findings reported VWD like protein as a potential member of metal tolerant  
335 gene family of eukaryotic community which increases the possibility of using such peptides as  
336 biomarker for metal contamination.

#### 337 **4. Conclusions**

338 The present study results conclude that the eukaryotic transcript PLCe10 have potential to  
339 tolerate the toxicity of different metals in the environment. Functional metatranscriptomic  
340 approach is suitable to characterize novel genes like VWD like protein whose over-expression  
341 restored tolerance to all the four heavy metals in hypersensitive yeast mutants. This study  
342 strongly supports the existence of some unknown mechanisms and biomolecules that play vital  
343 role in adaptation of biodiversity of an ecosystem against metal contaminated environment.  
344 Metatranscriptomics approach has major contribution in providing information about various

345 unknown phenomena occurring in soil and revealed high degree of diversity. Most of the work  
346 on vitellogenin is contributed to aquatic oviparous vertebrates and invertebrates, while in  
347 present study VWD like protein screened through metatranscriptomic cDNA library of metal  
348 polluted soil environment is novel finding of this approach. In this respect, VWD like protein  
349 could be used to probe gene related processes and to assess their involvement in the overall  
350 capacity of organisms to survive in heavy metal enriched environment. This gene could be  
351 developed as biomarker to determine heavy metals in the environment principally in soil and  
352 to exploit the levels of tolerance or resistance among the species of concern for the  
353 bioremediation of polluted lands.

#### 354 **Acknowledgements**

355 Authors are thankful to Indo-French Centre for the Promotion of Advanced Research for  
356 financial assistance (4709-1)

#### 357 **References**

- 358 Amdam GV, Norberg K, Fondrk MK, Page RE. 2004. Reproductive ground plan may mediate  
359 colony-level selection effects on individual foraging behavior in honey bees. *Proc Natl*  
360 *Acad Sci USA* 101, 11350-11355.
- 361 Ando S, Yanagida K. 1999. Susceptibility to oxidation of copper induced plasma lipoproteins  
362 from Japanese eel, protective effect of vitellogenin on the oxidation of very low density  
363 lipoprotein. *Comp Biochem Physiol* 123: 1-7.
- 364 Bailly J, Fraissinet-Tachet L, Verner MC, Debaud JC, Lemaire M, Wésolowski-Louvel M,  
365 Marmeisse R. 2007. Soil eukaryotic functional diversity, a metatranscriptomic approach.  
366 *ISME J* 1: 632.
- 367 Beaudoin E, Freier S, Wyatt JR, Claverie JM, Gautheret D. 2000. Patterns of variant  
368 polyadenylation signal usage in human genes. *Genome Res* 10: 1001-1010.
- 369 Berg C, Dupont CL, Asplund-Samuelsson J, Celepli NA, Eiler A, Allen AE, Ekman M,  
370 Bergman B, Ininbergs K. 2018. Dissection of microbial community functions during a  
371 cyanobacterial bloom in the Baltic Sea via metatranscriptomics. *Front Mar Sci*, 5: p.55.
- 372 Blindauer CA, Leszczyszyn OI. 2010. Metallothioneins: unparalleled diversity in structures  
373 and functions for metal ion homeostasis and more. *Nat Prod Rep* 27: 720-741.



374 Carvalhais LC, Dennis PG, Tyson GW, Schenk PM. 2012. Application of metatranscriptomics  
375 to soil environments. *J Microbiol Methods* 91: 246-251.

376 Clemens S. 2006. Toxic metal accumulation, responses to exposure and mechanisms of  
377 tolerance in plants. *Biochimie* 88: 1707-1719.

378 Cobbett C, Goldsbrough P. 2002. Phytochelatins and metallothioneins: roles in heavy metal  
379 detoxification and homeostasis. *Annu Rev Plant Biol*, 53: 159-182.

380 Conklin DS, McMaster JA, Culbertson MR, Kung C. 1992. COT1, a gene involved in cobalt  
381 accumulation in *Saccharomyces cerevisiae*. *Mol Cell Biol* 12: 3678-3688.

382 Corona M, Velarde RA, Remolina S, Moran-Lauter A, Wang Y, Hughes KA, Robinson GE.  
383 2007. Vitellogenin, juvenile hormone, insulin signaling, and queen honey bee longevity.  
384 *Pro Natl Acad Sci USA* 104: 7128-7133.

385 Crump BC, Wojahn JM, Tomas F, Mueller RS. 2018. Metatranscriptomics and Amplicon  
386 Sequencing Reveal Mutualisms in Seagrass Microbiomes. *Front Microbiol* 9: 388.

387 Dai X, Tian Y, Li J, Su X, Wang X, Zhao S, Wang J. 2015. Metatranscriptomic analyses of  
388 plant cell wall polysaccharide degradation by microorganisms in the cow rumen. *Appl*  
389 *Environ Microbiol* 81: 1375-1386.

390 Damon C, Lehembre F, Oger-Desfeux C, Luis P, Ranger J, Fraissinet-Tachet L, Marmeisse R.  
391 2012. Metatranscriptomics reveals the diversity of genes expressed by eukaryotes in  
392 forest soils. *PLoS one* 7: e28967.

393 Damon C, Vallon L, Zimmermann S, Haider MZ, Galeote V, Dequin S, Marmeisse R. 2011.  
394 A novel fungal family of oligopeptide transporters identified by functional  
395 metatranscriptomics of soil eukaryotes. *ISME J* 5: 1871- 1880.

396 de Menezes A, Clipson N, Doyle E. 2012. Comparative metatranscriptomics reveals  
397 widespread community responses during phenanthrene degradation in soil. *Environ*  
398 *Microbiol* 14: 2577-2588.

399 Denslow ND, Chow MC, Kroll KJ, Green L. 1999. Vitellogenin as a biomarker of exposure  
400 for estrogen or estrogen mimics. *Ecotoxicology* 8: 385-398.

401 Edgcomb VP, Pachiadaki MG, Mara P, Kormas KA, Leadbetter ER, Bernhard JM. 2016. Gene  
402 expression profiling of microbial activities and interactions in sediments under haloclines  
403 of E. Mediterranean deep hypersaline anoxic basins. *ISME J* 10: 2643-2657.

404 Esteban-Vasallo MD, Aragonés N, Pollan M, López- Abente G, Perez-Gomez B. 2012.  
405 Mercury, cadmium and lead levels in human placenta: a systematic review. *Environ.*  
406 *Health Perspect* 120: 1369-1377.

407 Fernández S, Cotos-Yáñez T, Roca-Pardiñas J, Ordóñez C. 2018. Geographically Weighted  
408 Principal Components Analysis to assess diffuse pollution sources of soil heavy metal:  
409 Application to rough mountain areas in Northwest Spain. *Geoderma* 311: 120-129.

410 Finn RN. 2007. Vertebrate yolk complexes and the functional implications of phosvitins and  
411 other subdomains in vitellogenins. *Biol. Reprod.* 76, 926-935.

412 Gonzalez E, Pitre FE, Pagé AP, Marleau J, Nissim WG, St-Arnaud M, Labrecque M, Joly S,  
413 Yergeau E, Brereton NJB. 2018. Trees, fungi and bacteria: tripartite metatranscriptomics  
414 of a root microbiome responding to soil contamination. *Microbiome* 6: 53.

415 Grant S, Grant WD, Cowan DA, Jones BE, Ma Y, Ventosa A, Heaphy, S. 2006. Identification  
416 of eukaryotic open reading frames in metagenomic cDNA libraries made from  
417 environmental samples. *Appl Environ Microbiol* 72: 135-143.

418 Hesse CN, Mueller RC, Vuyisich M, Gallegos-Graves LV, Gleasner CD, Zak DR, Kuske CR.  
419 2015. Forest floor community metatranscriptomes identify fungal and bacterial responses  
420 to N deposition in two maple forests. *Front Microbiol* 6: 337-406.

421 Hwang DS, Lee KW, Han J, Park HG, Lee J, Lee YM, Lee JS. 2010. Molecular characterization  
422 and expression of vitellogenin (Vg) genes from the cyclopoid copepod, *Paracyclops*  
423 *nana* exposed to heavy metals. *Comp Biochem Physiol* 151: 360-368.

424 Hwang DS, Lee KW, Lee JS. 2009. Cloning and expression of vitellogenin 2 gene from the  
425 intertidal copepod *Tigriopus japonicus*. *Ann. N. Y. Acad. Sci.* 1163, 417- 420.

426 Jubeaux G, Audourd-Combe F, Simon R, Tutundijn R, Salvador A, Geffard O, Chaumot A.  
427 2012. Vitellogenin- like proteins among invertebrate species diversity: Potential of  
428 proteomic mass spectrometry for biomarker development. *Environ Sci Technol* 46: 6315-  
429 6323.

430 Kalsotra T, Khullar S, Agnihotri R, Reddy MS. 2018. Metal induction of two metallothionein  
431 genes in the ectomycorrhizal fungus *Suillus himalayensis* and their role in metal  
432 tolerance. *Microbiol* 164: 868-876.

433 Kitson RE, Mellon MG. 1944. Colorimetric determination of phosphorus as  
434 molybdovanadophosphoric acid. *Ind Eng Chem* 16: 379-383.

435 Livak KJ, Schmittgen TD. 2001. Analysis of relative gene expression data using real-time  
436 quantitative PCR and the  $2^{-\Delta\Delta C_T}$  method. *Methods* 25: 402-408.

437 Lee JM, Nishimori Y, Hatakeyama M, Bae TW, Oishi K. 2000. Vitellogenin of the cicada  
438 *Graptopsaltria nigrofuscata* (Homoptera), Analysis of its primary structure. *Insect*  
439 *Biochem Mol Biol* 30: 1-7.

440 Lee-Pow CS, Yost EE, Aday DD, Kullman SW. 2016. Sharing the Roles: An Assessment of  
441 Japanese Medaka Estrogen Receptors in Vitellogenin Induction. *Environ Sci Technol* 50:  
442 8886-8895.

443 Lehembre F, Doillon D, David E, Perrotto S, Baude J, Foulon J, Harfouche L, Vallon L, Poulain  
444 J, Da Silva C, Wincker P, Oger-Desfeux C, Richaud P, Colpaert V, Chalot M, Fraissinet-  
445 Tachet L, Blaudez D, Marmeisse R. 2013. Soil metatranscriptomics for mining  
446 eukaryotic heavy metal resistance genes. *Environ Microbiol* 15: 2829-2840.

447 Li L, Kaplan J. 1998. Defects in the yeast high affinity iron transport system result in increased  
448 metal sensitivity because of the increased expression of transporters with a broad  
449 transition metal specificity. *J Biol Chem* 273: 22181-22187.

450 Li ZS, Lu YP, Zhen, RG, Szcypka M, Thiele DJ, Rea PA. 1997. A new pathway for vacuolar  
451 cadmium sequestration in *Saccharomyces cerevisiae*, YCF1 $\Delta$ -catalyzed transport of  
452 bis(glutathionato) cadmium. *Proc Natl Acad Sci USA* 94: 42-47.

453 Longo VD, Gralla EB, Valentine JS. 1996. Superoxide dismutase activity is essential for  
454 stationary phase survival in *Saccharomyces cerevisiae*. *J Biol Chem* 271: 12275–12280.

455 Marmeisse R, Kellner H, Fraissinet-Tachet L, Luis P, 2017. Discovering Protein-Coding Genes  
456 from the Environment: Time for the Eukaryotes? *Trends Biotechnol* 35: 824-835.

457 Mauchline TH, Hayat R, Clark IM, Hirsch PR. 2018. Old meets new: most probable number  
458 validation of metagenomic and metatranscriptomic datasets in soil. *Lett Appl*  
459 *Microbiol* 66: 14-18.

460 Minet M, Dufour ME, Lacroute F. 1992. Complementation of *Saccharomyces cerevisiae*  
461 auxotrophic mutants by *Arabidopsis thaliana* cDNAs. *Curr Genet* 2: 417-422.

462 Mitchell E, Frisbie S, Sarkar B. 2011. Exposure to multiple metals from groundwater-a global  
463 crisis: geology, climate. *Metallomics* 3: 874-908.

464 Montorzi M, Falchuk KH, Vallee BL. 1994. *Xenopus laevis* vitellogenin is a zinc protein.  
465 *Biochem Biophys Res Commun* 200: 1407-1413.

466 Moor C, Lymberopoulou T, Dietrich VJ. 2001. Determination of heavy metals in soils,  
467 sediments and geological materials by ICP-AES and ICP-MS. *Microchim Acta* 136: 123-  
468 128.

469 Olsen SR, Cole CV, Watanabe FS, Dean LA. 1954. Estimation of available phosphorus in soils  
470 by extraction with sodium bicarbonate. U.S. Government printing office. Washington  
471 (DC): US Department of Agriculture. Circular No. 939.

472 Pearson GA, Lago-Leston A, Cánovas F, Cox CJ, Verret F, Lasternas S, Serrão EA. 2015.  
473 Metatranscriptomes reveal functional variation in diatom communities from the  
474 Antarctic Peninsula. *ISME J* 9: 2275-2289.

475 Piper CS. 1966. Soil and plant analysis: a laboratory manual of methods for the examination  
476 of soils and the determination of the inorganic constituents of plants. Bombay: Hans  
477 Publications.

478 Puente-Sánchez F, Olsson S, Gómez-Rodríguez M, Souza-Egipsy V, Altamirano-Jeschke M,  
479 Amils R, Aguilera A. 2016. Solar radiation stress in natural acidophilic biofilms of  
480 *Euglena mutabilis* revealed by metatranscriptomics and PAM fluorometry. *Protist* 167:  
481 67-81.

482 Qi M, Wang P, O'Toole N, Barboza PS, Ungerfeld E, Leigh MB, Forster RJ. 2011. Snapshot  
483 of the eukaryotic gene expression in muskoxen rumen - A metatranscriptomic  
484 approach. *PloS one* 6: e20521.

485 Ramesh G, Podila GK, Gay G, Marmeisse R, Reddy MS. 2009. Different patterns of regulation  
486 for the copper and cadmium metallothioneins of the ectomycorrhizal fungus *Hebeloma*  
487 *cylindrosporum*. *Appl Environ Microbiol* 75: 2266-2274.

488 Reddy MS, Kour M, Aggarwal, S, Ahuja S, Marmeisse R, Fraissinet-Tachet, L. 2016. Metal  
489 induction of a *Pisolithus albus* metallothionein and its potential involvement in heavy  
490 metal tolerance during mycorrhizal symbiosis. *Environ Microbiol* 18: 2446-2454.

491 Ren M, Zhang Z, Wang X, Zhou Z, Chen D, Zeng H, Zhao S, Chen L, Hu Y, Zhang C, Liang  
492 Y. 2018. Diversity and Contributions to Nitrogen Cycling and Carbon Fixation of Soil  
493 Salinity Shaped Microbial Communities in Tarim Basin. *Front Microbiol* 9: 431.

494 Richards MP. 1989. Influence of egg production on zinc, copper and iron metabolism in the  
495 turkey hen (*Meleagris gallopavo*). *Comp Biochem Physiol* 93A: 811-817.

496 Salmela H, Stark T, Stucki D, Fuchs S, Freitak D, Dey A, Kent CF, Zayed A, Dhayegude K,  
497 Hokkanen H, Sundstorm L. 2016. Ancient Duplications have led to functional divergence  
498 of Vitellogenin- like genes potentially involved in inflammation and oxidative stress in  
499 honey bees. *Genome Biol Evol* 8: 495-506.

500 Seehuus SC, Norberg K, Gimsa U, Krekling T, Amdam GV. 2006. Reproductive protein  
501 protects functionally sterile honey bee workers from oxidative stress. *Proc Natl Acad Sci*  
502 *USA* 103: 962-967.

503 Shu Y, Zhou J, Tang W, Lu K, Zhou Q, Zhang G. 2009. Molecular characterization and  
504 expression pattern of *Spodoptera litura* (Lepidoptera, Noctuidae) vitellogenin, and its  
505 response to lead stress. *J Insect Physiol* 55: 608-616.

506 Suenaga H, Ohnuki T, Miyazaki K. 2007. Functional screening of a metagenomic library for  
507 genes involved in microbial degradation of aromatic compounds. *Environ Microbiol*, 9:  
508 2289-2297.

509 Tartar A, Wheeler MM, Zhou X, Coy MR, Boucias, DG, Scharf ME. 2009. Parallel  
510 metatranscriptome analyses of host and symbiont gene expression in the gut of the  
511 termite *Reticulitermes flavipes*. *Biotechnol Biofuels* 2: 1-19.

512 Teste MA, Duquenne M, François, JM, Parrou JL. 2009. Validation of reference genes for  
513 quantitative expression analysis by real-time RT-PCR in *Saccharomyces*  
514 *cerevisiae*. *BMC Mol Biol* 10: 99.

515 Thakur B, Yadav R, Fraissinet-Tachet L, Mermeisse R, Reddy MS. 2018. Isolation of multi-  
516 metal tolerant ubiquitin fusion protein from metal polluted soil by metatranscriptomic  
517 approach. *J Microbiol Methods* 152: 119-125.

518 Tchounwou PB, Yedjou, CG, Patlolla, AK, Sutton DJ. 2012. Heavy metal toxicity and the  
519 environment. *Molecular, clinical and environmental toxicology* 133-164.

520 Tufail M, Naeemullah M, Elmogy M, Sharma PN, Takeda M, Nakamura C. 2010, Molecular  
521 cloning, transcriptional regulation and differential expression profiling of vitellogenin in  
522 two wing morphs of the brown planthopper, *Nilaparvata ligens* Stal (Hemiptera,  
523 Delphacidae). *Insect Mol Biol* 19: 787-798.

524 Tufail M, Takeda M. 2008. Molecular characteristics of insect vitellogenins. *J. Insect Physiol*  
525 54: 1447-58.

526 Walkley A. 1947. A critical examination of a rapid method for determining organic carbon in  
527 soils- effect of variations in digestion conditions and of inorganic soil constituents. *Soil*  
528 *Sci* 63: 251-264.

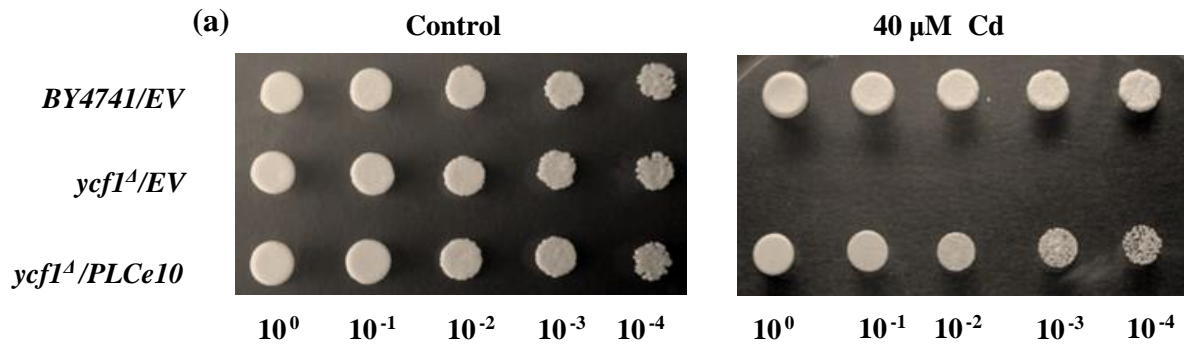
529 Wallace RA. 1985. Vitellogenesis and oocyte growth in non-mammalian vertebrates. *Dev Biol*  
530 1: 127-177.

531 Yadav RK, Barbi F, Ziller A, Luis P, Marmeisse R, Reddy MS, Fraissinet-Tachet L. 2014.  
532 Construction of sized eukaryotic cDNA libraries using low input of total environmental  
533 metatranscriptomic RNA. *BMC Biotechnol* 14: 80-86.

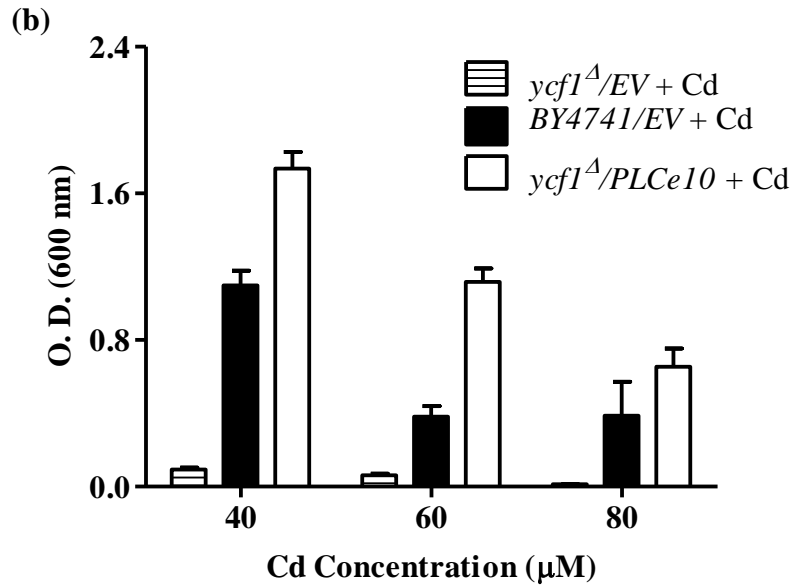
534 Zielinski BL, Allen AE, Carpenter EJ, Coles VJ, Crump BC, Doherty M, McCrow JP. 2016.  
535 Patterns of transcript abundance of eukaryotic biogeochemically-relevant genes in the  
536 Amazon River plume. *PloS one* 11: e0160929.

537 Žifčáková L, Větrovský T, Howe A, Baldrian P. 2016. Microbial activity in forest soil reflects  
538 the changes in ecosystem properties between summer and winter. *Environ Microbiol* 18:  
539 288-301.

540 Ziller A, Yadav RK, Capdevila M, Reddy MS, Vallon L, Marmeisse R, Atrian S, Palacios O,  
541 Frassiniet-Tachet L. 2016. Metagenomics analysis reveals a new metallothionein family,  
542 Sequence and metal-binding features of new environmental cysteine-rich proteins. *J.*  
543 *Inorg. Biochem* 167: 1-11.  
544



545  
546



547  
548  
549  
550  
551  
552

**Fig.1.** Functional complementation of PLCe10 (a) Drop assay of mutant strain *ycf1<sup>Δ</sup>* expressing PLCe10 on Cd 40  $\mu$ M. (b) Growth assay of mutant strain *ycf1<sup>Δ</sup>* expressing PLCe10 in SD-ura broth with different concentrations of Cd. Wild type BY4741 and *ycf1<sup>Δ</sup>* transformed with empty vector pFL61 (EV) were used as controls. Errors bars are  $\pm$ SD.

```

PLCe10      CGLCGHFDLESTDEFRNPDFTEQDIRQFYMNYLIKDGQKAPQQLTEVCESEECCKADRSSSSSSSS-S-----S
KHN77828   CGLCGHYDGEKLDLRAADNGLTDDVEEYSRSYFDKGDENIIEEDIVSEKNSYRFDDID-----ENFDIDDVEEE-
ERG83573   CGLCGHYDGEKQYELRMADNELTDDLEQYSRSYFNDEEDIEENIVKEKMNYRLEDDSN-----EFFGEDEDEEH-
CCD65563   CGLCGNNDDESTNEFYTSDNTEDEIEEFHRSYLLKNEEAEAEERLSEKKNYRKYERDEEQSDEYSSEETYDYE--QENT
ABV46673   CGLCGHFDEEDKDSEFYTPNKEYTSDVEEFHKSYPYLLT-DKVEVEKEFSFEKKDYAVETNEER--SDDWL-STYDDDNSTND-
CDJ93502   CGLCGHFDEEKDNEFYTPKKEYTDDIMEFHKSYPYLLN-DEVEVEKELPREKKHYRLEKSESSESSEEDWL-DFYENDDKRKEM
KHJ90957   CGLCGHYDEEKDNEFLTPKMERTTDIMEFHKSYPYLLD-EEEMEKEKELINEKKHYKLEEREH---DDDFL-DIFFEEDIDT-

```

```

PLCe10      SGSQESQSNEESEIPEKKTKEVIEIDDQLFSTVPIPQDEEDSYPMGEKEKRKVAVVICIDQDSQEAEDIERQARSGRRQIP
KHN77828   -----EEREIKKPILRTKIEHNYEVFSMKAVPHCPHKTAYARNEMKTRKKVPFVCLPRTNHEATNFINRA---LHEVL
ERG83573   -----TESNIKKPILKTKIEYNNVEFSMKPVKECPQQTAYAREDNKTQRKVPFACLPRSTYETTKLLNRA---RFEVL
CCD65563   KKSQKNQRSQKSDLVEKTIKEFQSHRIEFSVEPVAECPRGYE--VEQQQQRKIRFTCLQRHNRDASRLLKES---RQQPL
ABV46673   -----LKDFDEEPLKTTVMFPHRVFSLEPVRKCRKNEK--MDDMVEKKVRFTCLPRSSHETRQLLHKA---RTSVL
CDJ93502   RRTYKYRAESEEIILEKHHVIEYPHQVFSLEPVRVCRKNEV--KGMTMDKVRFTCLPRSSREARELLHKV---RKNVV
KHJ90957   -----KNKREKDEILKRNHVMEFSRVFSREPVRCRKNEE--MDEVVDMKVRFTCLPRTSHEARQLLHKV---RNNVL

```

```

PLCe10      ALNNRTPSFTRTEKIPEKPKKYSRN
KHN77828   DVNDYKKSFVEEVTVPDSRAAF---
ERG83573   NMNKYKASFTEEVSIPEKPKRTF---
CCD65563   QLDDYPVSFVESVKVPTAVAY---
ABV46673   ELNDYPISFVENLRVPLHALSTKLM
CDJ93502   DLSRHPISFVETIEVPRTTVY---
KHJ90957   NLHDYPVSFVETIQVPRTVVY---

```

553

554 **Fig.2.** Alignment of amino acid sequences of different organisms showing highly conserved  
555 GL/ICG motif (red highlighted) to the C-terminal ends. Accession numbers are in parenthesis  
556 *Toxocara canis* (KHN77828), *Ascaris suum* (ERG83573), *Caenorhabditis elegans*  
557 (CCD65563), *Angiostrongylus vasorum* (ABV46673), *Haemonchus contortus* (CDJ93502) and  
558 *Oesophagostomum dentatum* (KHJ90957). Cysteine residues are highlighted with green colour



MWEQLREYVQDDEDQAQ**C**SITSK**S**R**RQGRS**QVETFDGTKFSAPFTN**C**WVV 50  
LAKD**C**GSQQPKFVVMARK**S**ERGQDLKEVKIV**T**KKHRIQL**T**PD**S**AEY**N**SVK 100  
VQVNGQQYDPESDQQITENGQVVQIEKDQSTVQVQLPNTGLEVEFDGYAI 150  
NIQLSQAYRGQQ**CGLCG**HFDLESTDEFRNPDFTDEQDIRQFYMNYLIKDG 200  
Q**C**KAPQQLTEV**C**E**S**EE**C**DKADR**SSSSSSSSSSSS**G**S**Q**E**S**S**NEESEIPEKK 250  
TKVIEIDDQL**C**FSTVPIPQ**C**DEED**S**YPMGEKEKRKVAYV**C**IDQ**S**QEAEED 300  
IE**RQAR****S**GRRQIPALNNRTP**S**FTR**T**EKIPEK**C**KKY**S**RN 338

559

560

561

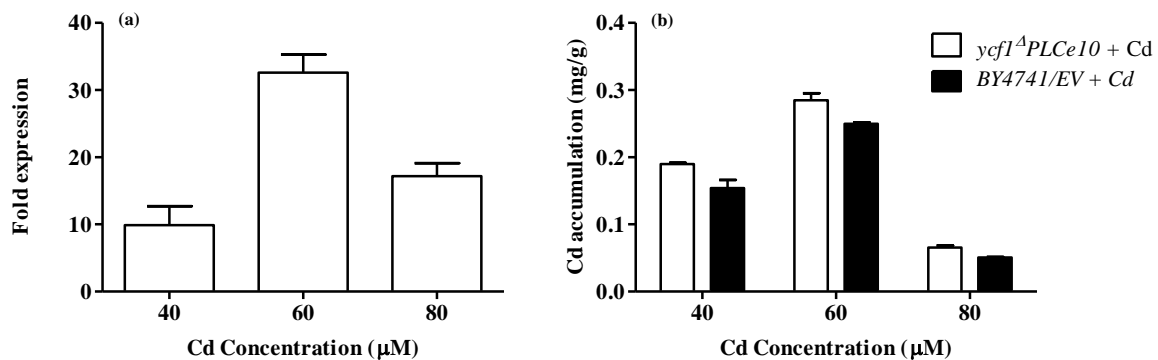
562

563

564

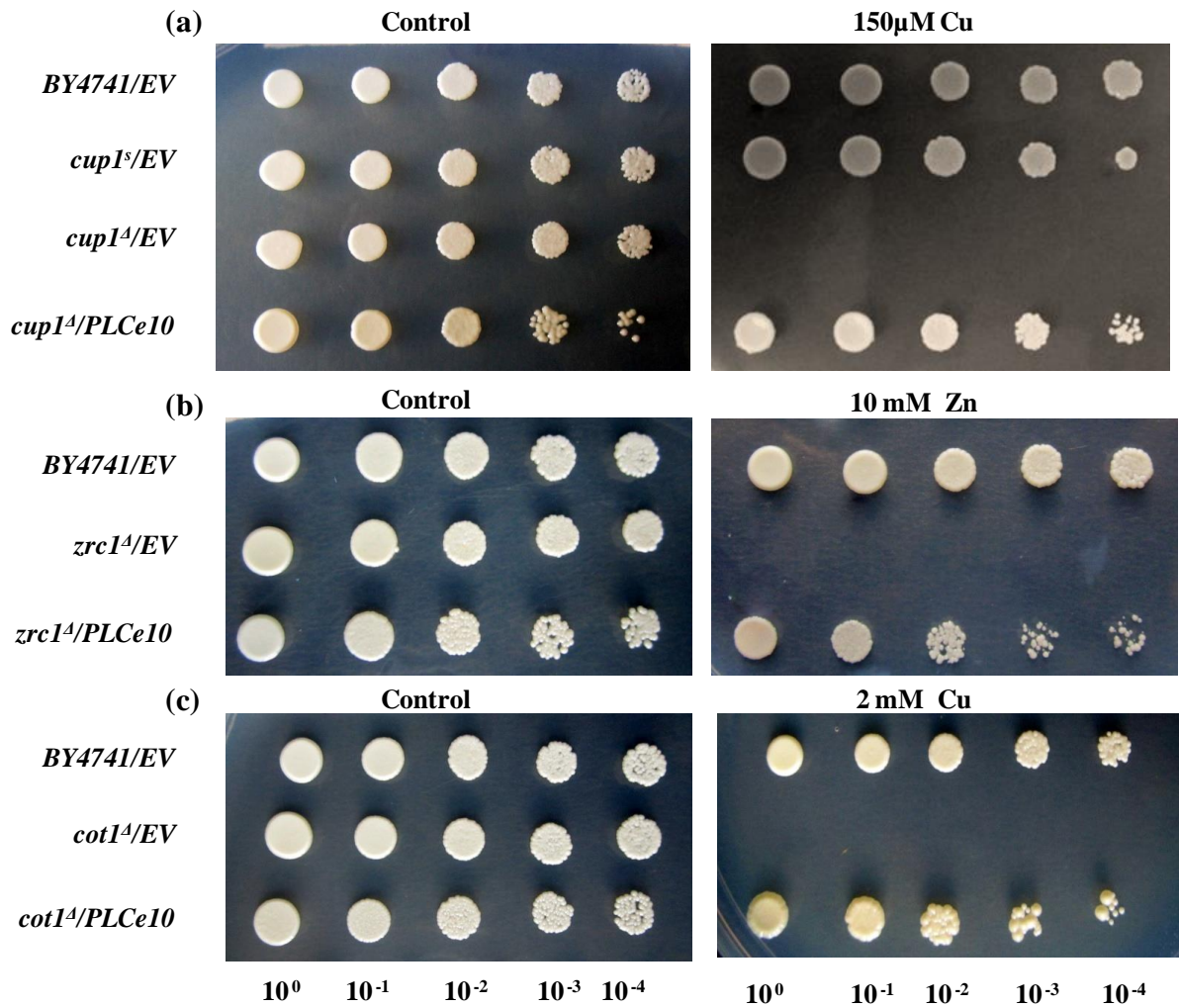
**Fig.3.** Deduced amino acid sequence of VWD protein: Highly conserved GL/ICG motif (highlighted in green) to the C- terminal end. Cysteine residues are given in bold red color while RXXR cleavage sites at the N- terminal end are highlighted in blue with white font. Bold and boxed amino acids are sites for O- linked glycosylation and grey highlighted are phosphorylation sites.

565



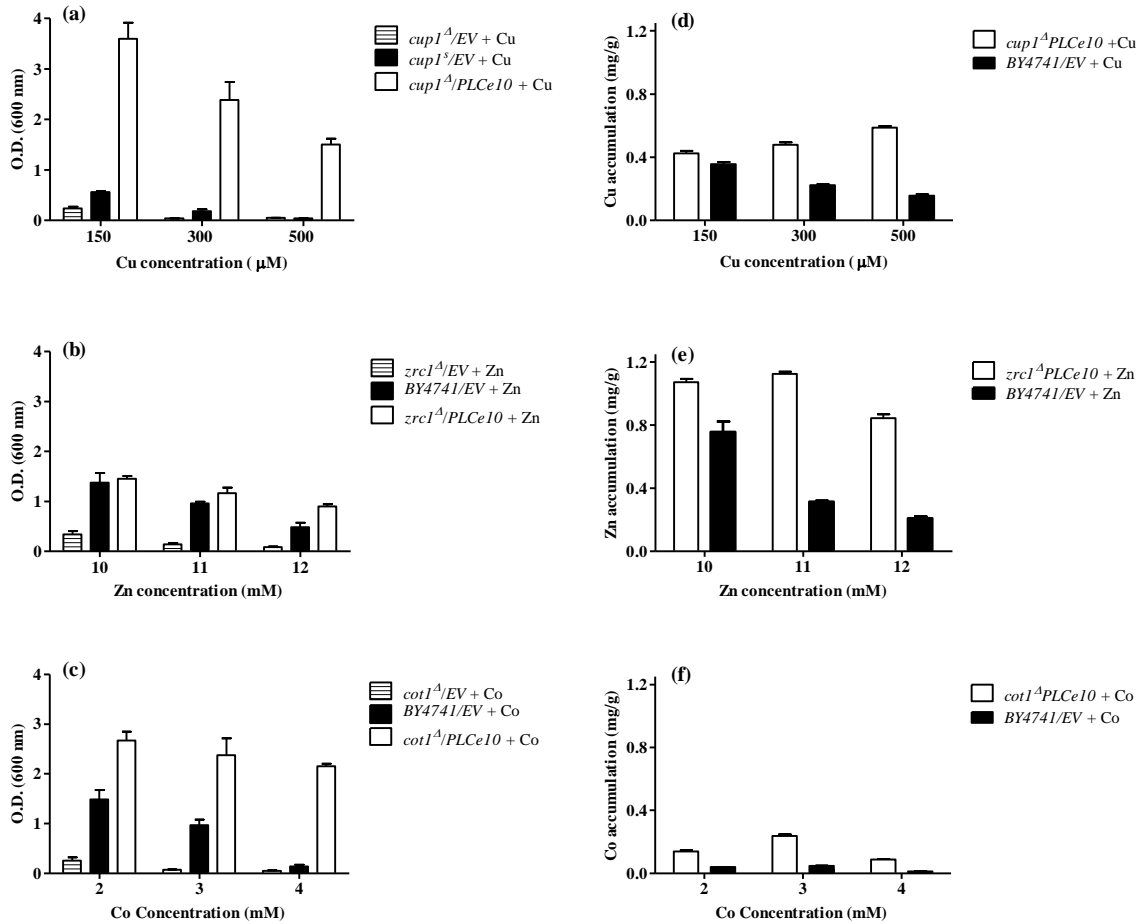
566  
 567 **Fig.4.** a) Fold increase in expression levels of cDNA PLCe10 in *ycf1 $\Delta$*  after 48 h incubation in  
 568 medium supplemented with different concentrations of Cd. Values plotted are referred to the  
 569 control condition (expression level in transformant *ycf1 $\Delta$* PLCe10 without metal treatment), b)  
 570 Accumulation of Cd in *ycf1 $\Delta$*  carrying cDNA PLCe10 and BY4741/EV. EV stands for empty  
 571 vector pFL61. Error bars are  $\pm$ SD.

572



573

574 **Fig.5.** Drop assay of metal sensitive mutant strains *cup1<sup>Δ</sup>*, *zrc1<sup>Δ</sup>* and *cot1<sup>Δ</sup>* expressing PLCe10  
 575 at various concentrations a) Cu 150 μM, b) Zn 10 mM and c) Co 2 mM . Wild strain BY4741  
 576 and *cup1<sup>s</sup>* are used as control and EV denotes to empty vector pFL61.



577  
 578 **Fig.6.** Growth assay of metal sensitive mutant strains *cup1* $\Delta$ , *zrc1* $\Delta$  and *cot1* $\Delta$  expressing  
 579 PLCe10 in SD-ura broth with metal supplementation for 48hrs. a) Cu 150 - 500  $\mu$ M, b) Zn 10  
 580 - 12 mM and c) Co 2 - 4 mM. Accumulation of d) Cu, e) Zn and f) Co in mutant yeast cells  
 581 *cup1* $\Delta$ , *zrc1* $\Delta$  and *cot1* $\Delta$  carrying cDNA PLCe10 and BY4741/EV. Wild type BY4741 and  
 582 mutant strain transformed with only empty vector pFL61 (EV) were used as controls. Error  
 583 bars are  $\pm$ SD.  
 584

585 **Table 1.** Primers used to amplify different genes in this study.  
 586

<b>pFL61 vector specific primers</b>	
pFL61 NF	5'-CAGATCATCAAGGAAGTAATTATCTAC-3'
pFL61 NR	5'-CAGAAAAGCAGGCTGGG AAGC-3'
<b>qPCR primers for yeast (<i>S. cerevisiae</i>)</b>	
Actin	qACTF: 5'-CGAATTGAGAGTTGCCCCAG-3' qACTR: 5'-CAAGGACAAAACGGCTTGGA-3'
TAF10	qTAF10F: 5'-GCTAACAACAGTCAGGCGAG-3' qTAF10R: 5'-GAGCCCGTATTCAGCAACAG-3'
TFC1	qTFC1F: 5'-ACGACGCTGCTTTGGAAAAT-3' qTFC1R: 5'-GCTTTTCATTGTTTCCGCGG-3'
PLCe10	qvitF: 5'-ACCGTCCAAGTTCAATTGCC-3' qvitR: 5'-GAATGTCCTGTTCGTCGGTG-3'

587

588