Screening of metallic pollution in complex environmental samples through a transcriptomic fingerprint method

*Mickaël Cregut12\*, Anna Hua1\*, Sulivan Jouanneau1, Ali Assaf1, Christophe B.Y. Cordella3, Gérald. Thouand1, Marie-José. Durand1\**

1 University of Nantes, Oniris, CNRS, GEPEA, UMR 6144, F-85000 La Roche sur Yon, France

2 University of Nantes, CAPACITES, Bd Vincent Gâche, 44 200 Nantes, France

3 University of Paris-Saclay, UMR AgroParisTech/INRA 914 PCNA, Rue Claude Bernard, 75 005 Paris, France

Supplemental Information

**Table S1 Transcripts specific to the fingerprint “Presence of metals”**

|  |  |  |
| --- | --- | --- |
|  | **Presence of metals (174)** | |
| **Ref n°** | **Transcript** | **Function** |
| *1* | *adiA* | arginine decarboxylase(adiA) |
| *2* | *ais* | putative LPS core heptose(II)-phosphate phosphatase(ais) |
| *3* | *aldB* | aldehyde dehydrogenase B(aldB) |
| *4* | *amiA* | N-acetylmuramoyl-l-alanine amidase I(amiA) |
| *5* | *araD* | L-ribulose-5-phosphate 4-epimerase(araD) |
| *6* | *argA* | amino acid N-acetyltransferase and inactive acetylglutamate kinase(argA) |
| *7* | *b0165* | unknown function |
| *8* | *b0257* | unknown function |
| *9* | *b0582* | unknown function |
| *10* | *b1172* | unknown function |
| *11* | *b2680* | unknown function |
| *12* | *blc* | outer membrane lipoprotein cell division and growth lipocalin(blc) |
| *13* | *cfa* | cyclopropane fatty acyl phospholipid synthase, SAM-dependent(cfa) |
| *14* | *chaC* | cation transport regulator(chaC) |
| *15* | *cirA* | colicin IA outer membrane receptor and translocator; ferric iron-catecholate transporter(cirA) |
| *16* | *codB* | cytosine transporter(codB) |
| *17* | *corA* | magnesium/nickel/cobalt transporter(corA) |
| *18* | *cspD* | inhibitor of DNA replication, cold shock protein homolog(cspD) |
| *19* | *cspI* | Qin prophage; cold shock protein(cspI) |
| *20* | *cysZ* | sulfate transporter, sulfite inhibited(cysZ) |
| *21* | *dksA* | transcriptional regulator of rRNA transcription; DnaK suppressor protein(dksA) |
| *22* | *dpbA* | unknown function |
| *23* | *ecpD* | polymerized tip adhesin of ECP fibers(ecpD) |
| *24* | *elaB* | putative membrane-anchored DUF883 family ribosome-binding protein(elaB) |
| *25* | *emrE* | DLP12 prophage; multidrug resistance protein(emrE) |
| *26* | *fadL* | long-chain fatty acid outer membrane transporter(fadL) |
| *27* | *fes* | enterobactin/ferrienterobactin esterase(fes) |
| *28* | *fhuA* | ferrichrome outer membrane transporter(fhuA) |
| *29* | *fhuF* | ferric iron reductase involved in ferric hydroximate transport(fhuF) |
| *30* | *fliL* | flagellar biosynthesis protein(fliL) |
| *31* | *focA* | formate channel(focA) |
| *32* | *folD* | bifunctional 5,10-methylene-tetrahydrofolate dehydrogenase/ 5,10-methylene-tetrahydrofolate cyclohydrolase(folD) |
| *33* | *fpr* | ferredoxin-NADP reductase; flavodoxin reductase(fpr) |
| *34* | *fxsA* | suppressor of F exclusion of phage T7(fxsA) |
| *35* | *gcvT* | aminomethyltransferase, tetrahydrofolate-dependent, subunit (T protein) of glycine cleavage complex(gcvT) |
| *36* | *glnH* | glutamine transporter subunit(glnH) |
| *37* | *glnL* | sensory histidine kinase in two-component regulatory system with GlnG(glnL) |
| *38* | *glnU* | tRNA(glnU) |
| *39* | *glxK* | glycerate kinase II(glxK) |
| *40* | *gmr* | cyclic-di-GMP phosphodiesterase; csgD regulator; modulator of RNase II stability(gmr) |
| *41* | *gntT* | gluconate transporter, high-affinity GNT I system(gntT) |
| *42* | *grxA* | glutaredoxin 1, redox coenzyme for ribonucleotide reductase (RNR1a)(grxA) |
| *43* | *gudP* | putative D-glucarate transporter(gudP) |
| *44* | *hcaT* | putative 3-phenylpropionic transporter(hcaT) |
| *45* | *hemN* | coproporphyrinogen III oxidase, SAM and NAD(P)H dependent, oxygen-independent(hemN) |
| *46* | *hisQ* | histidine ABC transporter permease(hisQ) |
| *47* | *hofM* | DNA catabolic putative pilus assembly protein(hofM) |
| *48* | *htpX* | putative endopeptidase(htpX) |
| *49* | *hyaA* | hydrogenase 1, small subunit(hyaA) |
| *50* | *kdgR* | KDG regulon transcriptional repressor(kdgR) |
| *51* | *kil* | unknown function |
| *52* | *lacA* | thiogalactoside acetyltransferase(lacA) |
| *53* | *lolA* | lipoprotein chaperone(lolA) |
| *54* | *lpxC* | UDP-3-O-acyl N-acetylglucosamine deacetylase(lpxC) |
| *55* | *maeB* | malic enzyme: putative oxidoreductase/phosphotransacetylase(maeB) |
| *56* | *mdtG* | putative drug efflux system protein(mdtG) |
| *57* | *metJ* | transcriptional repressor, S-adenosylmethionine-binding(metJ) |
| *58* | *mglB* | methyl-galactoside transporter subunit(mglB) |
| *59* | *mpaA* | murein peptide amidase A(mpaA) |
| *60* | *mraZ* | RsmH methytransferase inhibitor(mraZ) |
| *61* | *msbB* | unknown function |
| *62* | *mscL* | mechanosensitive channel protein, high conductance(mscL) |
| *63* | *murB* | UDP-N-acetylenolpyruvoylglucosamine reductase, FAD-binding(murB) |
| *64* | *mutH* | methyl-directed mismatch repair protein(mutH) |
| *65* | *nanA* | N-acetylneuraminate lyase(nanA) |
| *66* | *nanR* | sialic acid-inducible nan operon repressor(nanR) |
| *67* | *nanT* | sialic acid transporter(nanT) |
| *68* | *nfnB* | unknown function |
| *69* | *nupC* | nucleoside (except guanosine) transporter(nupC) |
| *70* | *nupG* | nucleoside transporter(nupG) |
| *71* | *osmB* | osmotically and stress inducible lipoprotein(osmB) |
| *72* | *otsB* | trehalose-6-phosphate phosphatase, biosynthetic(otsB) |
| *73* | *pheP* | phenylalanine transporter(pheP) |
| *74* | *phoP* | response regulator in two-component regulatory system with PhoQ(phoP) |
| *75* | *pitA* | phosphate transporter, low-affinity; tellurite importer(pitA) |
| *76* | *pmbA* | putative antibiotic peptide MccB17 maturation peptidase(pmbA) |
| *77* | *pmrD* | inactive two-component system connector protein(pmrD) |
| *78* | *pqiA* | paraquat-inducible, SoxRS-regulated inner membrane protein(pqiA) |
| *79* | *pstC* | phosphate ABC transporter permease(pstC) |
| *80* | *qseB* | quorum sensing DNA-binding response regulator in two-component regulatory system with QseC(qseB) |
| *81* | *rbsD* | D-ribose pyranase(rbsD) |
| *82* | *ribE* | riboflavin synthase beta chain(ribE) |
| *83* | *rplL* | 50S ribosomal subunit protein L7/L12(rplL) |
| *84* | *rrfE* | 5S ribosomal RNA of rrnE operon(rrfE) |
| *85* | *rrlA* | 23S ribosomal RNA of rrnA operon(rrlA) |
| *86* | *rrlB* | 23S ribosomal RNA of rrnB operon(rrlB) |
| *87* | *rrlD* | 23S ribosomal RNA of rrnD operon(rrlD) |
| *88* | *rssA* | putative patatin-like family phospholipase(rssA) |
| *89* | *sapA* | antimicrobial peptide transport ABC transporter periplasmic binding protein(sapA) |
| *90* | *shiA* | shikimate transporter(shiA) |
| *91* | *sodA* | superoxide dismutase, Mn(sodA) |
| *92* | *speC* | ornithine decarboxylase, constitutive(speC) |
| *93* | *ssuE* | NAD(P)H-dependent FMN reductase(ssuE) |
| *94* | *tdcG\_2* | unknown function |
| *95* | *tonB* | membrane spanning protein in TonB-ExbB-ExbD transport complex(tonB) |
| *96* | *trkA* | NAD-binding component of TrK potassium transporter(trkA) |
| *97* | *ubiF* | 2-octaprenyl-3-methyl-6-methoxy-1,4-benzoquinol oxygenase(ubiF) |
| *98* | *uhpA* | response regulator in two-component regulatory system wtih UhpB(uhpA) |
| *99* | *uspE* | stress-induced protein(uspE) |
| *100* | *uspF* | stress-induced protein, ATP-binding protein(uspF) |
| *101* | *uxaC* | uronate isomerase(uxaC) |
| *102* | *yabN* | unknown function |
| *103* | *yacH* | DUF3300 family protein(yacH) |
| *104* | *yadG* | putative ABC transporter ATPase(yadG) |
| *105* | *yadK* | putative fimbrial-like adhesin protein(yadK) |
| *106* | *yaiI* | UPF0178 family protein(yaiI) |
| *107* | *yaiZ* | DUF2754 family putative inner membrane protein(yaiZ) |
| *108* | *yajO* | 2-carboxybenzaldehyde reductase(yajO) |
| *109* | *yajR* | putative transporter(yajR) |
| *110* | *ybbL* | unknown function |
| *111* | *ybcK* | DLP12 prophage; putative recombinase(ybcK) |
| *112* | *ybhD* | putative DNA-binding transcriptional regulator(ybhD) |
| *113* | *ybhM* | BAX Inhibitor-1 family inner membrane protein(ybhM) |
| *114* | *ybiO* | mechanosensitive channel protein, intermediate conductance(ybiO) |
| *115* | *ybjG* | undecaprenyl pyrophosphate phosphatase(ybjG) |
| *116* | *ycaC* | putative isochorismatase family hydrolase(ycaC) |
| *117* | *ycaK* | putative NAD(P)H-dependent oxidoreductase(ycaK) |
| *118* | *ycaN* | LysR family putative transcriptional regulator(ycaN) |
| *119* | *ycbR* | unknown function |
| *120* | *ycfQ* | repressor for bhsA(ycfR)(ycfQ) |
| *121* | *yciC* | UPF0259 family inner membrane protein(yciC) |
| *122* | *ycjG* | L-Ala-D/L-Glu epimerase(ycjG) |
| *123* | *ycjW* | LacI family putative transcriptional repressor(ycjW) |
| *124* | *ycjZ* | unknown function |
| *125* | *ydcA* | putative periplasmic protein(ydcA) |
| *126* | *ydcI* | putative DNA-binding transcriptional regulator(ydcI) |
| *127* | *ydcJ* | putative metalloenzyme(ydcJ) |
| *128* | *yddE* | PhzC-PhzF family protein(yddE) |
| *129* | *yddM* | putative DNA-binding transcriptional regulator(yddM) |
| *130* | *yddW* | liprotein, glycosyl hydrolase homolog(yddW) |
| *131* | *yeaM* | putative DNA-binding transcriptional regulator(yeaM) |
| *132* | *yebG* | DNA damage-inducible protein regulated by LexA(yebG) |
| *133* | *yedW* | response regulator family protein(yedW) |
| *134* | *yeeR* | CP4-44 prophage; putative membrane protein(yeeR) |
| *135* | *yffS* | CPZ-55 prophage; uncharacterized protein(yffS) |
| *136* | *yfjR* | CP4-57 prophage; putative DNA-binding transcriptional regulator(yfjR) |
| *137* | *ygfU* | unknown function |
| *138* | *ygjG* | unknown function |
| *139* | *ygjH* | putative tRNA binding protein; putative tRNA corner chaperone(ygjH) |
| *140* | *yhbQ* | GIY-YIG nuclease superfamily protein(yhbQ) |
| *141* | *yhcG* | DUF1016 family protein in the PD-(D/E)XK nuclease superfamily(yhcG) |
| *142* | *yhcO* | putative barnase inhibitor(yhcO) |
| *143* | *yhfL* | small lipoprotein(yhfL) |
| *144* | *yhiJ* | DUF4049 family protein(yhiJ) |
| *145* | *yhjJ* | putative periplasmic M16 family chaperone(yhjJ) |
| *146* | *yidK* | putative transporter(yidK) |
| *147* | *yieE* | phosphopantetheinyl transferase superfamily protein(yieE) |
| *148* | *yihR* | putative sulphoquinovose mutarotase(yihR) |
| *149* | *yijO* | AraC family putative transcriptional activator(yijO) |
| *150* | *yjhF* | putative transporter(yjhF) |
| *151* | *ykgC* | unknown function |
| *152* | *ykgI* | unknown function |
| *153* | *ymdB* | O-acetyl-ADP-ribose deacetylase; RNase III inhibitor during cold shock; putative cardiolipin synthase C regulatory subunit(ymdB) |
| *154* | *ymjA* | DUF2543 family protein(ymjA) |
| *155* | *yncE* | ATP-binding protein, periplasmic, function unknown(yncE) |
| *156* | *yncG* | glutathione S-transferase homolog(yncG) |
| *157* | *yneE* | bestrophin family putative inner membrane protein(yneE) |
| *158* | *yneH* | unknown function |
| *159* | *yneI* | unknown function |
| *160* | *ynfC* | UPF0257 family lipoprotein(ynfC) |
| *161* | *ynfK* | putative dethiobiotin synthetase(ynfK) |
| *162* | *ynfL* | LysR family putative transcriptional regulator(ynfL) |
| *163* | *yniA* | fructosamine kinase family protein(yniA) |
| *164* | *yphH* | putative DNA-binding transcriptional regulator(yphH) |
| *165* | *ypjD* | cytochrome c assembly protein family inner membrane protein(ypjD) |
| *166* | *yqcD* | unknown function |
| *167* | *yqeG* | putative transporter(yqeG) |
| *168* | *yqiJ* | DUF1449 family inner membrane protein(yqiJ) |
| *169* | *yqjC* | DUF1090 family putative periplasmic protein(yqjC) |
| *170* | *yqjG* | putative S-transferase(yqjG) |
| *171* | *yqjH* | putative siderophore interacting protein(yqjH) |
| *172* | *yraQ* | putative inner membrane permease(yraQ) |
| *173* | *znuA* | zinc ABC transporter periplasmic binding protein(znuA) |
| *174* | *znuC* | zinc ABC transporter ATPase(znuC) |
| *175* | *Ctrl* | Control Strain with no promotor; pMSs201\_U66 |

In parenthesis, the number of regulated transcripts. Functional annotations from DAVID Bioinformatics (version 6.8), *Escherichia coli* organism.

Table S2: Regulated metabolic pathways in the presence of copper

|  |  |  |  |
| --- | --- | --- | --- |
| Function | C1 | C2 | C3 |
| Copper metabolism | | | |
| Copper export | *cusR* (6.2) | *cusR* (8.0) | *cusR* (28.5) |
|  | *cusC* (21.5) | *cusC* (47.5) |
| Iron homeostasis | *fes* (16.8) | *fes* (4.5) | *fes* (13.0) |
| Iron import | *cirA* (3.8) | *yqjH* (9.17) | *fhuA* (2.2) |
|  | *fecA* (4.0) | *fhuF* (3.3) |
| Oxidative stress | | | |
| *cpx* regulon |  | *htpX* (8.2) |  |
|  | *mraZ* (5.0) |  |
|  | *ppiA* (4.3) |  |
| *sox* regulon |  |  | *fpr* (16.1) |
|  |  | *sodA* (42.9) |
|  |  | *soxS* (4.6) |
| Flagellar biosynthesis |  |  | *flgM* (2.4) |
|  |  | *fliC* (8.4) |
| Secondary metabolism | | | |
| Polysaccharides | *glgB* (5.4) |  |  |
| Fatty acids | *fadE* (6.0) | *fabB* (3.3) |  |
| Amino acids |  | *argA* (3.0) |  |
| Coenzymes |  |  | *metK* (0.24) |

The values of the induction factor are in parentheses.

**Detailed Information regarding the table S2**

In this study, the toxicity of 6 metals was characterized with *E.coli* prior performing transcriptomic studies. Toxicity found for the different metals appears in accordance with the literature for *E. coli* K12MG1655 (Jouanneau et al., 2011). These concentrations were further used as reference values to assess the functionality of the transcriptomic fingerprint analysis to report on the toxicity of environmental samples. Exposing the fluorescent library composed of 1,870 strains to 3 increasing concentrations of metal allowed to determine transcripts impacted by the presence of a metallic contamination. Relevance of the transcriptomic changes was investigated on the physiological viewpoint, for copper for which toxicity is well decribed at the transcriptomic level. Transcriptomics changes due to the presence of copper were compared with the known regulations described in the literature (Table S2, Supplementary Information). Among the 184 transcripts forwhich expression is altered by Copper, a total of 3 mains metabolic functions can be found, genes implicated in Copper metabolism, genes implicated in response to the oxidative stress and finally those implicated in secondary metabolism. Among them, transcripts of the two copper homeostasis genes *cusC* and *cusR*, which code for an efflux pump and its regulatory protein, respectively, were both activated in a concentration-dependent manner (Munson et al. 2000; Yamamoto and Ishihama 2005). Similarly, due to the competition between copper and iron for iron-binding sites on proteins, genes involved in iron import (Arredondo and Núñez 2005), such as the ferrienterobactin esterase *fes*, the ferric iron-catecholate transporter *cirA* and the ferrichrome outer membrane transporter *fhuA*, were activated (McHugh et al. 2003). Finally, with increasing concentrations of copper, there was protein misfolding and production of oxygen radicals (Garrick et al. 2003). This behavior led to an increase in the Cpx regulon, as seen by activation of the membrane protease *htpX* (Shimohata et al. 2002) and of the peptidyl-prolyl isomerase *ppiA* (Pogliano et al. 1997)*.* Moreover, the Sox regulon was activated, as seen by the activation of the regulatory protein SoxS. Compared to the literature (Kershaw et al. 2005), this genome-wide bacterial sensor analysis led to the identification of the same mechanisms of resistance. This analysis validate the biological relevance of the transcriptomic method to detect toxicity.

**Associated References**

Jouanneau S, Durand MJ, Courcoux P, Blusseau T, Thouand G (2011) Improvement of the identification of four heavy metals in environmental samples by using predictive decision tree models coupled with a set of five bioluminescent bacteria. Environ SciTechnol 45(7):2925–2931. https://doi.org/10.1021/es1031757.

Munson GP, Lam DL, Outten FW, O’Halloran TV (2000) Identification of a copper-responsive two-component system on the chromosome of Escherichia coli K-12. J Bacteriol 182(20):5864–5871.

Yamamoto K, Ishihama A (2005) Transcriptional response of Escherichia coli to external copper. MolMicrobiol 56(1):215–227. https://doi.org/10.1111/j.1365-2958.2005.04532.x.

Arredondo M, Núñez MT (2005) Iron and copper metabolism. Mol Aspects Med 26(4):313–327. https://doi.org/10.1016/j.mam.2005.07.010.

McHugh JP, Rodríguez-Quiñones F, Abdul-Tehrani H, Svistunenko DA, Poole RK, Cooper CE, Andrews SC (2003) Global iron-dependent gene regulation in Escherichia coli a new mechanism for iron homeostasis. J BiolChem 278(32):29478–29486. https://doi.org/10.1074/jbc.M303381200.

Garrick MD, Núñez MT, Olivares M, Harris ED (2003) Parallels and contrasts between iron and copper metabolism. Biometals 16(1):1–8. https://doi.org/10.1023/A:1020735401734.

Shimohata N, Chiba S, Saikawa N, Ito K, Akiyama Y (2002) The Cpx stress response system of Escherichia coli senses plasma membrane proteins and controls HtpX, a membrane protease with a cytosolic active site. Genes Cells 7(7): 653–662. https://doi.org/10.1046/j.1365-2443.2002.00554.x

Pogliano J, Lynch AS, Belin D, Lin EC, Beckwith J (1997) Regulation of Escherichia coli cell envelope proteins involved in protein folding and degradation by the Cpx two-component system. Genes Dev 11(9):1169–1182. https://doi.org/10.1101/gad.11.9.1169.

Kershaw CJ, Brown NL, Constantinidou C, Patel MD, Hobman JL (2005) The expression profile of *Escherichia coli* K-12 in response to minimal, optimal and excess copper concentrations. Microbiology 151(4):1187–1198. https://doi.org/10.1099/mic.0.27650-0.

|  |  |
| --- | --- |
| A |  |
| B |  |

Figure S1: Arsenic toxicity test on the *E. coli* K12 MG1655 pMSs201\_U139 strain with growth kinetics of the *E. coli* K12 MG1655 pMSs201\_U139 strain in the presence of arsenic (A) and rate of growth inhibition depending on increasing concentrations of arsenic, calculated according to the Regtox macro (B). A total of 100 µL of bacteria were added to 25 µL of metal and cultured in HEPES medium at 37°C. The results are from three independent tests.