

## Proteogenomic insights into uranium tolerance of a Chernobyl's *Microbacterium* bacterial isolate

Nicolas Gallois, Béatrice Alpha-Bazin, Philippe Ortet, Mohamed Barakat, Laurie Piette, Justine Long, Catherine Berthomieu, Jean Armengaud, Virginie Chapon

► **To cite this version:**

Nicolas Gallois, Béatrice Alpha-Bazin, Philippe Ortet, Mohamed Barakat, Laurie Piette, et al.. Proteogenomic insights into uranium tolerance of a Chernobyl's *Microbacterium* bacterial isolate. Journal of Proteomics, Elsevier, 2018, 177, pp.148-157. 10.1016/j.jprot.2017.11.021 . cea-01944196

**HAL Id: cea-01944196**

**<https://hal-cea.archives-ouvertes.fr/cea-01944196>**

Submitted on 5 Dec 2018

**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 Proteogenomic insights into uranium tolerance of a Chernobyl's

## 2 *Microbacterium* bacterial isolate

3 Gallois Nicolas<sup>1</sup>, Alpha-Bazin Béatrice<sup>2</sup>, Ortet Philippe<sup>3</sup>, Barakat Mohamed<sup>3</sup>, Piette Laurie<sup>1</sup>, Long  
4 Justine<sup>3</sup>, Berthomieu Catherine<sup>1</sup>, Armengaud Jean<sup>2#</sup>, Chapon Virginie<sup>1</sup>

5 <sup>1</sup>CEA, CNRS, Aix-Marseille Université, UMR 7265 Biologie Végétale et Microbiologie

6 Environnementales, Laboratoire des Interactions Protéine Métal, 13108 Saint-Paul-lez-  
7 Durance, France.

8 <sup>2</sup>Laboratoire Innovations technologiques pour la Détection et le Diagnostic (Li2D), Service de

9 Pharmacologie et Immunoanalyse (SPI), CEA, INRA, F-30207 Bagnols sur Cèze, France.

10 <sup>3</sup>CEA, CNRS, Aix-Marseille Université, UMR 7265 Biologie Végétale et Microbiologie

11 Environnementales, Laboratoire d'écologie microbienne de la rhizosphère et

12 d'environnements extrêmes, 13108 Saint-Paul-lez-Durance, France

13

14 #Corresponding author at:

15 Laboratory «Innovative technologies for Detection and Diagnostics», CEA-Marcoule, DRF-Li2D, PRAE

16 Marcel Boiteux – Départementale 765, BP 17171, F-30200 Bagnols-sur-Cèze cedex, France

17 Tel: +33 4 66 79 62 77; Fax: +33 4 66 79 64 60.

18 E-mail address: jean.armengaud@cea.fr

19 URL: <http://ibitecs.cea.fr/dsv/ibitecs/Pages/services/spi/li2d.aspx>

20 Keywords

21 Radionuclides, toxicity, proteogenomics, bioinformatics, toxicoproteomics, *Microbacterium*

## 22 **Abstract**

23 *Microbacterium oleivorans* A9 is a uranium-tolerant actinobacteria isolated from the trench T22  
24 located near the Chernobyl nuclear power plant. This site is contaminated with different  
25 radionuclides including uranium. To observe the molecular changes at the proteome level occurring  
26 in this strain upon uranyl exposure and understand molecular mechanisms explaining its uranium  
27 tolerance, we established its draft genome and used this raw information to perform an in-depth  
28 proteogenomics study. High-throughput proteomics were performed on cells exposed or not to 10  
29  $\mu\text{M}$  uranyl nitrate sampled at three previously identified phases of uranyl tolerance. We  
30 experimentally detected and annotated 1,532 proteins and highlighted a total of 591 proteins for  
31 which abundances were significantly differing between conditions. Notably, proteins involved in  
32 phosphate and iron metabolisms show high dynamics. A large ratio of proteins more abundant upon  
33 uranyl stress, are distant from functionally-annotated known proteins, highlighting the lack of  
34 fundamental knowledge regarding numerous key molecular players from soil bacteria.

35

## 36 **Biological significance**

37 *Microbacterium oleivorans* A9 is an interesting environmental model to understand biological  
38 processes engaged in tolerance to radionuclides. Using an innovative proteogenomics approach, we  
39 explored its molecular mechanisms involved in uranium tolerance. We sequenced its genome,  
40 interpreted high-throughput proteomic data against a six-reading frame ORF database deduced from  
41 the draft genome, annotated the identified proteins and compared protein abundances from cells  
42 exposed or not to uranyl stress after a cascade search. These data show that a complex cellular  
43 response to uranium occurs in *Microbacterium oleivorans* A9, where one third of the experimental  
44 proteome is modified. In particular, the uranyl stress perturbed the phosphate and iron metabolic  
45 pathways. Furthermore, several transporters have been identified to be specifically associated to

46 uranyl stress, paving the way to the development of biotechnological tools for uranium  
47 decontamination.

48

## 49 **1. Introduction**

50 Radionuclides (RNs) are naturally present in the environment but are redistributed by anthropogenic  
51 activities *e.g.* chemical nitrogen fertilizer use, weapons manufacturing, nuclear research and nuclear  
52 fuel production. Several major nuclear accidents such as Chernobyl and Fukushima led to the  
53 dispersion of several RNs through the atmosphere in microparticles [1] and in the soil [2]. In  
54 Chernobyl, radionuclides-contaminated wastes were buried on site in about 800 trenches. Besides  
55  $^{137}\text{Cs}$  and  $^{90}\text{Sr}$ , other RNs are present in the Chernobyl trench T22, including  $^{60}\text{Co}$ ,  $^{154}\text{Eu}$ ,  $^{238, 239, 240}\text{Pu}$ ,  
56  $^{241}\text{Am}$  and  $^{235, 238}\text{U}$ .

57 Uranium (U) belongs to the actinide series. It exhibits both radiological and chemical toxicities, and a  
58 long half-life, resulting in environmental and human health major concern [3]. As a hard Lewis acid  
59 and according to the HSAB theory which qualitatively explains the stability of metal complexes [4],  
60 uranyl tends to interact with hard Lewis base such as phosphate, sulfate, carbonate or amine by ionic  
61 interactions. It can displace other hard Lewis acids like  $\text{Ca}^{2+}$  and  $\text{Fe}^{3+}$  in highly oxygenated site.  
62 Moreover, different studies have confirmed the affinity between uranium and phosphate or  
63 carbonate [5-9], or the replacement of iron by uranyl [10].

64 Bacteria can interact directly or indirectly with RNs and change their speciation, playing a major role  
65 in their mobility and transfer in the environment. The use of both molecular- and culture-based  
66 approaches to study bacterial diversity in the trench T22 demonstrated that RNs-contaminated soils  
67 host an unexpected diversity of bacteria [11].

68 During their sampling campaign at Chernobyl, Chapon and colleagues have constructed a large  
69 collection of cultured bacteria from the trench T22 soils offering the opportunity to study the RN

70 tolerance of these bacteria [11]. The strain *Microbacterium oleivorans* A9 (referred as  
71 *Microbacterium* sp. A9\_sp3\_-1\_2 in [12]) is one of the most uranium-tolerant isolate of this  
72 collection. In a previous paper, the interactions between uranium and this bacterium have been  
73 investigated and specific exposure conditions in which the bacteria were kept alive while being  
74 exposed to soluble forms of uranium have been set-up. With this tightly controlled exposure  
75 conditions, it has been shown that *Microbacterium oleivorans* A9 exhibit three sequential  
76 mechanisms involved in uranium detoxification: a rapid metal removal within the first 30 minutes,  
77 then an active U(VI) release in the exposure medium accompanied by a phosphate efflux and a final  
78 biomineralization step of uranium in autunite-like mineral phases [12]. For this reason, it constitutes  
79 a relevant model to study the mechanisms involved in uranium tolerance and in uranium-bacteria  
80 interactions. In bacteria, survival in metal-contaminated environment is mainly achieved through  
81 active efflux pumps [13]. If such systems are well-known for a wide variety of non-essential metals,  
82 involvement of an efflux system in detoxification of uranium has not yet been described, although  
83 up-regulation of genes encoding metal efflux pumps has already been reported for  
84 *Desulfotomaculum reducens* exposed to U(VI) in anaerobic conditions [14].

85 The cellular response of *Escherichia coli*, *Acidithiobacillus ferrooxidans* and two *Anabaena* strains  
86 upon uranium exposure has already been studied with 2D-PAGE based proteomic methodology [15-  
87 17]. These studies highlighted an impact of uranium on the bacterial proteome, but the low  
88 throughput of this approach allowed identifying only 11, 18, 45 and 27 proteins respectively. These  
89 may represent only a partial view of the proteins modulated upon uranium stress. Two studies using  
90 high-throughput tandem mass spectrometry have deepened the proteome changes occurring in two  
91 other bacteria, *Geobacter sulfurreducens* and *Caulobacter crescentus* [18, 19]. These papers show  
92 that resistance to uranium is a complex cellular response with induction of multiple stress response  
93 systems. These four studies, based on gram-negative organisms, have used different uranium  
94 concentrations (from 50 to 500  $\mu$ M) as well as different exposure conditions making any comparison  
95 difficult.

96 Proteogenomics, the alliance of genomics and proteomics, has proved its efficiency in supporting the  
97 existence of genes encoding hypothetical proteins or with unknown function, and even unannotated  
98 in the genome for several organisms [20]. This methodology is even more appropriate when a quick  
99 focus on mechanisms for which the main players are unknown is required [21-23]. Because,  
100 proteogenomics is based on the assignment of MS/MS spectra with a six-frame translation of the  
101 genome, time-consuming, resource-dependent and hard-working genome sequence completion  
102 steps are bypassed. For example, the annotation of *Deinococcus deserti* by means of proteogenomics  
103 has shown its added value to correct annotation errors in other *Deinococcus sp.* genomes [21, 24].  
104 The exploration of the halotolerance molecular mechanisms of *Tistlia consotensis* by proteogenomics  
105 based on a draft genome sequence straightforwardly highlighted major changes in response to hypo-  
106 osmotic and hyper-osmotic conditions [22, 25].

107 Here, the molecular changes occurring in *Microbacterium oleivorans* A9 upon uranium exposure  
108 were deciphered by means of a proteogenomics approach consisting in draft genome sequencing  
109 and high-throughput proteome coverage by next-generation proteomics. Proteomic changes of cells  
110 exposed to 10  $\mu$ M uranyl for 30 min, 4 h and 24 h were investigated. This experimental strategy  
111 allowed identification without *a priori* of soluble and membrane proteins with significantly differing  
112 abundances compared to controls, which could be players of the defense mechanism of  
113 *Microbacterium oleivorans* A9 against uranium. Furthermore, these genome and proteome datasets  
114 represent invaluable resources to gain insights into the physiology and role of *Microbacterium* in  
115 soils.

116

## 117 2. Experimental procedures

### 118 2.1 Uranium exposure

119 The experimental procedure used to expose the bacteria to uranium is described in [12].  
120 *Microbacterium oleivorans* A9 strain was routinely cultivated in 0.1 X Tryptic Soy Broth (TSB, Difco  
121 Laboratories) at 30°C with shaking. Cells at exponential growth phase were harvested by  
122 centrifugation for 10 min at 5000 *g*. From this stage on, samples were maintained at 25°C throughout  
123 the experiment. The resulting cell pellets were washed twice in 0.1 M NaCl pH 5.0 and were re-  
124 suspended at about  $6 \times 10^9$  cells mL<sup>-1</sup> in 0.1 M NaCl pH 5.0 with 0 (control) or 10 μM U(VI). U(VI) was  
125 added as uranyl nitrate UO<sub>2</sub>(NO<sub>3</sub>)<sub>2</sub>·6H<sub>2</sub>O (Sigma–Aldrich) from a 7.51 mM stock solution in 16 mM  
126 HNO<sub>3</sub>. No precipitate was observed along the exposure conditions. The nitrate concentration was  
127 adjusted to 0.416 mM by adding NaNO<sub>3</sub> when needed. Bacteria exposed to U(VI) and controls were  
128 incubated at 25°C with shaking. Four independent biological replicates were made for each condition  
129 and for each time point. Fractions of 1 ml of cell suspension were taken after 0.5, 4, and 24 h of  
130 uranium exposure. Samples were centrifuged for 5 min at 8000 *g*. The resulting supernatants were  
131 removed and the pellets were conserved at -80 °C until proteomic analysis.

### 132 2.2 Genome of *Microbacterium oleivorans* A9

133 Genomic DNA was sequenced on a HiSeq 2000 sequencing platform (Illumina) by the GenoScreen  
134 Company and genome *de novo* assembling was performed on the reads using ABYSS [26]. The Whole  
135 Genome Shotgun project data have been deposited at DDBJ/EMBL/GenBank under the accession  
136 number MTIO00000000. For each contig, the ORF translated sequences were extracted from stop  
137 codon to stop codon in each of the six possible reading frames. The ORF-derived polypeptide list was  
138 restricted to those exhibiting a length of at least 50 amino acids. Proteomic data and proteogenomic  
139 procedures were then used to discriminate between false, correct but undetected, and correct and  
140 detected protein sequences as described earlier [21, 27]. The ORF sequences for which at least one  
141 spectrum has been assigned were kept to create a second database. This step allowed us to identify

142 more proteins and to improve qualitatively and quantitatively the MS/MS spectra attribution. The  
143 protein sequences were given a cluster of orthologous groups (COG) assignment based on the  
144 reference database [28]. Furthermore, enzymes annotation using PRIAM profiles [29] associated to  
145 KEGG database [30] allowed the identification of metabolic pathways. An in-house batch program  
146 was developed to emphasize metabolic pathways highlighted by proteomic analysis.

### 147 2.3 Proteome sample preparation and trypsin in-gel proteolysis

148 Cell pellets (between 2.0 and 4.6 mg, 3.1 mg average, wet weight) were dissolved in a given volume  
149 (100  $\mu$ L for 1.7 mg of bacterial pellet) of 1X NuPAGE™ LDS Sample Buffer (Invitrogen) supplemented  
150 with 2%  $\beta$ -mercaptoethanol. Samples were sonicated for 5 min in a transonic 780H sonicator and  
151 subjected to a 5 min incubation at 99°C prior to SDS-PAGE. Samples were subjected to a 3 min short  
152 migration in denaturing conditions as previously described [31]. For this, samples were loaded onto a  
153 4-12% gradient 10-well NuPAGE Bis-Tris (Invitrogen) gel operated with NuPAGE MES (Invitrogen) as  
154 running buffer at 200 V. After gel staining with Simply Blue Safe Stain, a ready-to-use Coomassie G-  
155 250 stain from Invitrogen, the whole proteome content (comprising both soluble and membrane  
156 proteins) from each sample was excised as a single piece of gel polyacrylamide band. The 24  
157 resulting polyacrylamide bands were washed with MilliQ water and dehydrated with  $\text{CH}_3\text{CN}:\text{NH}_4\text{HCO}_3$   
158 50mM (1:1 v:v) before addition of 100%  $\text{CH}_3\text{CN}$ . Gel pieces were then dried for 2 to 5 min under  
159 vacuum. For in-gel digestion, dry gel pieces were rehydrated with 100 mM  $\text{NH}_4\text{HCO}_3$  containing 25  
160 mM DTT and incubated for 10 min at 56°C. After removal of the solution, the gel pieces were further  
161 treated with 100 mM  $\text{NH}_4\text{HCO}_3$  containing 55 mM iodoacetamide at room temperature and dried  
162 under vacuum as before. The gel bands were then processed for in-gel proteolysis with trypsin in  
163 presence of 0.01% ProteaseMax (Promega) as previously described [32].

### 164 2.4 NanoLC-MS/MS analysis

165 NanoLC–MS/MS experiments for the 24 resulting peptide mixtures were performed using a Q-  
166 Exactive HF mass spectrometer (ThermoFisher) coupled to an UltiMate 3000 LC system (Dionex-LC



167 Packings) in similar conditions as those previously described [33]. Peptide mixtures (10  $\mu$ l) were  
168 loaded and desalted on-line on a reverse phase precolumn (Acclaim PepMap 100 C18, 5  $\mu$ m bead  
169 size, 100  $\text{\AA}$  pore size, 5 mm  $\times$  300  $\mu$ m id) from LC Packings. Peptides were then resolved onto a  
170 reverse phase Acclaim PepMap 100 C18 column (3  $\mu$ m, 100  $\text{\AA}$ , 500 mm  $\times$  75  $\mu$ m id) at a flow rate of  
171 0.3  $\mu$ l/min with a 180 min linear gradient of  $\text{CH}_3\text{CN}/0.1\%$  formic acid and injected into the Q-Exactive  
172 HF mass spectrometer. The Q-Exactive HF instrument was operated according to a Top20 data-  
173 dependent acquisition method as previously described [33]. Briefly, a scan cycle was initiated with a  
174 full scan of peptide ions in the ultra-high-field Orbitrap analyzer, followed by selection of a single  
175 precursor and its dissociation in high energy collisional mode, and MS/MS scans on the 20 most  
176 abundant precursor ions. Full scan mass spectra were acquired with an Automatic Gain Control  
177 Target set at  $3 \times 10^6$  ions and a resolution of 60,000 from  $m/z$  350 to 1,800. MS/MS scan was initiated  
178 at a resolution of 15,000 when the ACG target reached  $1 \times 10^5$  ions with a threshold intensity of  
179 83,000 and potential charge states of 2+ and 3+. Precursor ions were selected with a dynamic  
180 exclusion of 10 sec for increasing the reliability of spectral count measurements.

## 181 2.5 Interpretation of mass spectrometry data

182 The recorded MS/MS spectra for the 24 samples were merged after being searched against the  
183 home-made ORF database. First, peak lists were generated with the MASCOT DAEMON software  
184 (version 2.3.2) from Matrix Science using the extract\_msn.exe data import filter from the Xcalibur FT  
185 package (version 2.0.7) proposed by ThermoFisher. Data import filter options were set as previously  
186 described [34] at 400 (minimum mass), 5000 (maximum mass), 0 (grouping tolerance), 0  
187 (intermediate scans), and 1000 (threshold). MS/MS spectra were searched with MASCOT against the  
188 ORF database with the following parameters: tryptic peptides with a maximum of 2 miss cleavages  
189 during proteolytic digestion, a mass tolerance of 5 ppm on the parent ion and 0.02 Da on the MS/MS,  
190 fixed modification for carbamidomethylated Cys (+57.0215) and variable modification for oxidized  
191 Met (+15.9949). All peptide matches with a peptide score above its query threshold set at  $p \leq 0.05$   
192 with the ORF database and rank 1 were parsed using the IRMa 1.31.1c software [35]. MS/MS spectra

193 assigned to several loci were systematically removed. The average error for the determination of  
194 peptide mass is 0.59 ppm and the mean MASCOT score is 50.7. A protein was considered validated  
195 when at least two different peptides were detected. False-positive identification of proteins was  
196 estimated using a reverse decoy database as below 0.1% with these parameters.

197 The mass spectrometry proteomics data have been deposited in a public repository at the  
198 ProteomeXchange Consortium (<http://proteomecentral.proteomexchange.org>) via the PRIDE partner  
199 repository [36] with the dataset identifier PXD005794 and project DOI 10.6019/PXD005794.

200 The number of MS/MS spectra per protein (spectral counts) was determined for the four replicates  
201 of each of the three time points for both conditions. The protein abundances were compared for  
202 each time point between the uranyl exposure and the control conditions. For each of these  
203 comparisons, the list of non-redundant proteins detected among the six corresponding datasets was  
204 established. The total spectral count of each polypeptide was used to rank the proteins from the  
205 highest to the lowest detection intensities. The statistical protein variation among the four replicates  
206 samples of the two specific exposure conditions compared was calculated using the T-Fold option of  
207 the PatternLab 2.0 software [37]. This module allows normalizing the spectral count datasets,  
208 calculating the average fold changes with statistics (t-test), and estimating the resulting theoretical  
209 false discovery rate. MS/MS data were compiled in Excel (Microsoft) and converted for PatternLab  
210 with an Excel home-designed macro. Normalization was done taking into account the total number  
211 of spectral count for each sample, taking at least two readings per protein. A minimum value of 1 was  
212 added systematically to all spectral count values in order to consider missing values as a standard  
213 PatternLab normalization. Parameters for the comparisons were as follows: minimum fold change of  
214 1.5, minimum p-value of 0.05 and BH-FDR Alfa of 0.15. Normalized spectral abundance factor (NSAF)  
215 for each protein was calculated using the formula: (spectral counts / theoretical molecular weight) x  
216 1000, see [34].

217

## 218 3. Results and discussion

### 219 3.1 *Microbacterium oleivorans* A9 genome annotation

220 The main characteristics of *Microbacterium oleivorans* A9 genome is described in [38]. Briefly, the  
221 genome size is around 2.95 Mbp. The G+C content of the genome is 68.33%. COG assignment for the  
222 resulting polypeptide dataset shows that two classes are predominant: proteins involved in  
223 metabolism of amino acids (74 proteins) on the first hand, and those involved in metabolism of  
224 carbohydrates (66 proteins) on the other. A search against the BacMet database [39], which  
225 comprises antibacterial biocides and metal resistance genes, highlights the presence of 104 unique  
226 metal resistance genes such as *actP* gene which encodes a copper-transporting P-type ATPase, *arsB*  
227 gene which encodes an arsenic pump membrane protein or *yieF* gene which encodes a chromate  
228 reductase. We found a 16S rRNA gene copy in the genome, exhibiting 99% sequence identity with  
229 *Microbacterium oleivorans* 16S rRNA gene. Moreover, the assembling results are strikingly similar to  
230 those obtained for *Microbacterium oleivorans* strain RIT293 (RefSeq NZ\_JFY000000000.1) [40] with a  
231 median total length of 2.9 Mbp, a G+C content of 69% and a total of 2,732 protein-coding genes. This  
232 genome version comprises 11 contigs with a N50 parameter at 467,109 bp.

### 233 3.2 Shotgun proteogenomics of *Microbacterium oleivorans* A9

#### 234 *Proteogenomics quick panorama*

235 **Figure 1** shows the strategy applied here for quickly deciphering the proteome dynamics from resting  
236 cells exposed or not to 10  $\mu$ M uranium. A total of 1,221,282 MS/MS spectra were recorded. Most of  
237 these spectra were of high quality as revealed by the high percentage of assignment obtained  
238 (61.1%) at p-value 0.05. For interpreting MS/MS spectra, a six reading frame translated ORF database  
239 was created. This proteogenomics database comprised a total of 30,853 putative polypeptide  
240 sequences, totaling 4,903,573 amino acids with an average of 159 amino acids per putative  
241 polypeptide and 2,064 as longest length. A total of 746,092 MS/MS spectra could be assigned in first  
242 intention, allowing certifying the presence of 1,504 proteins detected with at least two different

243 peptides (Supplementary data Table S1). Because of the unusual large size of the six reading frame  
244 ORF database(30,853 putative polypeptide sequences instead of 2813 annotated coding DNA  
245 sequences), a second round of search was initiated after reducing the database to the proteins  
246 detected in the first round as proposed previously for improving the sensitivity of proteogenomics  
247 studies [41]. The second search led to the assignment of 747,621 peptide-to-spectrum matches,  
248 certifying the detection of 17,027 unique peptide sequences and 1,532 non-redundant proteins  
249 identified with at least two different peptides (Supplementary data Table S2). This two-round search  
250 allows increasing the number of proteins validated with at least two different peptides (1532 instead  
251 of 1504) and enhances the number of assigned MS/MS spectra (747 621 instead of 746 092). This  
252 dataset compares favorably with previous studies of bacteria exposed to uranium, as a dataset of  
253 less than 950 proteins were reported from *Caulobacter crescentus* [19] and 1,363 proteins were  
254 listed from *Geobacter sulfurreducens* [18]. The proteomic dataset obtained here represents 54%  
255 (including 15% of membrane proteins) of the predicted protein-coding genes found in the draft  
256 genome of *Microbacterium oleivorans* A9 [38]. The distribution of the peptides detected during our  
257 experiment along the chimeric genome of *Microbacterium oleivorans* A9 is uniform and confirm the  
258 validity of this proteogenomics approach, as previously discussed for the proteogenomics mapping of  
259 the alphaproteobacterium *Tistlia consotensis* [22].

#### 260 *Functional categorization of the detected proteins and relative abundances*

261 The dataset of 1,532 identified proteins were classified into 22 COG categories. **Table 1** shows the  
262 functional classification of the mass-spectrometry certified global proteome. Most of the detected  
263 proteins (50%) are distributed in five classes: hypothetical proteins (15%), amino acid metabolism  
264 (11%), carbohydrate metabolism (9%), translation, ribosomal structure, and biogenesis (9%) and  
265 transcription (6%). The remaining 17 classes represent less than 5% each (**Table 1**). The prevalence of  
266 the same classes has been shown in previous studies [18, 19], but here, the large number of  
267 uncharacterized proteins is indicative of an important lack of knowledge regarding soil gram-positive  
268 bacteria. 36 proteins account for 25% of the total proteome detected, and thus are the key molecular

269 draft horses within the cells. These 36 most abundant proteins detected throughout the study as  
270 evaluated with their normalized spectral abundance factor (NSAF) are given in Supplementary data  
271 Table S2. The most abundant protein is the HU bacterial nucleoid DNA-binding protein involved in  
272 DNA structuration. Two of the top five most abundant proteins are the chaperonin GroEL (1.28%)  
273 and the co-chaperonin GroES (1.06%). These two proteins form the binary GroEL/GroES complex,  
274 which mediates proper folding of many proteins. The sub-unit GroEL is specifically known for its  
275 important role both in normal and stressful conditions as well as in the resistance towards toxic  
276 metals [42]. As expected the ribosomal proteins are abundant since the cells were collected during  
277 the exponential growth phase. The 100 most abundant proteins represent 46.1% of total NSAF.  
278 Amongst these, 16 proteins (8.3% of total NSAF) are components of the small ribosomal subunit and  
279 27 proteins (10.6% total NSAF) belong to the large ribosomal subunit. This abundance of ribosomal  
280 proteins has already been observed [23, 34]. Within the most abundant proteins, several are  
281 involved in energy metabolism such as the F<sub>0</sub>F<sub>1</sub>-type ATP synthase (0.78% and 0.57% respectively).

### 282 3.3 Proteome dynamics upon uranium stress

#### 283 *Common differentially produced proteins over the time course of the experiment*

284 Proteins with significant abundance changes (fold change  $\geq 1.5$  and p-value  $\leq 0.05$ ) in uranium-  
285 exposed condition versus control were listed at each time point. A total of 592 proteins met the  
286 criteria: 391 at 0.5h, 294 at 4h and 122 at 24h. The decrease of differentially produced proteins over  
287 time is due to the non-growing exposure conditions. **Figure 2** shows a Venn diagram representation  
288 where the numbers of these proteins are reported for each condition, highlighting the proteins  
289 specific of a given time point and those common in two or three time points. A relatively small set of  
290 24 proteins were shared along the whole kinetic. Remarkably seven of these 24 proteins are  
291 predicted to be ABC-type transporters. Their spectrum activities appeared to be broad as they share  
292 significant sequence similarities with dipeptide/oligopeptide/nickel, Fe<sup>3+</sup>-hydroxamate, multidrug or  
293 enterocholin transporters. A serine protease and a DNA-binding transcriptional regulator belonging

294 to the PadR family are also found more abundant in presence of uranium for the three time points:  
295 1.7, 1.6 and 4.6-fold; 2.1-fold, 2.6-fold and 2.0-fold, respectively. Members of the PadR family  
296 regulate different pathways such as multidrug resistance and detoxification [43].

297 A total of 128 proteins were found significantly more abundant in 0.5h and 4h samples of exposure  
298 to uranium compared to controls. These 128 proteins covered a large variety of molecular functions  
299 making difficult to draw specific uranium responsive pathways. Six ABC-type transporters were  
300 observed among this large panel of up-regulated proteins. They are predicted to be involved in  
301 amino acid, sugar, cobalamin, Fe<sup>3+</sup> siderophore and glycerol-3-phosphate transport. These data  
302 reinforce the hypothesis that bacteria deploy a rather large panel of transport machineries to  
303 respond to uranium exposure. Four proteins involved in cell division are also significantly more  
304 abundant at these two initial phases of uranium exposure: the cell division proteins FtsI and FtsQ,  
305 and two chromosome partitioning ATPases. Such observation reinforces the idea that cells are viable  
306 and active.

307 Once again, a link with iron metabolism is highlighted in the proteins common after 0.5h and 24h of  
308 uranium exposure with the presence of ABC-type Fe<sup>3+</sup>/spermidine/putrescine transport system (2.0  
309 and 1.8-fold), the ferredoxin-NADP reductase (1.9 and 1.9-fold), the deferrochelataase/peroxidase  
310 EfeB (1.7 and 1.8-fold) and the NADPH-dependent ferric siderophore reductase (1.8 and 1.6-fold).

311 Interestingly, a protein predicted to be a DNA-binding transcriptional regulator belonging to the ArsR  
312 family is found significantly more abundant at 0.5h but less abundant after 24h uranium exposure.  
313 This regulator is known to repress expression of stress-induced operons involved in toxic metal ion  
314 tolerance [44]. This suggests that *Microbacterium oleivorans* A9 first represses these genes when  
315 cells are at the initial stage of uranium contact, but need the expression of these genes when the  
316 uranium is intracellularly biomineralized and so less bioavailable.

317 A comparison of the significant changes of protein abundances between 4h and 24h of exposure  
318 shows that 17 proteins are in common for these two time points. Three are related to  
319 osmoregulation (DNA-binding response regulator of the OmpR family and of the AcrR family,

320 maltooligosyltrehalose synthase) which is consistent with the exposure condition used *i.e.* NaCl  
321 0.1M.

### 322 *Proteins specifically modulated at a given time point*

323 A large set of proteins (391) has been found significantly modulated in terms of abundance after 0.5h  
324 exposure to U(VI) as compared to the control. **Table 2** shows the proteins with the most important  
325 changes (fold-change  $\geq 3.5$ ). The chorismate mutase exhibited the highest fold-change (5.8-fold),  
326 while its homologue from *Geobacter sulfurreducens* has been previously shown to decrease in  
327 response to uranium [18]. This protein is involved in phenylalanine and tyrosine synthetic pathway  
328 through the conversion of chorismate in prephenate. As reported in **Figure 3**, five other proteins  
329 from the same pathway have an increased abundance upon uranium exposure, including the  
330 isochorismate synthase (2.6-fold) and the aminotransferase/4-amino-4-deoxychorismate lyase (2.0-  
331 fold). A protein involved in ABC-type glycerol-3-phosphate transport, two proteins responsible of acyl  
332 transfer to glycerol-3-phosphate and the CDP-glycerol glycerophosphotransferase were more  
333 abundant in uranium condition. Glycerol-3-phosphate appears as an important metabolite which  
334 may provide phosphate for complexing and immobilizing uranium [45]. Major changes affected the  
335 phosphate metabolism as evidenced by the important number (59) of proteins related to this  
336 metabolism. A large panel of phosphatases (16 polypeptides) is found in higher amounts after  
337 uranium stress. These enzymes are involved in phosphate removal mainly from sugar (fructose,  
338 trehalose) and from amino acids (serine, threonine).

339 Regarding the two other time points analyzed by high-throughput proteomics, the most modulated  
340 protein detected after 4h is also the protein with the higher abundance after 24h of uranium  
341 exposure compared to control (11-fold and 25-fold respectively). This protein shares some sequence  
342 similarities (63% identity) to the uncharacterized Ykol membrane protein of *Microbacterium*  
343 *oleivorans* LKL04 (**Table 2**). It presents a PepSY domain (pfam 03413) hypothetically involved in  
344 inhibition of a peptidase activity. Based on the TMHMM predictive tool [46], one transmembrane  
345 segment has been identified in the *Microbacterium oleivorans* A9 protein homolog from residues 39

346 to 61, suggesting that the longest part of this protein is extracellular (residues 62 to 247). Regarding  
347 the synteny in *Microbacterium oleivorans* A9 genome, a signal transduction histidine kinase is  
348 adjacent to this protein, indicating a possible involvement in signaling function. One of the four  
349 proteins for which the abundance decreased at 4h of uranium exposure is the glutaredoxin which can  
350 interact physically with a mercury/uranyl reductase (MerA) and plays a role in toxic metals resistance  
351 [47]. After 24h exposure, 12 transporters involved in the transport of diverse compounds  
352 (lipoprotein, Fe<sup>3+</sup>-hydroxamate, dipeptide/oligopeptide/nickel, enterochelin,  
353 Fe<sup>3+</sup>/spermidine/putrescine) are more abundant in the uranium condition compared to the control  
354 (**Table 2**). A DNA-binding response regulator of the NarL/FixJ family is also more detected (5.3-fold).  
355 The homologous regulator from *Escherichia coli* has been characterized and found to control nitrate-  
356 and nitrite-regulated gene expression. This is concordant with the lower detection (-2.1-fold) of a  
357 ferredoxin subunit of the nitrite reductase. In the present exposure condition, nitrate has been  
358 added in both control and uranium conditions and probably contributes to modify the production of  
359 these two proteins.

### 360 *Phosphate metabolism*

361 In our previous study, we have evidenced a phosphate efflux concomitant to the uranium efflux  
362 between 0.5h and 4h exposure and the intracellular formation of autunite, a mineral made of U, P  
363 and Ca, suggesting a strong connection between phosphate and uranium [12]. Among these  
364 proteins, the phosphohistidine phosphatase SixA is the most up-produced protein (3.5 fold-change at  
365 0.5h and 4h). Its precise role is still unclear but its involvement in a signal transduction circuitry  
366 through down-regulation of the ArcB-to-ArcA phosphorelay has been shown under anaerobic  
367 conditions in *Escherichia coli* [56]. On the other hand, the increase abundance of the broad specificity  
368 phosphatase PhoE only after 4h (2.3 fold-change) is congruent with the phosphate efflux observed in  
369 our previous experiment [12]. Several components of an ABC-type glycerol-3-phosphate transport  
370 system have also a significant positive fold-change. No such component is found more abundant  
371 after 24h exposure probably because the phosphate is complexed with uranium leading to autunite



372 accumulation. The glycerol-3-phosphate has been demonstrated as a source of phosphate for the  
373 precipitation of uranium [45]. A polyphosphate kinase shows an up-production after 0.5 h (fold-  
374 change of 1.8). It can act as a defense mechanism by constituting a phosphate reserve that will be  
375 used latter to form uranyl-phosphate complex. As reported in **Figure 4**, the genomic context of this  
376 kinase reveals a locus with seven adjacent genes, possibly organized as an operon. These genes are  
377 all related to phosphate metabolism. The operon comprises a transcriptional regulatory protein, the  
378 phosphohistidine phosphatase SixA and four components of an ABC-type phosphate transport  
379 system. Five of this six genes encode proteins which have been found more abundant in uranium  
380 condition but with low statistical confidence (p-value above 0.05) or with significant but low fold-  
381 change. Noteworthy, this cluster of genes is conserved amongst several *Microbacterium* species from  
382 our collection (data not shown). In this work, several proteins involved in phosphate metabolism  
383 were modulated upon uranium exposure. The relationships between uranium and phosphate  
384 metabolism have also been evidenced by others [57-59].

385 Reminding that uranium can bind to highly oxygenated sites such as phosphate and carbonate, it is  
386 interesting to point out that a component of an ABC-type nitrate/sulfonate/bicarbonate transport  
387 system shows a significant increase in terms of abundance (1.9 fold-change) after an exposure of 4h,  
388 in the same time as uranium and phosphate efflux occur. Several cations transporters ( $K^+$ ,  $Mn^{2+}$ ,  $Zn^{2+}$ ,  
389  $Mg^{2+}$ ,  $Co^{2+}$ ) are also induced upon uranium exposure. One can speculate that one of these  
390 transporters may be involved in uranyl efflux.

### 391 *Iron metabolism*

392 In our previous study and unlike to phosphate metabolism, we obtained no element suggesting an  
393 effect of uranium on iron metabolism [12]. Here, however, a strong impact on iron metabolism is  
394 highlighted through the large number of proteins related to iron metabolism more abundant under  
395 uranyl stress (19 out of 43 proteins classified in the “Inorganic ion metabolism and transport”  
396 category). Most of them are components of the siderophore iron uptake system, either ABC-  
397 transport type subunits or siderophore modification enzymes. Among these 19 proteins, two ABC-

398 type Fe<sup>3+</sup>-hydroxamate transport systems and one ABC-type enterocholin transport system are found  
399 systematically more abundant at the three time points (fold changes of 1.7, 2.1 and 2.8 respectively).  
400 Hydroxamate and enterocholin are siderophores, low molecular compounds produced by bacteria to  
401 scavenge iron (Fe(III)) in iron starvation conditions. The involvement of siderophores in tolerance to  
402 toxic metals has already been demonstrated [48]. The ATPase component of an ABC-type  
403 cobalamin/Fe<sup>3+</sup>-siderophore transport system is also found in higher proportion at 0.5h and 4h with  
404 fold changes of 1.6 and 1.7 respectively. Moreover, the abundance of a NADPH-dependent ferric  
405 siderophore reductase increased upon uranium exposure: fold change of a 1.8 and 1.6 at 0.5h and  
406 24h, respectively. This protein is involved in the release of iron from siderophore. Involvement of  
407 siderophores in the resistance to uranium has been poorly studied, but hydroxamate has been  
408 shown to bind uranium and to be able to chelate more efficiently uranium with the carbonate form  
409 [49]. In addition, induction of siderophore production in response to uranium stress has been shown  
410 for the marine cyanobacterium *Synechococcus elongatus* BDU 130911 [50]. Taken together, these  
411 data suggest that uranium could interact with siderophore in *Microbacterium oleivorans* A9. One  
412 attractive hypothesis is that uranium could then enter the cells via this pathway.

413 Two components (EfeO and EfeB) of the EfeUOB transporter were also detected in higher amounts in  
414 our dataset. This transporter is involved in iron uptake in *Bacillus subtilis* [51]. EfeB is a peroxidase  
415 involved in oxidation of ferrous iron to ferric iron which allows its binding by the EfeO component.  
416 This protein shows a regular increase at 0.5h, 4h, and 24h with fold changes of 1.7, 1.8 and 1.8,  
417 respectively. Moreover, EfeB presents the ability to reduce the reactive oxygen species formed in  
418 presence of ferrous iron and thus, confers protection to the cell. The prediction of both a cupredoxin-  
419 like domain and an imelysin peptidase with a highly conserved HXXE motif confirm that we are  
420 dealing with the EfeO component and not the EfeM for which the cupredoxin-like domain is absent  
421 [52]. The EfeO component is a ferric iron binding protein and is able to transfer its substrate to the  
422 permease EfeU. The variation in the abundance of EfeO during the experiment is not significant. The  
423 EfeU component remained undetected despite its prediction in the genome. The EfeUOB transporter

424 of *Escherichia coli* has been shown to be induced under low pH condition [53]. This three-component  
425 transporter is also found in *Bacillus subtilis* and was shown to be involved in high-affinity uptake of  
426 both ferrous (Fe(II)) and ferric iron (Fe(III)) [51]. Taken together, these results suggest that uranium  
427 exposure is perceived by *Microbacterium oleivorans* A9 as an iron starvation, thus enhancing the  
428 synthesis of iron uptake systems. This result is in line with what has been shown for *Geobacter*  
429 *sulfurreducens* in which uranium induces the Fur operon [18]. Other biological molecules involved in  
430 iron metabolism, e.g. transferrin, have been shown to be able to link uranium and can constitute a  
431 way for uranium to be transported in the cells [55].

#### 432 **4. Concluding remarks**

433 The high-throughput proteogenomic methodology applied to *Microbacterium oleivorans* A9 allows to  
434 quickly identify a large number of proteins and highlights those that could be involved in uranium  
435 tolerance. The proteome coverage obtained here is higher than previous proteomic studies devoted  
436 at exploring the proteome under uranium stress [18, 19]. *Microbacterium oleivorans* A9 modifies  
437 drastically its proteome upon uranium exposure. As previously hypothesized, proteins involved in  
438 phosphate metabolism are found in larger abundances under uranium stress due to higher synthesis  
439 or poorer degradation or both of these mechanisms. More surprisingly, proteins related to iron  
440 metabolism were detected with significant changes in abundance. The detailed relationships  
441 between uranium response and iron metabolism are unknown at the moment. However, based on  
442 our current results it can be hypothesized that uranium may enter the cells via siderophore  
443 transportation. This will be further investigated. The detection of 15% of functionally unassigned  
444 proteins highlights the lack of knowledge regarding proteins with potentially key roles in response to  
445 uranium stress. Focusing on the most interesting protein candidates could be done by a comparative  
446 analysis using high-throughput proteogenomics between *Microbacterium oleivorans* A9 and other  
447 *Microbacterium* isolates differing in uranium sensitivity such as *Microbacterium lemovicicum* ViU22  
448 [60] which is uranium-sensitive.

449

## 450 **Acknowledgments**

451 Nicolas Gallois was financially supported by the Commissariat à l'Energie Atomique et aux Energies  
452 Alternatives (CEA) by means of the program "Thèse Amont-Aval" directed by Yves Bréchet, Haut-  
453 Commissaire à l'Energie Atomique. The project was supported by CEA through the Toxicology  
454 program (BEnUr). We thank Microb&co for all the writing advices given during the 7<sup>th</sup> ICME .

455

## 456 **References**

457

458

- 459 [1] Y. Abe, Y. Iizawa, Y. Terada, K. Adachi, Y. Igarashi, I. Nakai, Detection of uranium and chemical  
460 state analysis of individual radioactive microparticles emitted from the Fukushima nuclear accident  
461 using multiple synchrotron radiation X-ray analyses, *Anal Chem* 86(17) (2014) 8521-8525.
- 462 [2] D. Bugai, V. Kashparov, L. Dewiére, Y. Khomutinin, S. Levchuk, V. Yoschenko, Characterization of  
463 subsurface geometry and radioactivity distribution in the trench containing Chernobyl clean-up  
464 wastes, *Environ Geol* 47(6) (2005) 869-881.
- 465 [3] C.R. Cothorn, W.L. Lappenbusch, J.A. Cotruvo, Health effects guidance for uranium in drinking  
466 water, *Health physics* 44 Suppl 1 (1983) 377-84.
- 467 [4] R.G. Pearson, Hard and Soft Acids and Bases, *J Am Chem Soc* 85(22) (1963) 3533-&.
- 468 [5] R.J. Reeder, M. Nugent, C.D. Tait, D.E. Morris, S.M. Heald, K.M. Beck, W.P. Hess, A. Lanzirotti,  
469 Coprecipitation of uranium(VI) with calcite: XAFS, micro-XAS, and luminescence characterization,  
470 *Geochim Cosmochim Acta* 65(20) (2001) 3491-3503.
- 471 [6] N. Renninger, R. Knopp, H. Nitsche, D.S. Clark, J.D. Keasling, Uranyl precipitation by *Pseudomonas*  
472 *aeruginosa* via controlled polyphosphate metabolism, *Appl Environ Microbiol* 70(12) (2004) 7404-12.
- 473 [7] P. Yong, L.E. Macaskie, Role of citrate as a complexing ligand which permits enzymically-mediated  
474 uranyl ion bioaccumulation, *Bull Environ Contam Toxicol* 54(6) (1995) 892-9.
- 475 [8] M.L. Merroun, S. Selenska-Pobell, Bacterial interactions with uranium: an environmental  
476 perspective, *J Contam Hydrol* 102(3-4) (2008) 285-295.
- 477 [9] A. Nakajima, T. Horikoshi, T. Sakaguchi, Studies on the Accumulation of Heavy-Metal Elements in  
478 Biological-Systems .5. Ion Effects on the Uptake of Uranium by *Chlorella-Regularis*, *Agr Biol Chem*  
479 *Tokyo* 43(3) (1979) 625-629.
- 480 [10] M. Bouby, I. Billard, M. J., I. Rossini, Complexation of uranium VI with the siderophore  
481 pyoverdine, *Radiochim Acta* 80 (1998) 95-100.
- 482 [11] V. Chapon, L. Piette, M.-H. Vesvres, F. Coppin, C.L. Marrec, R. Christen, N. Theodorakopoulos, L.  
483 Février, S. Levchuk, A. Martin-Garin, C. Berthomieu, C. Sergeant, Microbial diversity in contaminated  
484 soils along the T22 trench of the Chernobyl experimental platform, *Appl Geochem* 27(7) (2012) 1375-  
485 1383.
- 486 [12] N. Theodorakopoulos, V. Chapon, F. Coppin, M. Floriani, T. Vercoouter, C. Sergeant, V. Camilleri,  
487 C. Berthomieu, L. Fevrier, Use of combined microscopic and spectroscopic techniques to reveal

488 interactions between uranium and *Microbacterium* sp. A9, a strain isolated from the Chernobyl  
489 exclusion zone, *J Hazard Mater* 285 (2015) 285-293.

490 [13] D.H. Nies, Microbial heavy-metal resistance, *Appl Microbiol Biotechnol* 51(6) (1999) 730-50.

491 [14] P. Junier, E.D. Vecchia, R. Bernier-Latmani, The Response of *Desulfotomaculum reducens* MI-1 to  
492 U(VI) Exposure: A Transcriptomic Study, *Geomicrobiol J* 28(5-6) (2011) 483-496.

493 [15] A. Khemiri, M. Carriere, N. Bremond, M.A. Ben Mlouka, L. Coquet, I. Llorens, V. Chapon, T.  
494 Jouenne, P. Cosette, C. Berthomieu, *Escherichia coli* response to uranyl exposure at low pH and  
495 associated protein regulations, *PLoS one* 9(2) (2014) e89863.

496 [16] L. Dekker, F. Arsene-Ploetze, J.M. Santini, Comparative proteomics of *Acidithiobacillus*  
497 *ferrooxidans* grown in the presence and absence of uranium, *Res Microbiol* 167(3) (2016) 234-239.

498 [17] B. Panda, B. Basu, C. Acharya, H. Rajaram, S.K. Apte, Proteomic analysis reveals contrasting  
499 stress response to uranium in two nitrogen-fixing *Anabaena* strains, differentially tolerant to  
500 uranium, *Aquatic toxicology* 182 (2017) 205-213.

501 [18] R. Orellana, K.K. Hixson, S. Murphy, T. Mester, M.L. Sharma, M.S. Lipton, D.R. Lovley, Proteome  
502 of *Geobacter sulfurreducens* in the presence of U(VI), *Microbiology* 160(Pt 12) (2014) 2607-2617.

503 [19] M.C. Yung, J. Ma, M.R. Salemi, B.S. Phinney, G.R. Bowman, Y. Jiao, Shotgun proteomic analysis  
504 unveils survival and detoxification strategies by *Caulobacter crescentus* during exposure to uranium,  
505 chromium, and cadmium, *J Proteome Res* 13(4) (2014) 1833-1847.

506 [20] J. Armengaud, Next-generation proteomics faces new challenges in environmental  
507 biotechnology, *Curr Opin Biotechnol* 38 (2016) 174-182.

508 [21] A. de Groot, R. Dulermo, P. Ortet, L. Blanchard, P. Guerin, B. Fernandez, B. Vacherie, C. Dossat, E.  
509 Jolivet, P. Siguier, M. Chandler, M. Barakat, A. Dedieu, V. Barbe, T. Heulin, S. Sommer, W. Achouak, J.  
510 Armengaud, Alliance of proteomics and genomics to unravel the specificities of Sahara bacterium  
511 *Deinococcus deserti*, *PLoS genetics* 5(3) (2009) e1000434.

512 [22] C. Rubiano-Labrador, C. Bland, G. Miotello, P. Guerin, O. Pible, S. Baena, J. Armengaud,  
513 Proteogenomic insights into salt tolerance by a halotolerant alpha-proteobacterium isolated from an  
514 Andean saline spring, *J Proteomics* 97 (2014) 36-47.

515 [23] X. Zhu, S. Xie, J. Armengaud, W. Xie, Z. Guo, S. Kang, Q. Wu, S. Wang, J. Xia, R. He, Y. Zhang,  
516 Tissue-specific proteogenomic analysis of *Plutella xylostella* larval midgut using a multialgorithm  
517 pipeline, *Molecular & cellular proteomics* 15(6) (2016) 1791-807.

518 [24] M. Baudet, P. Ortet, J.C. Gaillard, B. Fernandez, P. Guerin, C. Enjalbal, G. Subra, A. de Groot, M.  
519 Barakat, A. Dedieu, J. Armengaud, Proteomics-based Refinement of *Deinococcus deserti* Genome  
520 Annotation Reveals an Unwanted Use of Non-canonical Translation Initiation Codons, *Molecular &*  
521 *Cellular Proteomics* 9(2) (2010) 415-426.

522 [25] C. Rubiano-Labrador, C. Bland, G. Miotello, J. Armengaud, S. Baena, Salt Stress Induced Changes  
523 in the Exoproteome of the Halotolerant Bacterium *Tistlia consotensis* Deciphered by  
524 Proteogenomics, *PLoS one* 10(8) (2015) e0135065.

525 [26] J.T. Simpson, K. Wong, S.D. Jackman, J.E. Schein, S.J. Jones, I. Birol, ABySS: a parallel assembler  
526 for short read sequence data, *Genome Res* 19(6) (2009) 1117-23.

527 [27] J. Armengaud, C. Bland, J. Christie-Oleza, G. Miotello, Microbial proteogenomics, gaining ground  
528 with the avalanche of genome sequences, *J Bacteriol Parasitol* S3-001 (2011).

529 [28] R.L. Tatusov, M.Y. Galperin, D.A. Natale, E.V. Koonin, The COG database: a tool for genome-scale  
530 analysis of protein functions and evolution, *Nucleic Acids Res* 28(1) (2000) 33-6.

531 [29] C. Claudel-Renard, C. Chevalet, T. Faraut, D. Kahn, Enzyme-specific profiles for genome  
532 annotation: PRIAM, *Nucleic Acids Res* 31(22) (2003) 6633-9.

533 [30] M. Kanehisa, Y. Sato, M. Kawashima, M. Furumichi, M. Tanabe, KEGG as a reference resource for  
534 gene and protein annotation, *Nucleic Acids Res* 44(D1) (2016) D457-62.

535 [31] E.M. Hartmann, F. Allain, J.C. Gaillard, O. Pible, J. Armengaud, Taking the shortcut for high-  
536 throughput shotgun proteomic analysis of bacteria, *Methods Mol Biol* 1197 (2014) 275-285.

537 [32] G. Clair, S. Roussi, J. Armengaud, C. Duport, Expanding the known repertoire of virulence factors  
538 produced by *Bacillus cereus* through early secretome profiling in three redox conditions, *Molecular &*  
539 *cellular proteomics* 9(7) (2010) 1486-98.

540 [33] G. Klein, C. Mathe, M. Biola-Clier, S. Devineau, E. Drouineau, E. Hatem, L. Marichal, B. Alonso,  
541 J.C. Gaillard, G. Lagniel, J. Armengaud, M. Carriere, S. Chedin, Y. Boulard, S. Pin, J.P. Renault, J.C.  
542 Aude, J. Labarre, RNA-binding proteins are a major target of silica nanoparticles in cell extracts,  
543 *Nanotoxicology* 10(10) (2016) 1555-1564.

544 [34] J.A. Christie-Oleza, B. Fernandez, B. Nogales, R. Bosch, J. Armengaud, Proteomic insights into the  
545 lifestyle of an environmentally relevant marine bacterium, *ISME J* 6(1) (2012) 124-35.

546 [35] V. Dupierriis, C. Masselon, M. Court, S. Kieffer-Jaquinod, C. Bruley, A toolbox for validation of  
547 mass spectrometry peptides identification and generation of database: IRMa, *Bioinformatics* 25(15)  
548 (2009) 1980-1.

549 [36] R.G. Cote, J. Griss, J.A. Dianes, R. Wang, J.C. Wright, H.W. van den Toorn, B. van Breukelen, A.J.  
550 Heck, N. Hulstaert, L. Martens, F. Reisinger, A. Csordas, D. Ovelleiro, Y. Perez-Rivevol, H. Barsnes, H.  
551 Hermjakob, J.A. Vizcaino, The PRoteomics IDentification (PRIDE) Converter 2 framework: an  
552 improved suite of tools to facilitate data submission to the PRIDE database and the  
553 ProteomeXchange consortium, *Molecular & cellular proteomics* 11(12) (2012) 1682-9.

554 [37] P.C. Carvalho, J. Hewel, V.C. Barbosa, J.R. Yates Iii, Identifying differences in protein expression  
555 levels by spectral counting and feature selection, *Genet Mol Res* 7(2) (2008) 342-356.

556 [38] P. Ortet, N. Gallois, J. Long, M. Barakat, V. Chapon, Draft genome sequence of *Microbacterium*  
557 *oleivorans* strain A9, a bacterium isolated from Chernobyl radionuclide-contaminated soil, *Genome*  
558 *announcement* 999(5) (2017) e00092-17.

559 [39] C. Pal, J. Bengtsson-Palme, C. Rensing, E. Kristiansson, D.G. Larsson, BacMet: antibacterial  
560 biocide and metal resistance genes database, *Nucleic Acids Res* 42(Database issue) (2014) 737-743.

561 [40] H.Y. Gan, H.M. Gan, M.A. Savka, A.J. Triassi, M.S. Wheatley, L.B. Smart, E.S. Fabio, A.O. Hudson,  
562 Whole-genome sequences of 13 endophytic bacteria isolated from shrub willow (*salix*) grown in  
563 geneva, new york, *Genome announcement* 2(3) (2014) e00288-14.

564 [41] P. Jagtap, J. Goslinga, J.A. Kooren, T. McGowan, M.S. Wroblewski, S.L. Seymour, T.J. Griffin, A  
565 two-step database search method improves sensitivity in peptide sequence matches for  
566 metaproteomics and proteogenomics studies, *Proteomics* 13(8) (2013) 1352-7.

567 [42] J.S. Weissman, H.S. Rye, W.A. Fenton, J.M. Beechem, A.L. Horwich, Characterization of the active  
568 intermediate of a GroEL-GroES-mediated protein folding reaction, *Cell* 84(3) (1996) 481-90.

569 [43] G. Fibriansah, A.T. Kovacs, T.J. Pool, M. Boonstra, O.P. Kuipers, A.M. Thunnissen, Crystal  
570 structures of two transcriptional regulators from *Bacillus cereus* define the conserved structural  
571 features of a PadR subfamily, *PloS one* 7(11) (2012) e48015.

572 [44] S. Silver, T. Phung le, A bacterial view of the periodic table: genes and proteins for toxic inorganic  
573 ions, *J Ind Microbiol Biotechnol* 32(11-12) (2005) 587-605.

574 [45] R.J. Martinez, M.J. Beazley, M. Taillefert, A.K. Arakaki, J. Skolnick, P.A. Sobecky, Aerobic uranium  
575 (VI) bioprecipitation by metal-resistant bacteria isolated from radionuclide- and metal-contaminated  
576 subsurface soils, *Environ Microbiol* 9(12) (2007) 3122-33.

577 [46] A. Krogh, B. Larsson, G. von Heijne, E.L. Sonnhammer, Predicting transmembrane protein  
578 topology with a hidden Markov model: application to complete genomes, *Journal of molecular*  
579 *biology* 305(3) (2001) 567-80.

580 [47] C. Cassier-Chauvat, F. Chauvat, Responses to oxidative and heavy metal stresses in  
581 cyanobacteria: recent advances, *Int J Mol Sci* 16(1) (2015) 871-886.

582 [48] M. Rajkumar, N. Ae, M.N.V. Prasad, H. Freitas, Potential of siderophore-producing bacteria for  
583 improving heavy metal phytoextraction, *Trends Biotechnol* 28(3) (2010) 142-149.

584 [49] K.F. Mo, Z. Dai, D.S. Wunschel, Production and characterization of desmalonichrome relative  
585 binding affinity for uranyl ions in relation to other siderophores, *J Nat Prod* (2016) 1492-9.

586 [50] V. Rashmi, M. Shylajanaciyar, R. Rajalakshmi, S.F. D'Souza, D. Prabakaran, L. Uma, Siderophore  
587 mediated uranium sequestration by marine cyanobacterium *Synechococcus elongatus* BDU 130911,  
588 *Bioresour Technol* 130 (2013) 204-210.

589 [51] M. Miethke, C.G. Monteferrante, M.A. Marahiel, J.M. van Dijl, The *Bacillus subtilis* EfeUOB  
590 transporter is essential for high-affinity acquisition of ferrous and ferric iron, *Biochim Biophys Acta*  
591 1833(10) (2013) 2267-2278.

592 [52] M.B. Rajasekaran, S.A. Mitchell, T.M. Gibson, R. Hussain, G. Siligardi, S.C. Andrews, K.A. Watson,  
593 Isolation and characterisation of EfeM, a periplasmic component of the putative EfeUOBM iron  
594 transporter of *Pseudomonas syringae* pv. *syringae*, *Biochem Bioph Res Co* 398(3) (2010) 366-371.  
595 [53] J. Cao, M.R. Woodhall, J. Alvarez, M.L. Cartron, S.C. Andrews, EfeUOB (YcdNOB) is a tripartite,  
596 acid-induced and CpxAR-regulated, low-pH Fe<sup>2+</sup> transporter that is cryptic in *Escherichia coli* K-12  
597 but functional in *E. coli* O157:H7, *Mol Microbiol* 65(4) (2007) 857-75.  
598 [54] B.D. Stewart, R.T. Amos, S. Fendorf, Effect of uranium(VI) speciation on simultaneous microbial  
599 reduction of uranium(VI) and iron(III), *J Environ Qual* 40(1) (2011) 90-97.  
600 [55] M. Hemadi, N.T. Ha-Duong, J.M. El Hage Chahine, Can uranium be transported by the iron-  
601 acquisition pathway? Ur uptake by transferrin, *J Phys Chem B* 115(14) (2011) 4206-4215.  
602 [56] M. Matsubara, T. Mizuno, The SixA phospho-histidine phosphatase modulates the ArcB  
603 phosphorelay signal transduction in *Escherichia coli*, *FEBS letters* 470(2) (2000) 118-24.  
604 [57] L.E. Macaskie, G. Basnakova, Microbially-enhanced chemisorption of heavy metals: a method for  
605 the bioremediation of solutions containing long-lived isotopes of neptunium and plutonium, *Environ*  
606 *Sci Technol* 32(1) (1998) 184-7.  
607 [58] M.J. Beazley, R.J. Martinez, P.A. Sobecky, S.M. Webb, M. Taillefert, Uranium biomineralization as  
608 a result of bacterial phosphatase activity: Insights from bacterial isolates from a contaminated  
609 subsurface, *Environ Sci Technol* 41(16) (2007) 5701-5707.  
610 [59] S. Kulkarni, C.S. Misra, A. Gupta, A. Ballal, S.K. Apte, Interaction of Uranium with Bacterial Cell  
611 Surfaces: Inferences from Phosphatase-Mediated Uranium Precipitation, *Appl Environ Microbiol*  
612 82(16) (2016) 4965-74.  
613 [60] L. Mondani, L. Piette, R. Christen, D. Bachar, C. Berthomieu, V. Chapon, *Microbacterium*  
614 *lemovicicum* sp. nov., a bacterium isolated from a natural uranium-rich soil, *Int J Syst Evol Microbiol*  
615 63(Pt 7) (2013) 2600-6.

616

617

618 **Table 1. Functional classification of the proteins detected by at least two peptides.**

<b>COG category</b>	<b>Relative abundance (%)</b>	<b>Number of proteins</b>
Hypothetical protein	14,6	223
Amino acid transport and metabolism	11,0	168
Translation, ribosomal structure and biogenesis	9,3	142
Carbohydrate transport and metabolism	8,8	135
Transcription	6,2	95
Energy production and conversion	5,5	84
Coenzyme transport and metabolism	5,3	82
General function	5,1	78
Lipid transport and metabolism	4,4	67
Nucleotide transport and metabolism	4,3	66
Posttranslational modification, protein turnover, chaperones	4,0	61
Cell wall/membrane/envelope biogenesis	3,9	60
Replication, recombination and repair	3,9	60
Signal transduction mechanisms	3,6	55
Function unknown	2,8	43
Inorganic ion transport and metabolism	2,8	43
Defense mechanism	1,3	20
Cell cycle control, cell division, chromosome partitioning	1,1	17
Secondary metabolites metabolism	1,0	16
Cell motility	0,5	8
Intracellular trafficking, secretion, and vesicular transport	0,5	8
RNA processing and modification	0,1	1

619



Table 2 - Key proteins with major changes after uranium exposure

ORFeome ID <sup>a</sup>	Functional annotation	Functional classification	Fold change <sup>b</sup>	p value	Best hit <sup>c</sup>
<b>After 0.5h</b>					
Contig10_25943_-2	Chorismate mutase	Amino acid metabolism	5.75	0.16E-03	CHMU_METJA
Contig1_343955_-2	Alpha-glucosidase, glycosyl hydrolase family GH31	Carbohydrate metabolism	4.50	0.02E-03	YCR2_ESCVU
Contig6_5891_-2	Acetyl-CoA acetyltransferase	Lipid metabolism	4.00	7.25E-06	BKTB_CUPNH
Contig11_20212_1	Beta-galactosidase/beta-glucuronidase	Carbohydrate metabolism	3.85	0.02E-03	BGAL_ARTSB
Contig3_31861_1	Sugar kinase of the NBD/HSP70 family	Carbohydrate metabolism	3.80	0.02E-03	XYLR_BACPZ
Contig3_278050_-1	DNA-binding transcriptional regulator, PadR family	Transcription	3.75	0.20E-03	YWZG_BACSU
Contig7_116064_-3	Dipeptidyl aminopeptidase/acylaminoacyl peptidase	Amino acid metabolism	3.75	0.15E-03	DAPB3_PSEMX
Contig8_12690_-3	tRNA A37 threonylcarbamoyladenine modification protein TsaB	Translation, ribosome	3.67	0.09E-03	Y3455_MYCBO
Contig2_261531_-3	Phosphohistidine phosphatase SixA	Signal transduction	3.50	0.10E-03	Y1276_MYCTO
Contig1_23131_1	Serine protease, subtilisin family	PTM, protein turnover	3.50	0.05E-03	HLY_NATA1
Contig8_9448_1	Hypothetical protein	Hypothetical	3.50	0.03E-03	Y3037_MYCTO
Contig1_334362_3	ABC-type glycerol-3-phosphate transport system	Carbohydrate metabolism	3.50	2.97E-03	YURO_BACSU
<b>After 4h</b>					
Contig11_87308_2	Uncharacterized membrane protein Ykol	Function unknown	10.75	3.99E-06	No results
Contig1_430924_1	D-arabinose 1-dehydrogenase, Zn-dependent alcohol dehydrogenase family	Carbohydrate metabolism	4.60	9.21E-03	ADH_SULTO
Contig3_147088_1	DNA-binding transcriptional regulator/sugar metabolism, DeoR/GlpR family	Transcription	4.58	1.72E-04	AGAR_ECO57
Contig5_94994_2	Hypothetical protein	Hypothetical	3.83	1.64E-03	No results
Contig11_20081_2	6-phosphogluconolactonase/Glucosamine-6-P isomerase/deaminase	Carbohydrate metabolism	3.80	2.68E-02	NAGB_STRCO
Contig5_120376_1	Hypothetical protein	Hypothetical	3.75	1.57E-02	No results
Contig5_105896_-2	Adenosylmethionine-8-amino-7-oxononanoate aminotransferase	Coenzyme metabolism	3.55	2.92E-03	Y3329_MYCTU
Contig3_278050_-1	DNA-binding transcriptional regulator, PadR family	Transcription	3.50	5.59E-03	YWZG_BACSU
Contig2_261531_-3	Phosphohistidine phosphatase SixA	Signal transduction	3.50	2.28E-02	Y1276_MYCTO
<b>After 24h</b>					
Contig11_87308_2	Uncharacterized membrane protein Ykol	Function unknown	25	1.15E-04	No results
Contig5_20481_3	ABC-type lipoprotein export system	Cell wall/membrane	8.18	7.26E-07	Y1508_METJA
Contig5_94994_2	Hypothetical protein	Hypothetical	5.25	6.46E-03	No results
Contig5_161786_-2	DNA-binding response regulator, NarI/FixJ family	Signal transduction	5.25	4.78E-04	DEGU_BREBE
Contig1_136842_3	Periplasmic serine protease, S1-C subfamily	PTM, protein turnover	4.33	8.18E-04	DEGP_SALTY
Contig3_196395_-3	Hypothetical protein	Hypothetical	4.28	2.74E-03	TITIN_MOUSE
Contig15_39811_-1	DNA-binding transcriptional regulator, AcrR family	Transcription	3.75	4.04E-03	MIFTR_MYCTO

<sup>a</sup> ORFeome ID refers to Supplementary Data Table S2.

<sup>b</sup> Fold changes represent the expression with uranium exposure normalized to expression without uranium.

<sup>c</sup> The best hit correspond to the first hit obtained by BLASTp against Swissprot database.

622 **Figure 1 – Shotgun nanoLC-MS/MS strategy used for the proteome analysis of *Microbacterium***  
623 ***oleivorans* A9 under uranium exposure or not.** A total of 24 samples were processed: four biological  
624 replicates for each time point at 0,5h, 4h and 24h of exposure. After lysis, the proteins were  
625 subjected to a short SDS-PAGE migration before trypsin proteolysis in-gel. A Q-Exactive HF mass  
626 spectrometer was used to identify the resulting peptides.

627 **Figure 2 – Venn diagram showing the number of proteins with significant abundance changes in**  
628 **the three sampling time points.**

629 **Figure 3 – Example of the phenylalanine, tyrosine and tryptophan biosynthesis pathways.** The  
630 boxes represent the branched metabolic pathways. The enzymes are identified through their EC  
631 number. Proteins detected by proteogenomics are highlighted in red, proteins more abundant upon  
632 uranium exposure are indicated in blue and proteins predicted in the genome but not detected by  
633 proteomics are in grey. The genes' ID corresponding to the enzymes are as follow: 1.3.1.12:  
634 MicroBact\_A9DB#Contig4\_204950\_-2, 2.4.2.18: MicroBact\_A9DB#Contig4\_125788\_-1, 2.5.1.19:  
635 MicroBact\_A9DB#Contig2\_13726\_-1, 2.5.1.54: MicroBact\_A9DB#Contig4\_169845\_-3, 2.6.1.1:  
636 MicroBact\_A9DB#Contig12\_70604\_2, 2.6.1.9: MicroBact\_A9DB#Contig7\_122729\_2, 2.7.1.71:  
637 MicroBact\_A9DB#Contig1\_401394\_-3, 4.1.1.48: MicroBact\_A9DB#Contig1\_371158\_-1, 4.2.1.10:  
638 MicroBact\_A9DB#Contig1\_399335\_-2, 4.2.1.20: MicroBact\_A9DB#Contig1\_369105\_-3, 4.2.1.51:  
639 MicroBact\_A9DB#Contig5\_42426\_3, 4.2.1.91: MicroBact\_A9DB#Contig3\_67122\_3, 4.2.3.5:  
640 MicroBact\_A9DB#Contig1\_401926\_-1, 5.4.99.5: MicroBact\_A9DB#Contig10\_25943\_-2, 6.1.1.1:  
641 MicroBact\_A9DB#Contig7\_15747\_3, 6.1.1.20: MicroBact\_A9DB#Contig7\_133127\_2.

642 **Figure 4 – Genomic context of the polyphosphate kinase.** Fold-change (FC) are indicated for the first  
643 time point of sampling (0.5h). Proteins for which the gene name is in bold were detected by  
644 proteomics.

645

646

647 **List of supplementary material**

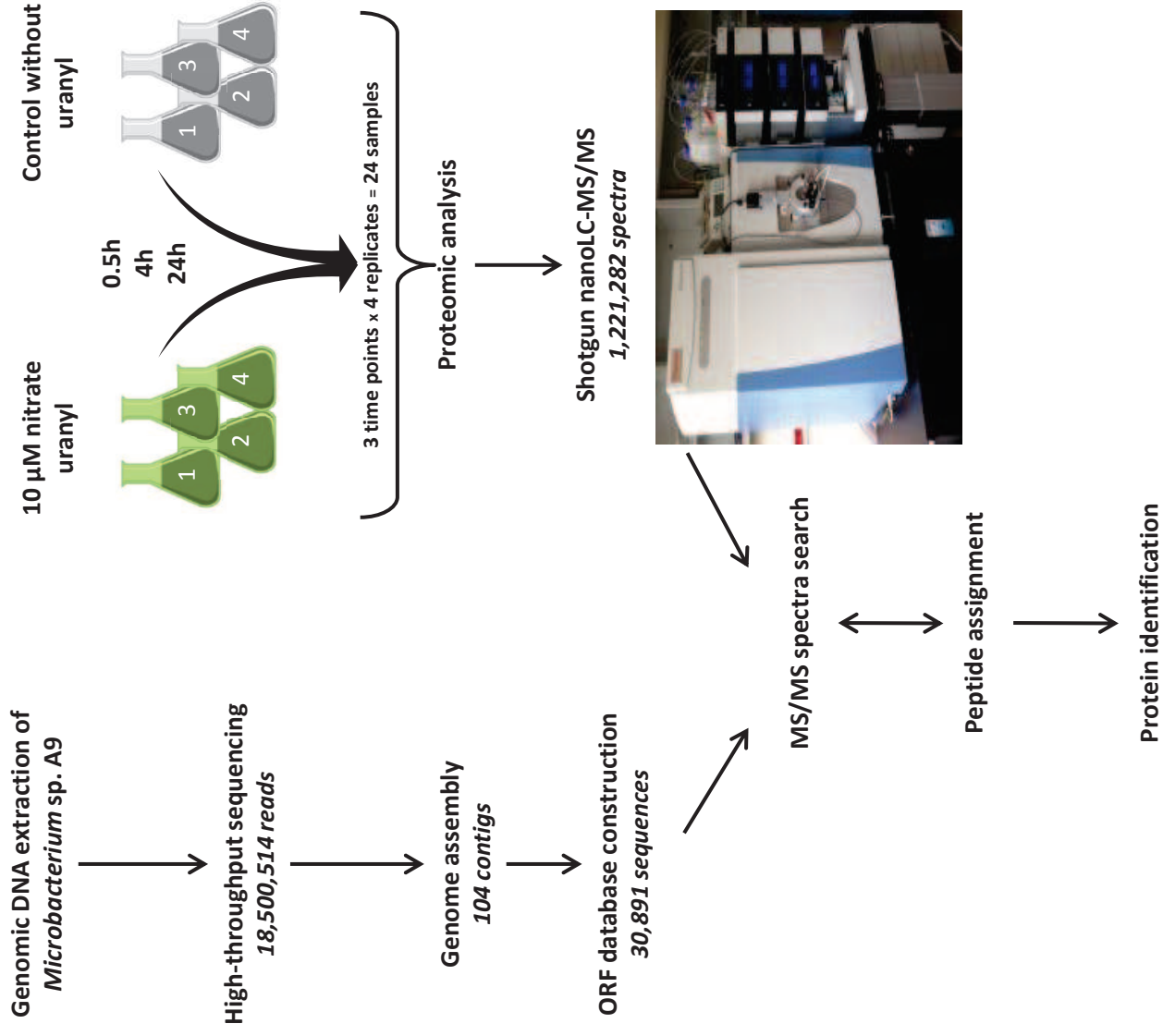
648 **Supplementary Table S1.** List of proteins and their spectral count in A9 samples in the first  
649 proteogenomic search.

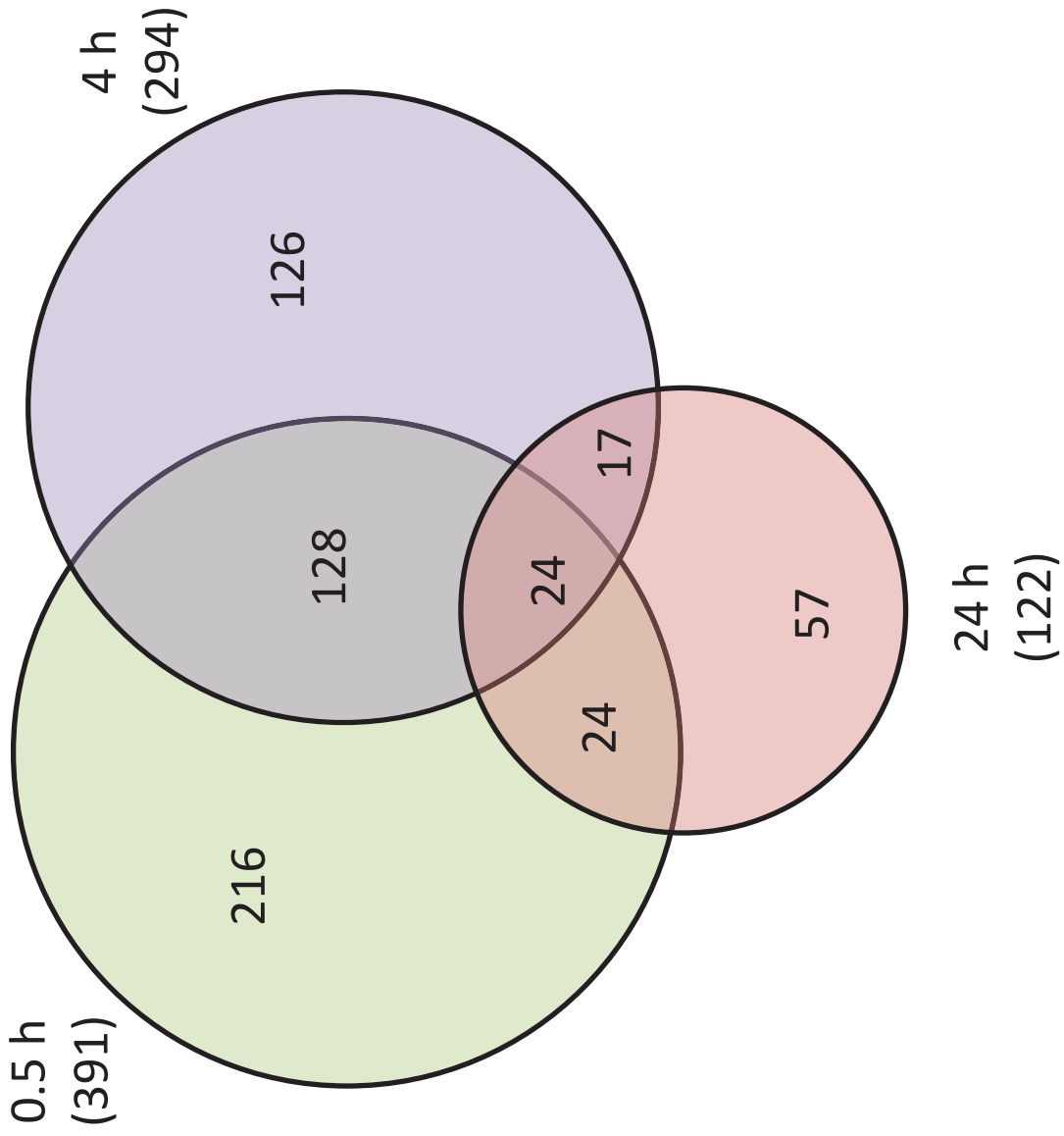
650 **Supplementary Table S2.** List of proteins and their spectral count in A9 samples in the  
651 second round search. NSAF: Normalized Spectral Abundance Factor; SC: Spectral Count

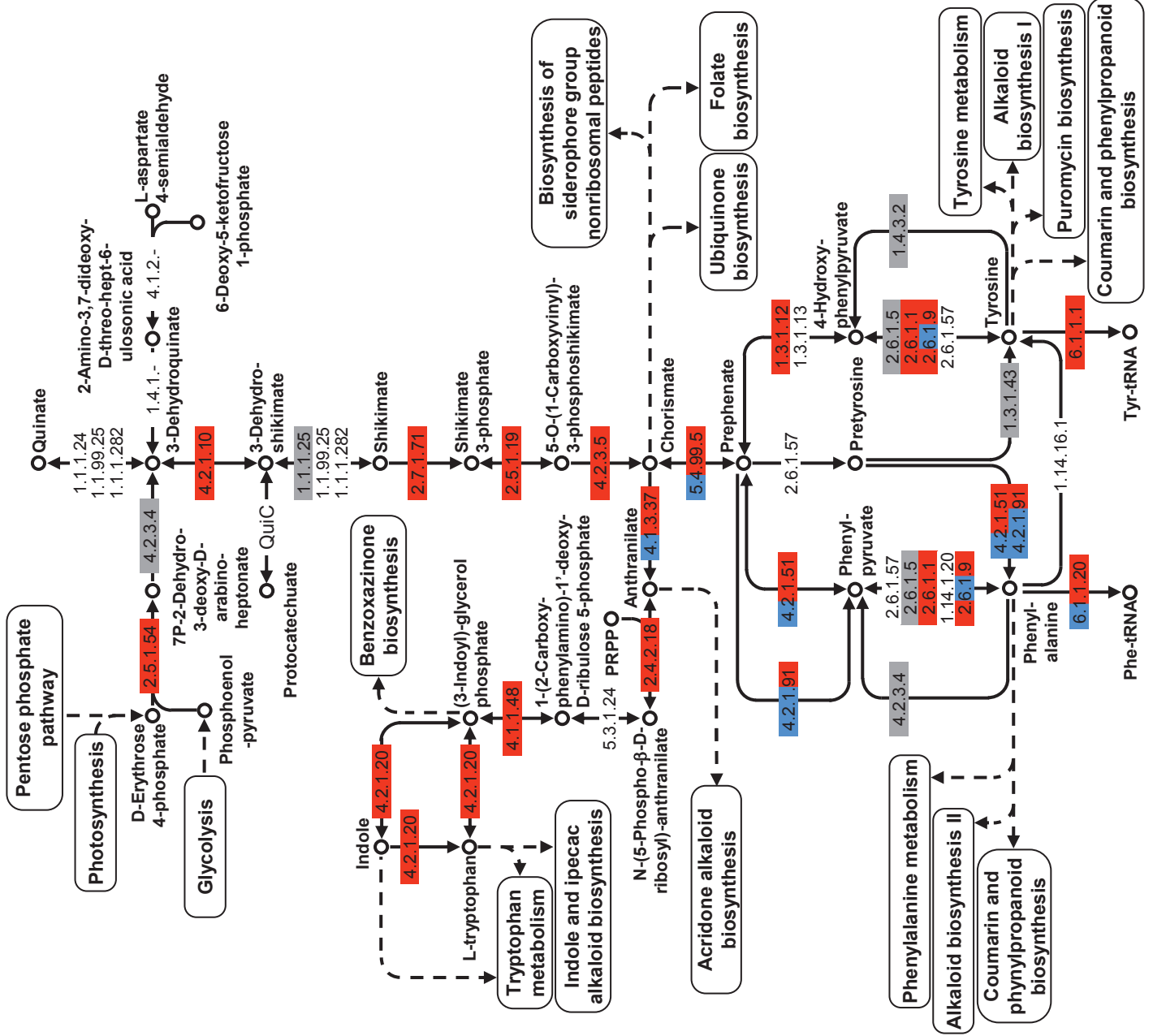
652

**Figure**

[Click here to download Figure: Proteogenomics of Microbacterium sp. A9\\_Four figures revised version.pptx](#)







## *Microbacterium oleivorans* A9 chromosome

