ORIGINAL ARTICLE

ISME Journal – Original article

Metaproteogenomic insights beyond bacterial response to naphthalene exposure and bio-stimulation

María-Eugenia Guazzaroni, Florian-Alexander Herbst, Iván Lores, Javier Tamames, Ana Isabel Peláez, Nieves López-Cortés, María Alcaide, Mercedes V. del Pozo, José María Vieites, Martin von Bergen, José Luis R. Gallego, Rafael Bargiela, Arantxa López-López, Dietmar H. Pieper, Ramón Rosselló-Móra, Jesús Sánchez, Jana Seifert and Manuel Ferrer

Supporting Online Material includes

Text (Supporting Materials and Methods)

Tables S1 to S9

Figures S1 to S7

SUPPORTING TEXT

Supporting Materials and Methods

Soil characterisation

Soil pH was measured in a suspension of soil and water (1:2.5) with a glass electrode, and electrical conductivity was measured in the same extract (diluted 1:5). Primary soil characteristics were determined using standard techniques, such as dichromate oxidation (organic matter content), the Kjeldahl method (nitrogen content), the Olsen method (phosphorus content) and a Bernard calcimeter (carbonate content). The Bouyoucos Densimetry method was used to establish textural data. Exchangeable cations (Ca, Mg, K and Na) extracted with 1 M NH_4Cl and exchangeable aluminium extracted with 1 M KCl were determined using atomic absorption/emission spectrophotometry with an AA200 PerkinElmer analyser. The effective cation exchange capacity (ECEC) was calculated as the sum of the values of the last two measurements (sum of the exchangeable cations and the exchangeable Al). Analyses were performed immediately after sampling.

Hydrocarbon analysis

Extraction (5 g of sample N and Nbs) was performed with dichloromethane:acetone (1:1) using a Soxtherm extraction apparatus (Gerhardt GmbH & Co. KG, Köln, Germany), followed by silica gel purification. The final extract was analysed via GCMS (Shimadzu QP2010Plus equipment) using a Factor Four VF-5 ms (5% phenyl 95% dimethylpolysiloxane) column (30 m x 0.25 mm, 0.25 μm) from Varian. The initial temperature was set to 70°C (maintained for 2 min), and the temperature was increased by 20°C/min up to 220°C and then by 10°C/min up to 300°C (maintained for 7 min). The injector temperature was maintained at 260°C working in splitless mode. Helium gas (purity of 99.999%) was used at a flow rate of 1 ml/min. Detection was performed in the EI (electronic ionisation) mode, and the ionisation source and the transfer line were maintained at 220°C and 260°C, respectively. For the quantification of metabolites, a selective ion monitoring mode (SIM) was used, selecting the *m*/*z* ratio characteristics of the PAHs of interest. Analyses were performed immediately after sampling.

DNA extraction

Six grams of soil N and Nbs were processed for DNA extraction. Highly purified suspension of soil microorganisms was obtained via density gradient centrifugation using Nycodenz (Axis-Shield PoC, Norway) resin, as described previously (Guazzaroni *et al.*, 2010). The resulting cell pellet was subjected to metagenomic DNA extraction using a Meta-Gnome® DNA extraction Kit (Epicentre, Madison, WI, USA), which incorporates mechanical lysis. The obtained DNA was visualised using 0.7% (wt/vol) agarose gel electrophoresis and quantified both spectrophotometrically and using PicoGreen (Molecular Probes, Carlsbad, CA). For enrichment cultures, 100 ml of the culture was centrifuged at 4000 g x for 15 min, and DNA was extracted using a Meta-Genome® DNA extraction Kit (Epicentre). Note: top soil sample N was sampled in February 2007 and stored at -20°C until DNA was isolated in March 2010; top soil sample Nbs was sampled 231 days later than sample N and stored at -20°C until DNA was isolated in February 2012; DNA from enrichment cultures CN1 and CN2 were extracted in March 2010 from stable cultures maintained by using serial batch culturing.

Construction of 16S RNA gene clone libraries, sequencing and phylogenetic analysis

Bacterial 16S RNA genes were amplified using the bacteria-specific primers F27 (5'-AGAGTTTGATCMTGGCTCAG-3') and R1492 (5'-CGGYTACCTTGTTACGACTT-3'). Amplification was performed in a 20 µl reaction volume with recombinant *Taq* DNA polymerase (Invitrogen, Germany) and reagents according to the basic PCR protocol at an annealing temperature of 45°C for 30 cycles. PCR amplicons were purified using electrophoresis on 0.8% (wt/vol) agarose gels, followed by isolation from excised bands using a QIAEX II Gel Extraction Kit (Qiagen, Germany). The purified PCR products were ligated into pGEM plasmid vectors (pGEM Cloning kit, Promega, Madison, USA), which were subsequently transformed into electrocompetent cells of *E. coli* (TOP 10) (Invitrogen, Germany). The obtained bacterial rDNA clones were sequenced via primer walking using the M13 forward (5'-GACGTTGTAAAACGACGGCCAG-3') and M13 reverse (5'-GAGGAAACAGCTATGACCATG-3') primers according to the protocol of the BigDye Terminator v3.1 Cycle Sequencing Kit from Applied Biosystems (USA). The sequencing reactions were analysed using AB 3730 xl equipment from Applied Biosystems (USA).

Phylogenetic inference of the 16S RNA gene sequences was carried out using the ARB software package (Ludwig *et al.*, 2004). The sequences were automatically aligned using the SINA aligner (E. Pruesse, unpublished) against the SILVA SSURef 102 (Pruesse *et al.*, 2007) and LTP s108 (Yarza *et al.*, 2008) reference alignments and were manually inspected to correct inaccurately situated bases. To improve resolution at lower taxonomic levels, two independent reference phylogenetic trees were constructed, one comprised of members of the phylum *Proteobacteria* and a second including the remaining bacterial phyla. The distinct datasets with nearly complete SSU sequences were first sieved with a 40% conservational filter, and the phylogeny was then reconstructed using the neighbour-joining algorithm with the Jukes-Cantor correction. The resulting tree topologies were carefully checked against the currently accepted classification of *Prokaryotes* (LPSN, http://www.bacterio.cict.fr) to verify the absence of incongruent phylogenetic relationships.

In addition to the clone affiliations, the partial sequences extracted from the metagenome of the sample Nbs were checked. From the >300 sequences identified in the metagenome, we just used those with a length >600 nucleotides. Shorter sequences were discarded as being not enough informative. Partial sequences were aligned as above mentioned, and inserted into the pre-existing Neighbour Joining tree using the ARB parsimony tool (Ludwig et al., 2004).

In order to recognize biologically meaningful units we grouped phylogenetic clades into Operational Phylogenetic Units (OPUs). We considered an OPU (López-López *et al.*, 2010) to be represented by each single group of clones that formed an independent clade in the tree without considering a rigid similarity cut-off value. As it can be seen in the supplementary **Table S2**, nearly all OPUs shared internal identity >97%, a value that nearly guarantees each of them to be considered a single species from the taxonomic point of view (Yarza et al., 2008). The OPUs were plotted in a rarefaction curve (**Figure S4**). Statistical analyses were performed using the PAST program (Hammer *et al*., 2001).

DGGE analysis

The V3–V5 variable regions of the 16S rDNA gene were amplified using 16F341-GC and 16R907 primers (Yu and Morrison, 2004). The F341-GC primer included a GC clamp at the 5′ end (5′-CGCCCGCCGCGCCCCGCGCCCGTCCCGCCGCCCCCGCCCG-3′). The PCR profile used with the F341GC/R907 primer set consisted of an initial denaturation of 5 min at 94°C, followed by 25 cycles of one denaturation step for 1 min at 94°C, one annealing step for 1 min at 57°C, and one elongation step for 45 s at 72°C. The last step involved an extension for 10 min at 72°C. Approximately 400 ng of purified PCR product was loaded onto a 6% (w/v) polyacrylamide gel (1 mm thick), with denaturing gradients ranging from 40% to 65% (100%

denaturant contains 7 M urea and 40% formamide (w/v)). The DGGE analysis was performed in a 0.5× Tris-acetate-ethylenediaminetetraacetic acid (TAE) buffer solution (40 mM Tris, 20 mM sodium acetate, 1 mM EDTA, pH 7.4) using a DCode Universal Mutation Detection System (BioRad Laboratories, Inc., CA). DGGE gels were stained for 45 min in water solution containing SybrGold (Molecular Probes, Inc., Eugene, OR, USA) and then scanned under UV using Gel Logic 200 Imaging Sytem (Kodak, Bionova, Canada). Individual bands were excised from the denaturing gradient gels and eluted in water (Martínez-Pascual *et al*., 2010), and some of the obtained products were subjected to sequencing.

Metagenomic setup: sequencing, assembly and gene prediction and analysis

Sequencing was performed using a Roche 5 GS FLX DNA sequencer (454 Life Sciences) (Life Sequencing S.L, Valencia, Spain) as described previously (Ferrer *et al*., 2011). Eleven micrograms of total DNA at an average concentration of 286 ng/μl were used for sequencing of the samples.

The reads from each sample were individually assembled to generate non-redundant metasequences using Newbler GS De Novo Assembler v.2.3 (Roche). An estimated error rate (incorrect bases/total number of expected nucleotides) of 0.49% was considered for the GS20 reads (Huse *et al*., 2007). The error rates for the GS20 reads were calculated using the Needleman-Wunsch algorithm (Needleman and Wunsch, 1970). The MetaGene (Noguchi *et al*., 2006) and MetaGeneMark programs, which are based on a hidden Markov model (HMM) algorithm, were employed to predict potential protein-coding regions (open reading frames, ORFs with ≥ 20 amino acids) from the sequences of each sample. When both predictions were coincident, the corresponding ORF start and end positions were kept. When the predictions did not agree, homology searching via BLASTX helped to resolve the inconsistencies by identifying the most complete gene. The start and end positions for these genes were adequately set. Homology searching was also used to infer genes that were not detected by the gene predictors. The complete sequences of the contigs were analysed via BLASTX searching against the GenBank non-redundant (nr) database, and nucleotide and amino acid sequences for the putative genes and proteins obtained were retrieved.

Each gene was ascribed to COG (Tatusov *et al*., 2001) and KEGG (Kanehisa *et al*., 2008) functional categories. The following procedure was used for this: homology searches were performed with the nucleotide sequences of the contigs using BLASTN and BLASTX against the GenBank nr database. The obtained hits were filtered based on a minimum e-value ($1e^{-03}$) and an alignment length of 75% minimum. The accession number of the hits was used to retrieve COG, KEGG and SEED annotations if available. The maximum expected bitscore for the gene was also obtained via self-alignment and was used to normalise all bitscores for the hits. We took the first five hits of each COG (Tatusov *et al*., 2001), KEGG (Kanehisa *et al*., 2008) or SEED (Overseek *et al*., 2005) and calculated its average normalised score (ANS). A gene was assigned to a functional category that corresponded to an ANS at least 20% higher than any other. If the ANS difference was less than 20%, no function was assigned. Our tests indicated that this procedure yielded more than 95% correct assignments (based on the application of a number of tests performed using controlled sets of sequences assigned to COGs, KEGG and SEED ffrom completely sequenced genomes). Functional assignment of the predicted genes involved in degradation was made on the basis of BLASTP analysis against an in-house database that contained gene sequences encoding enzymes that usually perform key catalysing steps in the aerobic degradation of pollutants via di- and trihydroxylated intermediates (**Table S8**). These enzymes included Rieske non-heme iron oxygenases, type I extradiol dioxygenases of the vicinal oxygen chelate superfamily, the type II or LigB superfamily extradiol dioxygenases, and the type III extradiol dioxygenases belonging to the cupin superfamily as well as intradiol dioxygenases (Pérez-Pantoja *et al*., 2011). Well-characterised key oxygenase sequences that have been documented as being involved in the catabolism of aromatics were used as initial seeds for similarity searches using the BLASTP program from the NCBI website (Altschul *et al.*, 1997). Translated protein sequences were aligned with MUSCLE using default values (Edgar, 2004). Phylogenetic trees were constructed with MEGA4 (Tamura *et al*., 2007) using the neighbour-joining algorithm (Saitou and Nei, 1987). The phylogenetic trees were inspected to detect proteins with similarities to proteins with documented functions.

Taxonomic assignment was performed in the same way as functional assignment. The complete taxonomy of the homologs was retrieved, and the ANS for each of the taxa at different phylogenetic ranks (kingdom, phylum, class, order, family and genus) was computed. Homology searches were also used in the recruitment analysis for particular genomes. Complete genome sequences were obtained from NCBI and used to map the contigs with the best hit against that particular genome. Common numbers of sequences between the different samples were obtained by clustering the sequences with the cd-hit program using different levels of identity. To identify potential metabolic pathways, genes were evaluated for similarity compared to the KEGG database. A match was counted if the similarity analysis resulted in an expectation E value below $1e^{-05}$. All obtained KEGG orthology (KO) numbers were mapped against KEGG pathway functional hierarchies and statistically analysed.

For calculating the over-representation of proteins and functions between metagenomes, we followed the z test for independent proportions proposed by Li (2009). Briefly, given N_A, N_B (the number of observations of a particular protein or function in metagenomes A and B), H_A, and H_B (the total number of observations in metagenomes A and B), there are three occurrence rates $P_A = H_A/N_A$, $P_B = H_B/N_B$, and $P = (H_A+H_B)/(N_A+N_B)$. The statistical significance between A and B can be described by:

$$z = \frac{P_A - P_B}{\sqrt{P(1-P)(\frac{1}{N_A}+\frac{1}{N_B})}}$$

Confidence levels of 0.95 and 0.99 are defined by *z* values of 1.96 and 2.58, and minimum H_A values of 4 and 7, respectively (Li, 2009).

The complete set of gene annotation features including, ORF ID, contig, contig length, GC content per contig, GC content per ORF, ORF length and ORF start and end positions, putative function, best hit, closest hit and putative phylogeny as well as COG, KEGG and SEED annotations are shown in details in **Table S9**.

Metaproteomic setup: protein extraction, separation and identification and data processing

Cells (triplicates) were harvested from the same enrichment cultures CN1 and CN2 (60 transfers) used also for DNA isolation, via centrifugation at 10 000 × g for 10 min. The pellets were suspended in 4 ml of the BugBuster® protein extraction reagent by vortexing. Then, 1 mg/ml of lysozyme and DNAse I were added, and the samples were incubated for 30 min on ice. Subsequently, 10 µl of lysonase was added, and the samples were incubated for 60 min on ice. Bacterial suspensions were subsequently ultra-sonicated for 2.5 min on ice, after which the extract was centrifuged for 10 min at 12 000 g to separate cell debris and intact cells. The supernatant was carefully aspirated (to avoid disturbing the pellet) and transferred to a new tube, and the protein extracts were then subjected to acetone precipitation (5 vol of acetone per 1 vol of sample solution) and stored at -86°C until use. Protein concentrations were determined by resuspending the precipitated proteins in 50 mM Tris/HCl (pH 7.0) and using the Bradford assay (Bradford, 1976). Then, 14 µg of three replicates of each the CN1 and CN2 samples was subjected to 1-D SDS PAGE, and the separated proteins were stained with Coomassie Blue. Ten sections of each lane were excised following SDS-PAGE, and the sections were de-stained and subsequently proteolytically digested overnight at 37°C using trypsin (Sigma, Munich, Germany) as previously described (Jehmlich *et al*., 2008). The eluted peptides were purified and concentrated using C18 Zip Tip columns (Millipore).

For LTQ-Orbitrap mass spectrometer analysis, the peptides were reconstituted in 0.1% formic acid. The samples were injected using the autosampler and concentrated in a trapping column (nanoAcquity UPLC column, C18, 180 µm x 2 cm, 5 µm, Waters) with water containing 0.1% formic acid at a flow rate of 15 µL min^{-1}. After 6 min, the peptides were eluted into a separation column (nanoAcquity UPLC column, C18, 75 µm x 10 cm, 1.75 µm, Waters). Chromatography was performed with 0.1% formic acid in solvents A (100% water) and B (100% acetonitrile). To elute the peptides, the solvent B gradient was set at 2 to 20% in the first 54 min and subsequently at 20 to 85% for 28 min using a nano-high pressure liquid chromatography (nano-HPLC) system (nanoAcquity, Waters) coupled with an LTQ-Orbitrap mass spectrometer (Thermo Fisher Scientific). To achieve an unbiased analysis, continuous scanning of the eluted peptide ions was carried out between 300 – 2000 *m/z*, automatically switching to MS/MS collision-induced dissociation (CID) mode for ions that exceeded an intensity of 3000. For MS/MS CID measurements, a dynamic precursor exclusion of 3 min was enabled.

For identification and retrieval of protein intensities, MaxQuant version 1.1.1.25 was used (Cox and Mann, 2008). Analyses of all fractions were performed simultaneously, such that ten gel slices per replicate were combined into one experiment (2 samples x 3 replicates x 10 gel slices). Peptide identification was performed using Andromeda implemented in MaxQuant (Cox *et al.*, 2011) against the CN1 and CN2 pyro-sequences in the form of forward and reverse entries as decoys. Cysteine carbamidomethylation was searched for as a fixed modification, whereas methionine oxidation was searched for as a variable modification. The maximal false discovery rate was set to 1% for the peptides. The offered *first search* was performed with a mass tolerance of 20 ppm. Proteins were regarded as convincingly identified if at least two unique peptides were found. In addition to activating the *match between runs* with a time window of 2 min and a unique peptide count of at least 2, the standard MaxQuant settings were used.

The raw protein intensities corresponded to the sum of the peptide intensities. Thus, larger proteins, which contain significantly more tryptic peptides, have higher intensities. To eliminate the influence of the number of peptide hits, the protein intensities were divided by the number of peptides, which were dedicated to the protein of interest. These *per peptide* intensities were summed for every replicate, and the averages of the *summed* intensities of the three replicates for CN1 and CN2 were calculated. The average total intensities for CN1 and CN2 were divided by the corresponding *summed* replicate intensities. The obtained quotient served as a normalisation factor. To correct for different protein loads, all *per peptide* intensities were multiplied by their normalisation factor. If normalised *per peptide* intensities based on at least two peptides could be established in at least two replicates, the average was calculated and reported as the quantitative value for the protein of interest.

References

Altschul SF, Madden TL, Schäffer AA, Zhang J, Zhang Z, Miller W, Lipman, DJ. (1997). Gapped BLAST and PS I-BLAST: a new generation of protein database search programs. *Nucleic Acids Res* **25:** 3389-3402.

Bradford MM. (1976). A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal Biochem* **72:** 248–254.

Cox J, Mann M. (2008). MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nat Biotechnol* **26:** 1367–1372.

Cox J, Neuhauser N, Michalski A, Scheltema RA, Olsen JV, Mann M. (2011). Andromeda - a peptide search engine integrated into the MaxQuant environment. *J Proteome Res* **10:** 1794–1805.

Edgar RC. (2004). MUSCLE: a multiple sequence alignment method with reduced time and space complexity. *BMC Bioinformatics* **5:** 113.

Ferrer M, Guazzaroni ME, Richter M, García-Salamanca A, Yarza P, Suárez-Suárez A. (2011). Taxonomic and functional metagenomic profiling of the microbial community in the anoxic sediment of a sub-saline shallow lake (Laguna de Carrizo, Central Spain). *Microb Ecol* **62:** 824–837.

Guazzaroni M-E, Golyshin PN, Ferrer M. (2010). Molecular methods to study complex microbial communities. In: Marco, D (ed). *Metagenomics: Theory, Methods and Applications.* Caister Academic Press (Portland), pp 55–77.

Hammer Ø, Harper D, Ryan P. (2001). PAST: Paleontological statistics software package for education and data analysis. *Palaeontologia Electronica* **4:** 1–9.

Huse SM, Huber JA, Morrison HG, Sogin ML, Weich DM (2007) Accuracy and quality of massively parallel DNA pyrosequencing. *Genome Biol* **8:** R143.

Jehmlich N, Schmidt F, Hartwich M, von Bergen M, Richnow HH, Vogt C. (2008). Incorporation of carbon and nitrogen atoms into proteins measured by protein-based stable isotope probing (Protein-SIP). *Rapid Commun Mass Spectrom* **22:** 2889–2897.

Kanehisa M, Araki M, Goto S, Hattori M, Hirakawa M, Itoh M *et al*. (2008). KEGG for linking genomes to life and the environment. *Nucleic Acids Res* **36:** D480–D484.

Li W. (2009). Analysis and comparison of very large metagenomes with fast clustering and functional annotation. *BMC Bioinformatics* **10:** 359.

Ludwig W, Strunk O, Westram R, Richter L, Meier H, Yadhukumar *et al*. (2004) ARB: a software environment for sequence data. *Nucleic Acids Res* **32:** 1363–1371.

Needleman SB, Wunsch CD. (1970). A general method applicable to the search for similarities in the amino acid sequence of two proteins. *J Mol Biol* **48:** 443–453.

Noguchi H, Park J, Takagi T. (2006). MetaGene: prokaryotic gene finding from environmental genome shotgun sequences. *Nucleic Acids Res* **34:** D5623–5630.

Overbeek R, Begley T, Butler RM, Choudhuri JV, Chuang HY, Cohoon M, *et al*. (2005). The subsystems approach to genome annotation and its use in the project to annotate 1000 genomes. *Nucleic Acids Res* **33:** 5691-5702.

Pruesse E, Quast C, Knittel K, Fuchs BM, Ludwig W, Peplies J, Glöckner FO. (2007). SILVA: a comprehensive online resource for quality checked and aligned ribosomal RNA sequence data compatible with ARB. *Nucleic Acids Res* **35:** 7188-7196.

Saitou N, Nei M. (1987). The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol Biol Evol* **4:** 406–425.

Schloss PD, Handelsman J. (2005). Introducing DOTUR, a computer program for defining operational taxonomic units and estimating species richness. *Appl Environ Microbiol* **71:** 1501–1506.

Tamura K, Dudley J, Nei M, Kumar S. (2007). MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. *Mol Biol Evol* **24:** 1596–1599.

Tatusov RL, Natale DA, Garkavtsev IV, Tatusova TA, Shankavaram UT, Rao BS *et al*. (2001). The COG database: new developments in phylogenetic classification of proteins from complete genomes. *Nucleic Acids Res* **29:** 22–28.

Yu Z, Morrison M (2004). Comparisons of different hypervariable regions of rrs genes for use in fingerprinting of microbial communities by PCR-denaturing gradient gel electrophoresis. *Appl Environ Microbiol* **70:** 4800–4806

SUPPORTING TABLES

Table S1 Initial soil characteristics. All the values are an average of three measurements on distinct replicas of the original soil sample N.

Texture	Sand (2 mm – 63 µm)	43%
	Silt (63 – 2 µm)	19%
	Clay (< 2 µm)	39%
Chemical parameters[1]	pH	8.2
	Conductivity	0.13 mS/cm
	Organic Matter	0.20 % w/w
	Total Nitrogen	0.03 % w/w
	C/N ratio	3.9
	Phosphorus	<2.6 mg/kg
	Carbonate Content	0.72 % w/w
	Potassium	64 mg/kg
	Magnessium	80 mg/kg
	Calcium	3160 mg/kg
	Sodium	19 mg/kg

[1] Soil characteristics and chemical parameters were determined using standard techniques described in the **Supplementary Materials and Methods**.

Table S2 Prokaryotic diversity and microbial composition blueprints of N, Nbs, CN1 and CN2 communities. Tentative phylogenetic affiliations are specifically shown.

OPUS CN1[1]	**OPUS CN2**[1]	**OPUS N**[1]	**OPUS Nbs 16S rDNA clones**[1]	**OPUS Nbs 454 partial sequences**[1]	**OPUS Nbs total**[1]	**OPU**[1]	**Clones**[1]	**Minimal internal identity**[2]	**Closest related species**[2]	**Phylum**[2]	**Family**[2]
	48	**1**	**12**	**12**	**24**	**1**	N-144, CN2-28, CN2-72, CN2-58, CN2-78, CN2-15, CN2-26, CN2-99, CN2-42, CN2-21, CN2-02, CN2-07, CN2-12, CN2-46, CN2-74, CN2-108, CN2-16, CN2-71, CN2-106, CN2-114, CN2-68, CN2-56, CN2-102, CN2-63, CN2-64, CN2-81, CN2-110, CN2-76, CN2-10, CN2-112, CN2-09, CN2-65, CN2-104, CN2-05, CN2-11, CN2-31, CN2-06, CN2-48, CN2-91, CN2-111, CN2-60, CN2-83, CN2-103, CN2-82, CN2-107, CN2-24, CN2-36, CN2-88, CN2-35, CN2-88, Nbs-59, Nbs-61, Nbs-63, Nbs-30, Nbs-70, Nbs-19, Nbs-89, Nbs-25, Nbs-69, Nbs-84, Nbs-10, Nbs-21, Nbs261, Nbs222, Nbs202, Nbs123, Nbs271, Nbs238, Nbs106, Nbs 165, Nbs120, Nbs201, Nbs212, Nbs147	99.2	*Pseudomonas stutzeri* GV 1	*Gammaproteobacteria*	*Pseudomonadaceae*
			1	**2**	**3**	**2**	Nbs-14, Nbs110, Nbs209	0	*Pseudomonas alcaligenes*	*Gammaproteobacteria*	*Pseudomonadaceae*
		6	**1**	**3**	**4**	**3**	N-154, N-75, N-178, N-25, N-68, N-216, N-173, Nbs-15, Nbs223, Nbs270, Nbs241	99.8	*Pseudomonas stutzeri* GV 3	*Gammaproteobacteria*	*Pseudomonadaceae*
	1		**1**	**7**	**8**	**4**	CN2-40, Nbs-73, Nbs158, Nbs260, Nbs236, Nbs105, Nbs, 176, Nbs118, Nbs174	99.8	*Pseudomonas* sp.	*Gammaproteobacteria*	*Pseudomonadaceae*
		5	**2**	**8**	**10**	**5**	N-110, N-55, N-128, N-151, N-119, Nbs-93, Nbs-75, Nbs115, Nbs254, Nbs224, Nbs, 107, Nbs256, Nbs150, Nbs177, Nbs275	99.2	*Pseudomonas stutzeri* M14C	*Gammaproteobacteria*	*Pseudomonadaceae*
			6	**15**	**21**	**6**	Nbs76, Nbs-48, Nbs-11, Nbs-88, Nbs-64, Nbs-2, Nbs252, Nbs230, Nbs233, Nbs145, Nbs272, Nbs221 Nbs253, Nbs269, Nbs251, Nbs231, Nbs264, Nbs152, Nbs207, Nbs136, Nbs243		*Pseudomonas tuomuerensis*	*Gammaproteobacteria*	*Pseudomonadaceae*
		1				**7**	N-179	0	*Pseudomonas argentinensis*	*Gammaproteobacteria*	*Pseudomonadaceae*
			2	**6**	**8**	**8**	Nbs-24, Nbs-72, Nbs114, Nbs193, Nbs166, Nbs226, Nbs204, Nbs183	98.6	*Pseudomonas pertucinogena*	*Gammaproteobacteria*	*Pseudomonadaceae*
			3	**17**	**20**	**9**	Nbs-12, Nbs-53, Nbs-27, Nbs101, Nbs242, Nbs205, Nbs219, Nbs245, Nbs140, Nbs217, Nbs157, Nbs253, Nbs239, Nbs168, Nbs186, Nbs195, Nbs139, Nbs279, Nbs151, Nbs210	99.1	*Pseudomonas* sp.	*Gammaproteobacteria*	*Pseudomonadaceae*
		4	**1**		**1**	**10**	N-103, N-74, N-33, N-02, Nbs-44	99.4	*Pseudomonas guineae*	*Gammaproteobacteria*	*Pseudomonadaceae*

		65				**11**	N-120, N-152, N-130, N-170, N-116, N-212, N-159, N-57, N-10, N-177, N-34, N-208, N-86, N-81, N-89, N-100, N-95, N-98, N-72, N-183, N-50, N-35, N-194, N-119, N-142, N-37, N-44, N-01, N-77, N-211, N-47, N-105, N-111, N-118, N-166, N-24, N-69, N-61, N-49, N-195, N-60, N-87, N-181, N-82, N-141, N-205, N-171, N-53, N-201, N-218, N-123, N-190, N-13, N-140, N-129, N-200, N-156, N-174, N-16, N-91, N-143, N-199, N-97, N150, N-210	99.2	*Pseudomonas amygdali*	*Gammaproteobacteria*	*Pseudomonadaceae*
		1				**12**	N-193	0	*Pseudomonas brassicacearum*	*Gammaproteobacteria*	*Pseudomonadaceae*
		22				**13**	N-28, N-94, N-146, N-113, N-180, N-101, N-121, N-07, N-56, N-67, N-21, N-71, N-20, N-09, N-26, N-39, N-169, N-42, N-163, N-167, N-22, N-153	99.2	*Pseudomonas putida*	*Gammaproteobacteria*	*Pseudomonadaceae*
			1	**2**	**3**	**14**	Nbs-86	0	*Pseudomonas* sp.	*Gammaproteobacteria*	
		1			**0**	**15**	N-05	0	*Azotobacter vinelandii*	*Gammaproteobacteria*	*Pseudomonadaceae*
			1	**2**	**3**	**16**	Nbs-90	0	*Pseudomonas* sp.	*Gammaproteobacteria*	
		1				**17**	N-46	0	*Alkanidiges illinoisensis*	*Gammaproteobacteria*	*Moraxellaceae*
11						**18**	CN1-108, CN1-40, CN1-86, CN1-52, CN1-72, CN1-91, CN1-110, CN1-84, CN1-88, CN1-45,CN1-14, CN1-104, CN1-78, CN1-02, CN1-17, CN1-34, CN1-31, CN1-24, CN1-24, CN1-89, CN1-50, CN1-92, CN1-64, CN1-82, CN1-83, CN1-59	98.9	*Legionella lytica*	*Gammaproteobacteria*	*Legionellaceae*
	15					**19**	CN2-29, CN2-98, CN2-01, CN2-59, CN2-89, CN2-44, CN2-33, CN2-34, CN2-113, CN2-57, CN2-13, CN2-80, CN2-109, CN2-54, CN2-86	99.3	*Pseudomonas geniculata*	*Gammaproteobacteria*	*Pseudomonadaceae*
		12				**20**	N-114, N-133, N-70, N-132, N-90, N-32, N-172, N-17, N-165, N-96, N-40, N-168	98.7	*Stenotrophomonas chelatiphaga*	*Gammaproteobacteria*	*Xanthomonadaceae*
		8				**21**	N-36, N-59, N-43, N-160, N-196, N-52, N-64, N-164	99.3	*Stenotrophomonas rhizophila*	*Gammaproteobacteria*	*Xanthomonadaceae*
		28	**1**	**3**	**4**	**22**	N-88, N-107, N-108, N-109, N-187, N-176, N-137, N-213, N-148, N-182, N-185, N-139, N-79, N-27, N-149, N-158, N-204, N-186, N-162, N-65, N-203, N-136, N-92, N-115, N-126, N62, N84, N117, Nbs-22, Nbs278, Nbs104, Nbs218	97.1	*Pseudoxanthomonas spadix*	*Gammaproteobacteria*	*Xanthomonadaceae*
6						**23**	CN1-19, CN1-94, CN1-87, CN1-77, CN1-79, CN1-09,	99.5	*Pseudoxhanthomonas japonensis*	*Gammaproteobacteria*	*Xanthomonadaceae*
		2				**24**	N-198, N-04	99.6	*Pseudoxanthomonas yeongjuensis*	*Gammaproteobacteria*	*Xanthomonadaceae*
		1				**25**	N-03	0	*Pseudoxhanthomonas broegbernens*	*Gammaproteobacteria*	*Xanthomonadaceae*
		1				**26**	N-15	0	*Thermomonas lysobacter oryzae*	*Gammaproteobacteria*	*Xanthomonadaceae*
			1		**1**	**27**	Nbs-3	0	*Lysobacter sp.*	*Gammaproteobacteria*	

7	**29**			**1**	**1**	**28**	CN1-97, CN1-101, CN1-55, CN1-30, CN1-95, CN1-80, CN1-85, CN2-41, CN2-90, CN2-03, CN2-55, CN2-101, CN2-67, CN2-105, CN2-49, CN2-39, CN2-66, CN2-69, CN2-93, CN2-87, CN2-30, CN2-79, CN2-62, CN2-51, CN2-18, CN2-37, CN2-92, CN2-97,CN2-77, CN2-32, CN2-75, CN2-70, CN2-50, CN2-95, CN2-27, CN2-84, Nsb160	99.3	*Achromobacter spanius*	*Betaproteobacteria*	*Alcaligenaceae*
	1					**29**	CN2-17	0	*Pigmentiphaga litoralis*	*Betaproteobacteria*	*Alcaligenaceae*
		1		**1**	**1**	**30**	N-99	0	*Achromobacter xylosoxidans*	*Betaproteobacteria*	*Alcaligenaceae*
			8	**3**	**11**	**31**	Nbs-46, Nbs-74, Nbs-26, Nbs-54, Nbs-77, Nbs-45, Nbs-60, Nbs-96, Nbs171, Nbs112, Nbs170	99.1	*Tetrathiobacter kashmirensis*	*Betaproteobacteria*	
			2	**1**	**3**	**32**	Nbs-20, Nbs-67, Nbs135	96.2	*Tetrathiobacter sp.*	*Betaproteobacteria*	*Alcaligenaceae*
	7					**33**	CN2-43, CN2-19, CN2-61, CN2-73, CN2-20, CN2-22, CN2-25	99.8	*Diaphorobacter nitroreducens*	*Betaproteobacteria*	*Comamonadaceae*
	1		**3**	**9**	**12**	**34**	CN2-47, Nbs-16, Nbs-78, Nbs4, Nbs203, Nbs134, Nbs228, Nbs108, Nbs198, Nbs122, Nbs125, Nbs126, Nbs116	0	*Acidovorax defluvii*	*Betaproteobacteria*	*Comamonadaceae*
		1				**35**	N-122	0	*Acidovorax valerianellae*	*Betaproteobacteria*	*Comamonadaceae*
10						**36**	CN1-66, CN1-22, CN1-100, CN1-35, CN1-62, CN1-21, CN1-32, CN1-75, CN1-20, CN1-67	99.9	*Comamonas composti*	*Betaproteobacteria*	*Comamonadaceae*
		1				**37**	N-11	0	*Aquicella siphonis*	*Gammaproteobacteria*	*Coxiellaceae*
		1				**38**	N-18	0	*Hydrocarboniphaga effusa*	*Gammaproteobacteria*	*Sinobacteraceae*
			1		**1**	**39**	Nbs-95		*Hydrocarboniphaga sp.*	*Gammaproteobacteria*	
		4				**40**	N-209, N-112, N-189, N-12	99.5	*Rhizobium cellulosilyticum*	*Alphaproteobacteria*	*Rhizobiaceae*
		1				**41**	N-41	0	*Rhizobium radiobacter*	*Alphaproteobacteria*	*Rhizobiaceae*
1						**42**	CN1-05	0	*Mesorhizobium septentrionale*	*Alphaproteobacteria*	*Phyllobacteriaceae*
		1				**43**	N-54, N-104	98.1	*Kaistia adipata*	*Alphaproteobacteria*	*Rhizobiaceae*
4						**44**	CN1-01, CN1-73, CN1-103, CN1-12	99.7	*Devosia insulae*	*Alphaproteobacteria*	*Hyphomicrobiaceae*
		1				**45**	N-131	0	*Pseudolabrys taiwanensis*	*Alphaproteobacteria*	*Xanthobacteraceae*
	1					**46**	CN2-14	0	*Anderseniella baltica*	*Alphaproteobacteria*	*Rhodobiaceae*
		1	**1**	**1**	**2**	**47**	N-217, Nbs-36, Nbs247	0	*Sphingomonas yanoikuyae*	*Alphaproteobacteria*	*Sphingobacteriaceae*
		1				**48**	N-73	0	*Sphingobium aromaticiconverten*	*Alphaproteobacteria*	*Sphingobacteriaceae*
		3				**49**	N-191, N-66, N-145	99	*Novosphingobium resinovorum*	*Alphaproteobacteria*	*Sphingobacteriaceae*

		2				**50**	N-29, N-63	98.3	*Sphingomonas sanxanigenens*	*Alphaproteobacteria*	*Sphingobacteriaceae*
		3				**51**	N-93, N-147, N-51	99.3	*Brevundimonas aurantiaca*	*Alphaproteobacteria*	*Caulobacteraceae*
		1				**52**	N-184	0	*Phenylobacterium haematophilum*	*Alphaproteobacteria*	*Caulobacteraceae*
		1				**53**	N-102	0	*Rhodobacter veldkampii*	*Alphaproteobacteria*	*Rhodobacteraceae*
		1				**54**	N-188	0	*Rhodobacter changlensis*	*Alphaproteobacteria*	*Rhodobacteraceae*
		2				**55**	N-175, N-80	99.7		*Alphaproteobacteria*	*Incertae sedis* (no assigned to any family)
35	**1**					**56**	CN1-11, CN1-16, CN1-39, CN1-26, CN1-63, CN1-54, CN1-107, CN1-41, CN1-43, CN1-60, CN1-44, CN1-70, CN1-98, CN1-58, CN1-25, CN1-29, CN1-51, CN1-53, CN1-76, CN1-42, CN1-47, CN1-109, CN1-85, CN1-48, CN1-102, CN1-33, CN1-18, CN1-06, CN1-96, CN1-08, CN1-53, CN1-23, CN1-69, CN1-56, CN1-03, CN1-28, CN2-53	99.2	*Azospirillum oryzae*	*Alphaproteobacteria*	*Rhodospirillaceae*
7						**57**	CN1-38, CN1-61, CN1-90, CN1-65, CN1-71, CN1-04, CN1-74	99.2	Uncultured *alpha*	*Alphaproteobacteria*	Uncultured *alpha*
1						**58**	CN1-93	0	*Caedibacter caryophilus*	*Alphaproteobacteria*	*Thiotrichales*
		1				**59**	N-08	0	*Sinorickettsia chlamys*	*Alphaproteobacteria*	Uncultured *alpha*
	1					**60**	CN2-04	0	*Endosymbiont Acanthamoeba*	*Alphaproteobacteria*	Uncultured *alpha*
			2		**2**	**61**	Nbs-40, Nbs-35	99.8	*Uncultured alpha*	*Alphaproteobacteria*	Uncultured *alpha*
		1				**62**	N-161	0	*Roseomonas terrae*	*Alphaproteobacteria*	*Rhodobacteraceae*
		1				**63**	N-58	0	*Arcobacter mytili*	*Epsilonproteobacteria*	*Campylobacteraceae*
		1	**1**	**2**	**3**	**64**	N-207, Nbs87, Nbs191, Nbs131	0	*Delsulfocapsa thiozymogenes*	*Deltaproteobacteria*	*Desulfobulbaceae*
		1		**1**	**1**	**65**	N-206	0	*Geobacter thiogenes*	*Deltaproteobacteria*	*Geobacteraceae*
			2	**3**	**5**	**66**	Nbs-28, Nbs-32, Nbs199, Nbs154, Nbs225	99	*Soehngenia saccharolytica*	*Firmicutes*	*Clostridiales*
			3	**3**	**6**	**67**	Nbs-13, Nbs-85, Nbs-29, Nbs156, Nbs109, Nbs263	97.6	*Sedimentibacter saalensis*	*Firmicutes*	*Clostridiales*
		1				**68**	N-83	0	*Parvimonas micra*	*Firmicutes*	*Incertae sedis* (no assigned to any family)
			1		**1**	**69**	Nbs-23	0	Uncultured *Firmicutes*	*Firmicutes*	Uncultured *Firmicutes*
			2		**2**	**70**	Nbs-58, Nbs-91	100	Uncultured *Firmicutes*	*Firmicutes*	Uncultured *Firmicutes*

		9				**71**	N-30, N-38, N-48, N-14, N-157, N-85, N-78, N-76, N-127	99.6	*Haloplasma contractile*	*Tenericutes*	*Haloplasmataceae*
		1				**72**	N-124	0	*Streptococcus sanguinis*	*Firmicutes*	*Streptococcaceae*
			1			**73**	Nsb-56	0	Uncultured *Actinobacteria*	*Actinobacteria*	Uncultured *Actinobacteria*
			1		**1**	**74**	Nsb-82	0	*Uncultured Actinobacteria*	*Actinobacteria*	Uncultured *Actinobacteria*
1						**75**	CN1-27	0	*Microbacterium oxydans*	*Actinobacteria*	*Microbacteriaceae*
		1				**76**	N-31	0	*Microbacterium invictum*	*Actinobacteria*	*Microbacteriaceae*
			15	**24**	**39**	**77**	Nbs-10, Nbs-92, Nbs-6, Nbs-65, Nbs-72, Nbs-94, Nbs-83, Nbs-41, Nbs-57, Nbs-38, Nbs-68, Nbs-66, Nbs-06, Nbs-43, Nbs-34, Nbs155, Nbs102, Nbs215, Nbs149, Nbs113, Nbs274, Nbs128, Nbs121, Nbs141, Nbs130, Nbs133, Nbs229, Nbs206, Nbs249, Nbs137, Nbs162, Nbs172, Nbs265, Nbs169, Nbs181, Nbs250, Nbs138	97.1	*Proteiniphilum acetatigenes*	*Bacteroidetes*	*Porphyromonadaceae*
			1		**1**	**78**	Nbs-42	0	*Parabacteroides goldsteinii*	*Bacteroidetes*	*Porphyromonadaceae*
			2	**2**	**4**	**79**	Nsb-37, Nsb-81, Nbs178, Nbs144		Uncultured *Bacteroidetes*	*Bacteroidetes*	Uncultured *Bacteroidetes*
	2	**1**				**80**	CN2-94, CN2-45, N-45	99.8	*Mucilaginibacter oryzae*	*Bacteroidetes*	*Sphingobacteriaceae*
		1				**81**	N-106	0	*Nubsella zeaxanthinifaciens*	*Bacteroidetes*	*Sphingobacteriaceae*
		2				**82**	N-138, N-192	99.1	*Olivibacter soli*	*Bacteroidetes*	*Sphingobacteriaceae*
2						**83**	CN1-36, CN1-37	99.7	*Terrimonas ferruginea*	*Bacteroidetes*	*Chitinophagaceae*
		1				**84**	N-155	0		*Cyanobacteria*	*Chloroplast*
			4	**2**	**6**	**85**	Nbs-8, Nbs-9, Nbs-52, Nbs-1, Nbs234, Nbs188	95	Uncultured *Chloroflexi*	*Chloroflexi*	Uncultured *Chloroflexi*
			1		**1**	**86**	Nbs-39	0	Uncultured *Spirochaeta*	*Spirochaetae*	Uncultured *Spirochaetae*
		1				**87**	N-125	0	*Luteolibacter pohnpeiensis*	*Verrucomicrobia*	*Verrucomicrobiaceae*
			1		**1**	**88**	Nbs-47	0	*Candidatus Xiphinematobacter americani*	*Verrucomicrobia*	*Verrucomicrobiaceae*
		1				**89**	N-134	0	*Opitutus terrae*	*Verrucomicrobia*	*Opitutaceae*
2						**90**	CN1-13, CN1-57	99.9	*Singulisphaera acidiphila*	*Planctomycetes*	*Planctomycetaceae*
		1				**91**	N-135	0	*Singulisphaera acidiphila*	*Planctomycetes*	*Planctomycetaceae*

3						**92**	CN1-46, CN1-105, CNI-99	99.5	*Planctomycetes maris*	*Planctomycetes*	*Planctomycetaceae*
			1			**93**	Nbs-80	0	Uncultured *candidate division* OP8	Candidate division OP8	Candidate division OP8
		1				**94**	N-202	0	*Verrucomicrobium spinosum*	*Verrucomicrobia*	*Verrucomicrobiaceae*
		1				**95**	N-06	0	*Proteiniphilum acetatigenes*	*Bacteroidetes*	*Porphyromonadaceae*
				1	**1**	**96**	Nbs132	0	*Pseudomonas* sp.	*Gammaproteobacteria*	*Pseudomonadaceae*
				1	**1**	**97**	Nbs117	0	*Pseudomonas* sp.	*Gammaproteobacteria*	*Pseudomonadaceae*
				2	**2**	**98**	Nbs266, Nbs175	99.7	*Alkanivorax* sp.	*Gammaproteobacteria*	*Alcanivoraceae*
				2	**2**	**99**	Nbs182, Nbs197	98	*Diaphorobacter* sp.	*Betaproteobacteria*	*Comamonadaceae*
				4	**4**	**100**	Nbs267, Nbs164, Nbs244, Nbs146	97.9	*Brevundimonas* sp.	*Betaproteobacteria*	*Caulobacteraceae*
				1	**1**	**101**	Nbs119	0	*Sphingomonas* sp.	*Alphaproteobacteria*	*Sphingomonadaceae*
				1	**1**	**102**	Nbs192	0	Uncultured *alpha*	*Alphaproteobacteria*	
				1	**1**	**103**	Nbs200	0	Uncultured *epsilon*	*Epsilonproteobacteria*	
				1	**1**	**104**	Nbs187	0	*Acholeplasma sp.*	*Tenericutes*	*Acholeplasmataceae*
				1	**1**	**105**	Nbs246	0	*Actinotalea sp.*	*Actinobacteria*	*Cellulomonadaceae*
				1	**1**	**106**	Nbs127	0	*Georgenia sp.*	*Actinobacteria*	*Bogoriellaceae*
				1	**1**	**107**	Nbs258	0	*Arcanobacterium sp.*	*Actinobacteria*	*Actinomycetaceae*
				1	**1**	**108**	Nbs211	0	Uncultured *Actinobacteria*	*Actinobacteria*	
				4	**4**	**109**	Nbs148, Nbs259, Nbs220, Nbs248	95	*Desulosforosinus sp.*	*Firmicutes*	*Peptococcaceae*
				4	**4**	**110**	Nbs143, Nbs190, Nbs276, Nbs257	95.4	*Sedimentibacter saalensis*	*Firmicutes*	*Clostridiales*
				2	**2**	**111**	Nbs237, Nbs153	95.2	*Anaerovorax* sp.	*Firmicutes*	*Clostridiales*
				1	**1**	**112**	Nbs255	0	*Fusibacter sp.*	*Firmicutes*	*Clostridiales*
				2	**2**	**113**	Nbs124, Nbs189	100	Uncultured *clostridiales*	*Firmicutes*	*Clostridiales*
				3	**3**	**114**	Nbs129, Nbs273, Nbs213	93	Uncultured *clostridiales*	*Firmicutes*	*Clostridiales*
				1	**1**	**115**	Nbs227	0	Uncultured *clostridiales*	*Firmicutes*	*Clostridiales*
				2	**2**	**116**	Nbs103, Nbs268	98.8	Uncultured *Bacteroidetes*	*Bacteroidetes*	
				1	**1**	**117**	Nbs180	0	Uncultured *Bacteroidetes*	*Bacteroidetes*	
				1	**1**	**118**	Nbs167	0	*Dysgonomonas sp.*	*Bacteroidetes*	*Porphyromonadaceae*
				1	**1**	**119**	Nbs196	0	Uncultured Candidate division TM7	Candidate division TM7	
				1	**1**	**120**	Nbs161	0	Uncultured	*Firmicutes*	

									bacterium		
				1	**1**	**121**	Nbs235	0	Uncultured *syntrophomonas*	*Firmicutes*	
				1	**1**	**122**	Nbs142	0	Uncultured *Chloroflexi*	*Chloroflexi*	
				1	**1**	**123**	Nbs111	0	Uncultured *bacterium*	Uncultured *bacteria*	
				1	**1**	**124**	Nbs184	0	Uncultured *bacterium*	Candidate division WCHB1	

[1] The Table provides the analysis of 670 16S rDNA sequences (N: 212; Nbs: 261 (86 clones + 175 454-partial sequences); CN1: 90; CN2: 107). Clones were grouped based on their affiliation to different operational phylogenic units (OPUs) as described in the **Supplementary Materials and Methods**. We considered an OPU to be represented by each single group of clones that formed an independent clade in the tree without considering a similarity cut-off value.

[2] Phylogenetic inference of the 16S RNA gene sequences was performed as described in **Supplementary Materials and Methods**; closest related specie (and sequence identity) associated to the sequence of interest is specifically shown.

Table S3 General features of the metagenomes of the four soil-derived communities that were examined*

	Sample			
	N	Nbs	CN1	CN2
Contigs	4335	16032	20809	9915
Total bps	2599359	17925186	20031687	13002964
Average contig size (bps)	599	1118	962	1311
Average GC content (%)	61.03	53.36	57.19	63.68
ncRNAs	14	50	20	55
rRNAs	41	85	79	20
tRNAs	39	208	170	141
Total RNAs	94	343	269	216
ORFs (at least 50 amino acids)	5211	21130	27124	18391
Average ORF size (bps)	405	631	566	632
ORFs with predicted function	4776	19867	25037	16527
Hypothetical	435	1891	2087	1864
Assigned to COGs	4630	19243	24322	15779
Different COGs	2089	2746	3961	3454
Average COG size	2.8	7.5	7.8	5.8
KEGG	1538	2751	2775	2507

*We isolated the microbial DNA, directly pyro-sequenced using a Roche GS FLX DNA sequencer, and produced sequences from which potential protein-coding genes (≥ 20 amino acids) and functional assignments was made on the basis of BLASTP analysis against reference dataset for “Clusters of Orthologous Groups (COG)” and “Kyoto Encyclopedia of Genes and Genomes (KEGG)” assignments. Full details for metagenomic setup (sequencing, assembly and gene prediction) are available in the **Supplementary Materials and Methods.**

Table S4 Enrichment in COG functional classes by meaning of Z-scores (A) and percentage of genes belonging to enriched COGs (B)

(A) Pair-pair COG comparison by meaning of z-scores

Enriched in <first set> compared with <second set>	Z-score of significance[1]	Class[2]	Confidence[1]	Instances in set 1[3]	Instances in set 2[3]
Nbs-CN1	6.13	COG3436L:Transposase and inactivated derivatives	0.99	54	19
Nbs-CN1	4.831	COG3039L:Transposase and inactivated derivatives, IS5 family	0.99	27	7
Nbs-CN1	4.77	COG1373R:Predicted ATPase (AAA+ superfamily)	0.99	28	8
Nbs-CN1	4.169	COG0553KL:Superfamily II DNA/RNA helicases, SNF2 family	0.99	36	18
Nbs-CN1	4.02	COG3250G:Beta-galactosidase/beta-glucuronidase	0.99	28	12
Nbs-CN1	3.972	COG0610V:Type I site-specific restriction-modification system, R (restriction) subunit and related helicases	0.99	24	9
Nbs-CN1	3.493	COG1002V:Type II restriction enzyme, methylase subunits	0.95	16	5
Nbs-CN1	3.477	COG2963L:Transposase and inactivated derivatives	0.99	32	19
Nbs-CN1	3.424	COG0286V:Type I restriction-modification system methyltransferase subunit	0.99	23	11
Nbs-CN1	3.372	COG0334E:Glutamate dehydrogenase/leucine dehydrogenase	0.95	14	4
Nbs-CN1	3.341	COG3464L:Transposase and inactivated derivatives	0.99	19	8
Nbs-CN1	3.3	COG2826L:Transposase and inactivated derivatives, IS30 family	0.95	15	5
Nbs-CN1	3.197	COG3507G:Beta-xylosidase	0.99	17	7
Nbs-CN1	3.13	COG0732V:Restriction endonuclease S subunits	0.99	19	9
Nbs-CN1	3.13	COG3547L:Transposase and inactivated derivatives	0.99	19	9
Nbs-CN1	3.046	COG0668M:Small-conductance mechanosensitive channel	0.95	15	6
Nbs-CN1	3.005	COG3177S:Uncharacterized conserved protein	0.99	16	7
Nbs-CN1	2.947	COG0643NT:Chemotaxis protein histidine kinase and related kinases	0.99	18	9
Nbs-CN1	2.888	COG0789K:Predicted transcriptional regulators	0.99	27	18
Nbs-CN1	2.858	COG0582L:Integrase	0.99	46	39
Nbs-CN1	2.732	COG0793M:Periplasmic protease	0.99	26	18
Nbs-CN1	2.72	COG4584L:Transposase and inactivated derivatives	0.99	23	15
Nbs-CN1	2.669	COG1479S:Uncharacterized conserved protein	0.95	12	5

Nbs-CN1	2.669	COG4206H:Outer membrane cobalamin receptor protein	0.95	12	5
Nbs-CN1	2.628	COG0348C:Polyferredoxin	0.95	13	6
Nbs-CN1	2.628	COG0084L:Mg-dependent DNase	0.95	13	6
Nbs-CN1	2.598	COG1092R:Predicted SAM-dependent methyltransferases	0.99	14	7
Nbs-CN1	2.564	COG0776L:Bacterial nucleoid DNA-binding protein	0.95	24	17
Nbs-CN1	2.544	COG0515RTKL:Serine/threonine protein kinase	0.95	20	13
Nbs-CN1	2.494	COG2801L:Transposase and inactivated derivatives	0.95	57	56
Nbs-CN1	2.487	COG0569P:K+ transport systems, NAD-binding component	0.95	10	4
Nbs-CN1	2.383	COG4147R:Predicted symporter	0.95	13	7
Nbs-CN1	2.377	COG1484L:DNA replication protein	0.95	21	15
Nbs-CN1	2.362	COG0543HC:2-polyprenylphenol hydroxylase and related flavodoxin oxidoreductases	0.95	19	13
Nbs-CN1	2.359	COG0826O:Collagenase and related proteases	0.95	15	9
Nbs-CN1	2.355	COG0057G:Glyceraldehyde-3-phosphate dehydrogenase/erythrose-4-phosphate dehydrogenase	0.95	16	10
Nbs-CN1	2.234	COG1826U:Sec-independent protein secretion pathway components	0.95	9	4
Nbs-CN1	2.2	COG0366G:Glycosidases	0.95	20	15
Nbs-CN1	2.196	COG0798P:Arsenite efflux pump ACR3 and related permeases	0.95	10	5
Nbs-CN1	2.196	COG1974KT:SOS-response transcriptional repressors (RecA-mediated autopeptidases)	0.95	10	5
Nbs-CN1	2.196	COG0082E:Chorismate synthase	0.95	10	5
Nbs-CN1	2.196	COG1555L:DNA uptake protein and related DNA-binding proteins	0.95	10	5
Nbs-CN1	2.158	COG2890J:Methylase of polypeptide chain release factors	0.95	12	7
Nbs-CN1	2.155	COG0784T:FOG: CheY-like receiver	0.95	37	35
Nbs-CN1	2.152	COG0635H:Coproporphyrinogen III oxidase and related Fe-S oxidoreductases	0.95	15	10
Nbs-CN1	2.149	COG3335L:Transposase and inactivated derivatives	0.95	14	9
Nbs-CN1	2.149	COG0534V:Na+-driven multidrug efflux pump	0.95	14	9
Nbs-CN1	2.149	COG0835NT:Chemotaxis signal transduction protein	0.95	14	9
Nbs-CN1	2.074	COG0494LR:NTP pyrophosphohydrolases including oxidative damage repair enzymes	0.95	22	18
Nbs-CN1	1.964	COG4942D:Membrane-bound metallopeptidase	0.95	8	4
Nbs-CN1	1.964	COG0503F:Adenine/guanine phosphoribosyltransferases and related PRPP-binding proteins	0.95	8	4
Nbs-CN1	1.964	COG1066O:Predicted ATP-dependent serine protease	0.95	8	4
Nbs-CN1	1.964	COG0763M:Lipid A disaccharide synthetase	0.95	8	4
Nbs-CN1	1.964	COG0005F:Purine nucleoside phosphorylase	0.95	8	4
Nbs-CN1	1.964	COG0392S:Predicted integral membrane protein	0.95	8	4
Nbs-CN1	1.964	COG1825J:Ribosomal protein L25 (general stress protein Ctc)	0.95	8	4

Nbs-CN2	5.788	COG3436L:Transposase and inactivated derivatives	0.99	54	10
Nbs-CN2	5.071	COG0673R:Predicted dehydrogenases and related proteins	0.95	36	5
Nbs-CN2	3.083	COG0793M:Periplasmic protease	0.99	26	9
Nbs-CN2	3.031	COG3464L:Transposase and inactivated derivatives	0.95	19	5
Nbs-CN2	2.999	COG3507G:Beta-xylosidase	0.95	17	4
Nbs-CN2	2.833	COG0209F:Ribonucleotide reductase, alpha subunit	0.99	21	7
Nbs-CN2	2.814	COG0610V:Type I site-specific restriction-modification system, R (restriction) subunit and related helicases	0.99	24	9
Nbs-CN2	2.678	COG0826O:Collagenase and related proteases	0.95	15	4
Nbs-CN2	2.562	COG3177S:Uncharacterized conserved protein	0.95	16	5
Nbs-CN2	2.466	COG0286V:Type I restriction-modification system methyltransferase subunit	0.95	23	10
Nbs-CN2	2.328	COG4775M:Outer membrane protein/protective antigen OMA87	0.95	13	4
Nbs-CN2	2.328	COG1752R:Predicted esterase of the alpha-beta hydrolase superfamily	0.95	13	4
Nbs-CN2	2.3	COG3547L:Transposase and inactivated derivatives	0.95	19	8
Nbs-CN2	2.3	COG1136V:ABC-type antimicrobial peptide transport system, ATPase component	0.95	19	8
Nbs-CN2	2.232	COG0366G:Glycosidases	0.95	20	9
Nbs-CN2	2.172	COG1484L:DNA replication protein	0.95	21	10
Nbs-CN2	2.141	COG0249L:Mismatch repair ATPase (MutS family)	0.95	12	4
Nbs-CN2	2.141	COG4206H:Outer membrane cobalamin receptor protein	0.95	12	4
Nbs-CN2	2.141	COG0060J:Isoleucyl-tRNA synthetase	0.95	12	4
Nbs-CN2	2.141	COG0156H:7-keto-8-aminopelargonate synthetase and related enzymes	0.95	12	4
Nbs-CN2	2.12	COG1131V:ABC-type multidrug transport system, ATPase component	0.95	26	14
Nbs-CN2	2.076	COG0050J:GTPases - translation elongation factors	0.95	19	9
Nbs-CN2	2.022	COG0776L:Bacterial nucleoid DNA-binding protein	0.95	24	13
Nbs-N	2.117	COG0673R:Predicted dehydrogenases and related proteins	0.95	36	4
Nbs-N	2.104	COG3436L:Transposase and inactivated derivatives	0.95	54	8
CN1-Nbs	5.308	COG3181S:Uncharacterized protein conserved in bacteria	0.99	143	35
CN1-Nbs	5.277	COG0583K:Transcriptional regulator	0.99	285	100
CN1-Nbs	4.859	COG1629P:Outer membrane receptor proteins, mostly Fe transport	0.99	215	72
CN1-Nbs	4.554	COG0747E:ABC-type dipeptide transport system, periplasmic component	0.95	50	5
CN1-Nbs	4.223	COG0683E:ABC-type branched-chain amino acid transport systems, periplasmic component	0.99	109	30
CN1-Nbs	4.146	COG0765E:ABC-type amino acid transport system, permease component	0.95	41	4
CN1-Nbs	4.105	COG2197TK:Response regulator containing a CheY-like receiver domain and an HTH DNA-binding domain	0.99	76	17

CN1-Nbs	3.643	COG1012C:NAD-dependent aldehyde dehydrogenases	0.99	120	40
CN1-Nbs	3.339	COG1167KE:Transcriptional regulators containing a DNA-binding HTH domain and an aminotransferase domain (MocR family) and their eukaryotic orthologs	0.99	42	8
CN1-Nbs	3.331	COG0559E:Branched-chain amino acid ABC-type transport system, permease components	0.99	64	17
CN1-Nbs	3.287	COG1609K:Transcriptional regulators	0.99	49	11
CN1-Nbs	3.265	COG0596R:Predicted hydrolases or acyltransferases (alpha/beta hydrolase superfamily)	0.99	94	31
CN1-Nbs	3.259	COG3279KT:Response regulator of the LytR/AlgR family	0.99	41	8
CN1-Nbs	3.229	COG1802K:Transcriptional regulators	0.99	38	7
CN1-Nbs	3.1	COG1804C:Predicted acyl-CoA transferases/carnitine dehydratase	0.99	49	12
CN1-Nbs	3.027	COG1020Q:Non-ribosomal peptide synthetase modules and related proteins	0.95	33	6
CN1-Nbs	2.979	COG0697GER:Permeases of the drug/metabolite transporter (DMT) superfamily	0.99	59	17
CN1-Nbs	2.937	COG0531E:Amino acid transporters	0.95	32	6
CN1-Nbs	2.908	COG0601EP:ABC-type dipeptide/oligopeptide/nickel transport systems, permease components	0.95	29	5
CN1-Nbs	2.908	COG4948MR:L-alanine-DL-glutamate epimerase and related enzymes of enolase superfamily	0.95	29	5
CN1-Nbs	2.791	COG0154J:Asp-tRNAAsn/Glu-tRNAGln amidotransferase A subunit and related amidases	0.99	33	7
CN1-Nbs	2.786	COG0308E:Aminopeptidase N	0.95	25	4
CN1-Nbs	2.751	COG1335Q:Amidases related to nicotinamidase	0.95	30	6
CN1-Nbs	2.735	COG4177E:ABC-type branched-chain amino acid transport system, permease component	0.99	49	14
CN1-Nbs	2.681	COG0841V:Cation/multidrug efflux pump	0.99	122	50
CN1-Nbs	2.661	COG2207K:AraC-type DNA-binding domain-containing proteins	0.99	80	29
CN1-Nbs	2.655	COG0179Q:2-keto-4-pentenoate hydratase/2-oxohepta-3-ene-1,7-dioic acid hydratase (catechol pathway)	0.99	34	8
CN1-Nbs	2.598	COG0318IQ:Acyl-CoA synthetases (AMP-forming)/AMP-acid ligases II	0.99	77	28
CN1-Nbs	2.579	COG1304C:L-lactate dehydrogenase (FMN-dependent) and related alpha-hydroxy acid dehydrogenases	0.95	23	4
CN1-Nbs	2.558	COG0624E:Acetylornithine deacetylase/Succinyl-diaminopimelate desuccinylase and related deacylases	0.95	49	15
CN1-Nbs	2.541	COG0183I:Acetyl-CoA acetyltransferase	0.95	70	25
CN1-Nbs	2.477	COG0665E:Glycine/D-amino acid oxidases (deaminating)	0.95	48	15
CN1-Nbs	2.471	COG0843C:Heme/copper-type cytochrome/quinol oxidases, subunit 1	0.95	22	4
CN1-Nbs	2.467	COG1250I:3-hydroxyacyl-CoA dehydrogenase	0.95	32	8
CN1-Nbs	2.455	COG2373R:Large extracellular alpha-helical protein	0.95	27	6
CN1-Nbs	2.426	COG0410E:ABC-type branched-chain amino acid transport systems, ATPase component	0.95	43	13
CN1-Nbs	2.409	COG1593G:TRAP-type C4-dicarboxylate transport system, large permease component	0.95	29	7
CN1-Nbs	2.404	COG0715P:ABC-type nitrate/sulfonate/bicarbonate transport systems, periplasmic components	0.95	36	10
CN1-Nbs	2.404	COG0365I:Acyl-coenzyme A synthetases/AMP-(fatty) acid ligases	0.95	36	10
CN1-Nbs	2.36	COG1171E:Threonine dehydratase	0.95	21	4

CN1-Nbs	2.33	COG1024I:Enoyl-CoA hydratase/carnithine racemase	0.95	71	27
CN1-Nbs	2.312	COG1846K:Transcriptional regulators	0.95	46	15
CN1-Nbs	2.296	COG1670J:Acetyltransferases, including N-acetylases of ribosomal proteins	0.95	23	5
CN1-Nbs	2.269	COG1414K:Transcriptional regulator	0.95	39	12
CN1-Nbs	2.245	COG3523S:Uncharacterized protein conserved in bacteria	0.95	20	4
CN1-Nbs	2.245	COG2274V:ABC-type bacteriocin/lantibiotic exporters, contain an N-terminal double-glycine peptidase domain	0.95	20	4
CN1-Nbs	2.245	COG1653G:ABC-type sugar transport system, periplasmic component	0.95	20	4
CN1-Nbs	2.198	COG0840NT:Methyl-accepting chemotaxis protein	0.95	100	43
CN1-Nbs	2.184	COG0251J:Putative translation initiation inhibitor, yjgF family	0.95	22	5
CN1-Nbs	2.136	COG3203M:Outer membrane protein (porin)	0.95	24	6
CN1-Nbs	2.101	COG1522K:Transcriptional regulators	0.95	35	11
CN1-Nbs	2.069	COG0507L:ATP-dependent exoDNAse (exonuclease V), alpha subunit - helicase superfamily I member	0.95	21	5
CN1-Nbs	2.069	COG3842E:ABC-type spermidine/putrescine transport systems, ATPase components	0.95	21	5
CN1-Nbs	2.067	COG0329EM:Dihydrodipicolinate synthase/N-acetylneuraminate lyase	0.95	28	8
CN1-Nbs	2.039	COG0451MG:Nucleoside-diphosphate-sugar epimerases	0.95	51	19
CN1-Nbs	2.005	COG1475K:Predicted transcriptional regulators	0.95	34	11
CN1-Nbs	2.004	COG1807M:4-amino-4-deoxy-L-arabinose transferase and related glycosyltransferases of PMT family	0.95	18	4
CN1-Nbs	1.989	COG1529C:Aerobic-type carbon monoxide dehydrogenase, large subunit CoxL/CutL homologs	0.95	25	7
CN1-CN2	4.345	COG3279KT:Response regulator of the LytR/AlgR family	0.95	41	4
CN1-CN2	4.046	COG2931Q:RTX toxins and related Ca2+-binding proteins	0.99	75	19
CN1-CN2	4.041	COG3209M:Rhs family protein	0.95	40	5
CN1-CN2	4.041	COG0673R:Predicted dehydrogenases and related proteins	0.95	40	5
CN1-CN2	3.136	COG0183I:Acetyl-CoA acetyltransferase	0.99	70	23
CN1-CN2	3.085	COG0612R:Predicted Zn-dependent peptidases	0.95	29	5
CN1-CN2	3.061	COG0318IQ:Acyl-CoA synthetases (AMP-forming)/AMP-acid ligases II	0.99	77	27
CN1-CN2	2.811	COG1609K:Transcriptional regulators	0.99	49	15
CN1-CN2	2.599	COG0438M:Glycosyltransferase	0.99	76	30
CN1-CN2	2.585	COG0683E:ABC-type branched-chain amino acid transport systems, periplasmic component	0.99	109	48
CN1-CN2	2.542	COG0443O:Molecular chaperone	0.95	33	9
CN1-CN2	2.527	COG1197LK:Transcription-repair coupling factor (superfamily II helicase)	0.95	26	6
CN1-CN2	2.471	COG0451MG:Nucleoside-diphosphate-sugar epimerases	0.95	51	18
CN1-CN2	2.418	COG0308E:Aminopeptidase N	0.95	25	6
CN1-CN2	2.396	COG3523S:Uncharacterized protein conserved in bacteria	0.95	20	4

CN1-CN2	2.275	COG0399M:Predicted pyridoxal phosphate-dependent enzyme apparently involved in regulation of cell wall biogenesis	0.95	19	4
CN1-CN2	2.149	COG0463M:Glycosyltransferases involved in cell wall biogenesis	0.95	51	20
CN1-CN2	2.148	COG2373R:Large extracellular alpha-helical protein	0.95	27	8
CN1-CN2	2.141	COG0559E:Branched-chain amino acid ABC-type transport system, permease components	0.95	64	27
CN1-CN2	2.123	COG0410E:ABC-type branched-chain amino acid transport systems, ATPase component	0.95	43	16
CN1-CN2	2.119	COG0265O:Trypsin-like serine proteases, typically periplasmic, contain C-terminal PDZ domain	0.95	33	11
CN1-CN2	2.119	COG0664T:cAMP-binding proteins - catabolite gene activator and regulatory subunit of cAMP-dependent protein kinases	0.95	41	15
CN1-CN2	2.075	COG0702MG:Predicted nucleoside-diphosphate-sugar epimerases	0.95	22	6
CN1-CN2	2.047	COG1846K:Transcriptional regulators	0.95	46	18
CN1-CN2	2.027	COG4585T:Signal transduction histidine kinase	0.95	28	9
CN1-CN2	1.966	COG2197TK:Response regulator containing a CheY-like receiver domain and an HTH DNA-binding domain	0.95	76	35
CN1-N	2.779	COG0845M:Membrane-fusion protein	0.99	96	7
CN1-N	2.353	COG2197TK:Response regulator containing a CheY-like receiver domain and an HTH DNA-binding domain	0.95	76	6
CN1-N	2.096	COG0438M:Glycosyltransferase	0.95	76	7
CN2-Nbs	9.306	COG0583K:Transcriptional regulator	0.99	298	100
CN2-Nbs	6.302	COG3181S:Uncharacterized protein conserved in bacteria	0.99	118	35
CN2-Nbs	5.387	COG3451U:Type IV secretory pathway, VirB4 components	0.95	47	6
CN2-Nbs	4.794	COG1629P:Outer membrane receptor proteins, mostly Fe transport	0.99	151	72
CN2-Nbs	4.208	COG1012C:NAD-dependent aldehyde dehydrogenases	0.99	93	40
CN2-Nbs	4.13	COG0747E:ABC-type dipeptide transport system, periplasmic component	0.95	31	5
CN2-Nbs	3.939	COG2207K:AraC-type DNA-binding domain-containing proteins	0.99	72	29
CN2-Nbs	3.359	COG1167KE:Transcriptional regulators containing a DNA-binding HTH domain and an aminotransferase domain (MocR family) and their eukaryotic orthologs	0.99	30	8
CN2-Nbs	3.121	COG1414K:Transcriptional regulator	0.99	35	12
CN2-Nbs	3.041	COG1028IQR:Dehydrogenases with different specificities (related to short-chain alcohol dehydrogenases)	0.99	151	95
CN2-Nbs	2.989	COG0154J:Asp-tRNAAsn/Glu-tRNAGln amidotransferase A subunit and related amidases	0.99	25	7
CN2-Nbs	2.964	COG0765E:ABC-type amino acid transport system, permease component	0.95	19	4
CN2-Nbs	2.964	COG0601EP:ABC-type dipeptide/oligopeptide/nickel transport systems, permease components	0.95	21	5
CN2-Nbs	2.935	COG0277C:FAD/FMN-containing dehydrogenases	0.99	43	18
CN2-Nbs	2.911	COG0665E:Glycine/D-amino acid oxidases (deaminating)	0.99	38	15
CN2-Nbs	2.85	COG0477GEPR:Permeases of the major facilitator superfamily	0.99	119	73
CN2-Nbs	2.825	COG0843C:Heme/copper-type cytochrome/quinol oxidases, subunit 1	0.95	18	4

CN2-Nbs	2.734	COG1404O:Subtilisin-like serine proteases	0.99	23	7
CN2-Nbs	2.69	COG0841V:Cation/multidrug efflux pump	0.99	86	50
CN2-Nbs	2.628	COG0596R:Predicted hydrolases or acyltransferases (alpha/beta hydrolase superfamily)	0.99	59	31
CN2-Nbs	2.576	COG0840NT:Methyl-accepting chemotaxis protein	0.95	75	43
CN2-Nbs	2.571	COG0582L:Integrase	0.95	79	46
CN2-Nbs	2.509	COG1475K:Predicted transcriptional regulators	0.95	28	11
CN2-Nbs	2.436	COG0028EH:Thiamine pyrophosphate-requiring enzymes [acetolactate synthase, pyruvate dehydrogenase (cytochrome), glyoxylate carboligase, phosphonopyruvate decarboxylase]	0.95	29	12
CN2-Nbs	2.429	COG3481R:Predicted HD-superfamily hydrolase	0.95	19	6
CN2-Nbs	2.321	COG1024I:Enoyl-CoA hydratase/carnithine racemase	0.95	50	27
CN2-Nbs	2.304	COG1192D:ATPases involved in chromosome partitioning	0.95	31	14
CN2-Nbs	2.298	COG0642T:Signal transduction histidine kinase	0.95	138	95
CN2-Nbs	2.282	COG1020Q:Non-ribosomal peptide synthetase modules and related proteins	0.95	18	6
CN2-Nbs	2.249	COG2197TK:Response regulator containing a CheY-like receiver domain and an HTH DNA-binding domain	0.95	35	17
CN2-Nbs	2.243	COG0251J:Putative translation initiation inhibitor, yjgF family	0.95	16	5
CN2-Nbs	2.236	COG0745TK:Response regulators consisting of a CheY-like receiver domain and a winged-helix DNA-binding domain	0.95	76	47
CN2-Nbs	2.136	COG0697GER:Permeases of the drug/metabolite transporter (DMT) superfamily	0.95	34	17
CN2-Nbs	2.129	COG1049C:Aconitase B	0.95	17	6
CN2-Nbs	2.086	COG0179Q:2-keto-4-pentenoate hydratase/2-oxohepta-3-ene-1,7-dioic acid hydratase (catechol pathway)	0.95	20	8
CN2-Nbs	2.041	COG4993G:Glucose dehydrogenase	0.95	13	4
CN2-Nbs	2.041	COG1305E:Transglutaminase-like enzymes, putative cysteine proteases	0.95	13	4
CN2-Nbs	2.028	COG1802K:Transcriptional regulators	0.95	18	7
CN2-Nbs	2.028	COG0693R:Putative intracellular protease/amidase	0.95	18	7
CN2-Nbs	2.028	COG2303E:Choline dehydrogenase and related flavoproteins	0.95	18	7
CN2-Nbs	1.966	COG0625O:Glutathione S-transferase	0.95	45	26
CN2-CN1	5.818	COG0582L:Integrase	0.99	79	39
CN2-CN1	5.172	COG0583K:Transcriptional regulator	0.99	298	285
CN2-CN1	4.915	COG3451U:Type IV secretory pathway, VirB4 components	0.99	47	20
CN2-CN1	4.159	COG0784T:FOG: CheY-like receiver	0.99	57	35
CN2-CN1	4.143	COG0553KL:Superfamily II DNA/RNA helicases, SNF2 family	0.99	38	18
CN2-CN1	3.732	COG0642T:Signal transduction histidine kinase	0.99	138	128
CN2-CN1	3.364	COG3039L:Transposase and inactivated derivatives, IS5 family	0.99	19	7
CN2-CN1	3.295	COG3210U:Large exoproteins involved in heme utilization or adhesion	0.99	41	27

CN2-CN1	3.079	COG0789K:Predicted transcriptional regulators	0.99	30	18
CN2-CN1	3.061	COG0668M:Small-conductance mechanosensitive channel	0.95	16	6
CN2-CN1	3.008	COG1049C:Aconitase B	0.99	17	7
CN2-CN1	2.93	COG2199T:FOG: GGDEF domain	0.99	41	30
CN2-CN1	2.928	COG0550L:Topoisomerase IA	0.99	19	9
CN2-CN1	2.869	COG3706T:Response regulator containing a CheY-like receiver domain and a GGDEF domain	0.95	15	6
CN2-CN1	2.725	COG1974KT:SOS-response transcriptional repressors (RecA-mediated autopeptidases)	0.95	13	5
CN2-CN1	2.626	COG2890J:Methylase of polypeptide chain release factors	0.99	15	7
CN2-CN1	2.626	COG1289S:Predicted membrane protein	0.99	15	7
CN2-CN1	2.575	COG3311K:Predicted transcriptional regulator	0.95	11	4
CN2-CN1	2.575	COG1113E:Gamma-aminobutyrate permease and related permeases	0.95	11	4
CN2-CN1	2.563	COG0526OC:Thiol-disulfide isomerase and thioredoxins	0.95	39	31
CN2-CN1	2.543	COG0629L:Single-stranded DNA-binding protein	0.95	18	10
CN2-CN1	2.511	COG3971Q:2-keto-4-pentenoate hydratase	0.95	12	5
CN2-CN1	2.511	COG2826L:Transposase and inactivated derivatives, IS30 family	0.95	12	5
CN2-CN1	2.511	COG5001T:Predicted signal transduction protein containing a membrane domain, an EAL and a GGDEF domain	0.95	12	5
CN2-CN1	2.396	COG1062C:Zn-dependent alcohol dehydrogenases, class III	0.95	15	8
CN2-CN1	2.345	COG1651O:Protein-disulfide isomerase	0.95	18	11
CN2-CN1	2.342	COG3419NU:Tfp pilus assembly protein, tip-associated adhesin PilY1	0.95	10	4
CN2-CN1	2.342	COG1450NU:Type II secretory pathway, component PulD	0.95	10	4
CN2-CN1	2.336	COG2963L:Transposase and inactivated derivatives	0.95	26	19
CN2-CN1	2.33	COG1028IQR:Dehydrogenases with different specificities (related to short-chain alcohol dehydrogenases)	0.95	151	170
CN2-CN1	2.323	COG2801L:Transposase and inactivated derivatives	0.95	59	56
CN2-CN1	2.214	COG1734T:DnaK suppressor protein	0.95	13	7
CN2-CN1	2.214	COG4773P:Outer membrane receptor for ferric coprogen and ferric-rhodotorulic acid	0.95	13	7
CN2-CN1	2.18	COG2010C:Cytochrome c, mono- and diheme variants	0.95	25	19
CN2-CN1	2.166	COG1404O:Subtilisin-like serine proteases	0.95	23	17
CN2-CN1	2.153	COG0515RTKL:Serine/threonine protein kinase	0.95	19	13
CN2-CN1	2.096	COG0460E:Homoserine dehydrogenase	0.95	9	4
CN2-CN1	2.096	COG0334E:Glutamate dehydrogenase/leucine dehydrogenase	0.95	9	4
CN2-CN1	2.096	COG0471P:Di- and tricarboxylate transporters	0.95	9	4
CN2-CN1	2.049	COG2165NU:Type II secretory pathway, pseudopilin PulG	0.95	10	5
CN2-CN1	2.049	COG1002V:Type II restriction enzyme, methylase subunits	0.95	10	5
CN2-CN1	2.049	COG0439I:Biotin carboxylase	0.95	10	5

CN2-CN1	2.017	COG0348C:Polyferredoxin	0.95	11	6
CN2-CN1	1.995	COG1894C:NADH:ubiquinone oxidoreductase, NADH-binding (51 kD) subunit	0.95	12	7
CN2-CN1	1.995	COG0270L:Site-specific DNA methylase	0.95	12	7
CN2-CN1	1.976	COG3501S:Uncharacterized protein conserved in bacteria	0.95	19	14
CN2-CN1	1.971	COG2378K:Predicted transcriptional regulator	0.95	14	9
CN2-CN1	1.971	COG0835NT:Chemotaxis signal transduction protein	0.95	14	9
CN2-CN1	1.971	COG1236J:Predicted exonuclease of the beta-lactamase fold involved in RNA processing	0.95	14	9
CN2-CN1	1.971	COG0643NT:Chemotaxis protein histidine kinase and related kinases	0.95	14	9
CN2-CN1	1.971	COG2003L:DNA repair proteins	0.95	18	13
CN2-CN1	1.971	COG0693R:Putative intracellular protease/amidase	0.95	18	13
CN2-CN1	1.967	COG0810M:Periplasmic protein TonB, links inner and outer membranes	0.95	17	12
CN2-CN1	1.967	COG2200T:FOG: EAL domain	0.95	17	12
CN2-CN1	1.967	COG1230P:Co/Zn/Cd efflux system component	0.95	15	10
CN2-N	2.461	COG0845M:Membrane-fusion protein	0.95	60	7
CN2-N	2.087	COG0277C:FAD/FMN-containing dehydrogenases	0.95	43	5
N-Nbs	5.556	COG0840NT:Methyl-accepting chemotaxis protein	0.99	43	43
N-Nbs	5.497	COG0583K:Transcriptional regulator	0.99	73	100
N-Nbs	5.193	COG1012C:NAD-dependent aldehyde dehydrogenases	0.99	39	40
N-Nbs	4.185	COG1020Q:Non-ribosomal peptide synthetase modules and related proteins	0.95	12	6
N-Nbs	4.012	COG1167KE:Transcriptional regulators containing a DNA-binding HTH domain and an aminotransferase domain (MocR family) and their eukaryotic orthologs	0.99	13	8
N-Nbs	3.557	COG0531E:Amino acid transporters	0.95	10	6
N-Nbs	3.544	COG0596R:Predicted hydrolases or acyltransferases (alpha/beta hydrolase superfamily)	0.99	25	31
N-Nbs	3.223	COG4644L:Transposase and inactivated derivatives, TnpA family	0.99	13	12
N-Nbs	3.223	COG2200T:FOG: EAL domain	0.99	13	12
N-Nbs	3.218	COG0112E:Glycine/serine hydroxymethyltransferase	0.95	9	6
N-Nbs	3.145	COG0834ET:ABC-type amino acid transport/signal transduction systems, periplasmic component/domain	0.99	20	25
N-Nbs	3.125	COG4977K:Transcriptional regulator containing an amidase domain and an AraC-type DNA-binding HTH domain	0.95	8	5
N-Nbs	2.992	COG0642T:Signal transduction histidine kinase	0.99	51	95
N-Nbs	2.945	COG1538MU:Outer membrane protein	0.99	26	39
N-Nbs	2.93	COG0174E:Glutamine synthetase	0.99	12	12
N-Nbs	2.92	COG2202T:FOG: PAS/PAC domain	0.99	18	23

N-Nbs	2.745	COG0477GEPR:Permeases of the major facilitator superfamily	0.99	40	73
N-Nbs	2.627	COG3523S:Uncharacterized protein conserved in bacteria	0.95	6	4
N-Nbs	2.627	COG0843C:Heme/copper-type cytochrome/quinol oxidases, subunit 1	0.95	6	4
N-Nbs	2.479	COG3264M:Small-conductance mechanosensitive channel	0.95	7	6
N-Nbs	2.454	COG0085K:DNA-directed RNA polymerase, beta subunit/140 kD subunit	0.95	11	13
N-Nbs	2.425	COG4638PR:Phenylpropionate dioxygenase and related ring-hydroxylating dioxygenases, large terminal subunit	0.95	13	17
N-Nbs	2.392	COG1250I:3-hydroxyacyl-CoA dehydrogenase	0.95	8	8
N-Nbs	2.291	COG0443O:Molecular chaperone	0.95	11	14
N-Nbs	2.237	COG0753P:Catalase	0.95	7	7
N-Nbs	2.237	COG1802K:Transcriptional regulators	0.95	7	7
N-Nbs	2.237	COG2201NT:Chemotaxis response regulator containing a CheY-like receiver domain and a methylesterase domain	0.95	7	7
N-Nbs	2.182	COG0308E:Aminopeptidase N	0.95	5	4
N-Nbs	2.182	COG0439I:Biotin carboxylase	0.95	5	4
N-Nbs	2.182	COG0765E:ABC-type amino acid transport system, permease component	0.95	5	4
N-Nbs	2.182	COG1793L:ATP-dependent DNA ligase	0.95	5	4
N-Nbs	2.182	COG1737K:Transcriptional regulators	0.95	5	4
N-Nbs	2.182	COG1391OT:Glutamine synthetase adenylyltransferase	0.95	5	4
N-Nbs	2.182	COG2234R:Predicted aminopeptidases	0.95	5	4
N-Nbs	2.153	COG0402FR:Cytosine deaminase and related metal-dependent hydrolases	0.95	9	11
N-Nbs	2.071	COG1570L:Exonuclease VII, large subunit	0.95	6	6
N-Nbs	2.035	COG3696P:Putative silver efflux pump	0.95	14	22
N-Nbs	2.015	COG1352NT:Methylase of chemotaxis methyl-accepting proteins	0.95	7	8
N-Nbs	1.986	COG0057G:Glyceraldehyde-3-phosphate dehydrogenase/erythrose-4-phosphate dehydrogenase	0.95	11	16
N-CN1	4.643	COG4644L:Transposase and inactivated derivatives, TnpA family	0.99	13	12
N-CN1	4.643	COG2200T:FOG: EAL domain	0.99	13	12
N-CN1	4.568	COG2202T:FOG: PAS/PAC domain	0.99	18	23
N-CN1	4.301	COG0057G:Glyceraldehyde-3-phosphate dehydrogenase/erythrose-4-phosphate dehydrogenase	0.99	11	10
N-CN1	4.261	COG5001T:Predicted signal transduction protein containing a membrane domain, an EAL and a GGDEF domain	0.95	8	5
N-CN1	4.141	COG0840NT:Methyl-accepting chemotaxis protein	0.99	43	100
N-CN1	4.073	COG0642T:Signal transduction histidine kinase	0.99	51	128
N-CN1	3.942	COG0174E:Glutamine synthetase	0.99	12	14
N-CN1	3.579	COG3501S:Uncharacterized protein conserved in bacteria	0.99	11	14
N-CN1	3.524	COG2113E:ABC-type proline/glycine betaine transport systems, periplasmic components	0.95	7	6

N-CN1	3.48	COG0050J:GTPases - translation elongation factors	0.99	12	17
N-CN1	3.438	COG0582L:Integrase	0.99	20	39
N-CN1	3.366	COG0543HC:2-polyprenylphenol hydroxylase and related flavodoxin oxidoreductases	0.99	10	13
N-CN1	3.303	COG1570L:Exonuclease VII, large subunit	0.95	6	5
N-CN1	3.303	COG2826L:Transposase and inactivated derivatives, IS30 family	0.95	6	5
N-CN1	3.286	COG0835NT:Chemotaxis signal transduction protein	0.99	8	9
N-CN1	3.265	COG0834ET:ABC-type amino acid transport/signal transduction systems, periplasmic component/domain	0.99	20	41
N-CN1	3.203	COG0210L:Superfamily I DNA and RNA helicases	0.99	12	19
N-CN1	3.068	COG2766T:Putative Ser protein kinase	0.95	5	4
N-CN1	3.031	COG3706T:Response regulator containing a CheY-like receiver domain and a GGDEF domain	0.95	6	6
N-CN1	2.995	COG0784T:FOG: CheY-like receiver	0.99	17	35
N-CN1	2.901	COG2199T:FOG: GGDEF domain	0.99	15	30
N-CN1	2.836	COG0643NT:Chemotaxis protein histidine kinase and related kinases	0.99	7	9
N-CN1	2.812	COG0112E:Glycine/serine hydroxymethyltransferase	0.99	9	14
N-CN1	2.787	COG4773P:Outer membrane receptor for ferric coprogen and ferric-rhodotorulic acid	0.95	6	7
N-CN1	2.767	COG0082E:Chorismate synthase	0.95	5	5
N-CN1	2.767	COG0439I:Biotin carboxylase	0.95	5	5
N-CN1	2.767	COG3459G:Cellobiose phosphorylase	0.95	5	5
N-CN1	2.66	COG0402FR:Cytosine deaminase and related metal-dependent hydrolases	0.99	9	15
N-CN1	2.641	COG1352NT:Methylase of chemotaxis methyl-accepting proteins	0.99	7	10
N-CN1	2.622	COG0553KL:Superfamily II DNA/RNA helicases, SNF2 family	0.99	10	18
N-CN1	2.564	COG0469G:Pyruvate kinase	0.95	6	8
N-CN1	2.561	COG4977K:Transcriptional regulator containing an amidase domain and an AraC-type DNA-binding HTH domain	0.95	8	13
N-CN1	2.505	COG1012C:NAD-dependent aldehyde dehydrogenases	0.95	39	120
N-CN1	2.502	COG0668M:Small-conductance mechanosensitive channel	0.95	5	6
N-CN1	2.502	COG0196H:FAD synthase	0.95	5	6
N-CN1	2.492	COG1538MU:Outer membrane protein	0.95	26	72
N-CN1	2.475	COG1322S:Uncharacterized protein conserved in bacteria	0.95	4	4
N-CN1	2.475	COG3285L:Predicted eukaryotic-type DNA primase	0.95	4	4
N-CN1	2.475	COG0190H:5,10-methylene-tetrahydrofolate dehydrogenase/Methenyl tetrahydrofolate cyclohydrolase	0.95	4	4
N-CN1	2.475	COG1966T:Carbon starvation protein, predicted membrane protein	0.95	4	4
N-CN1	2.46	COG0753P:Catalase	0.95	7	11
N-CN1	2.36	COG2831U:Hemolysin activation/secretion protein	0.95	6	9
N-CN1	2.36	COG0520E:Selenocysteine lyase	0.95	6	9

N-CN1	2.36	COG3547L:Transposase and inactivated derivatives	0.95	6	9
N-CN1	2.29	COG0465O:ATP-dependent Zn proteases	0.95	7	12
N-CN1	2.29	COG0617J:tRNA nucleotidyltransferase/poly(A) polymerase	0.95	7	12
N-CN1	2.264	COG0047F:Phosphoribosylformylglycinamidine (FGAM) synthase, glutamine amidotransferase domain	0.95	5	7
N-CN1	2.264	COG3039L:Transposase and inactivated derivatives, IS5 family	0.95	5	7
N-CN1	2.258	COG4584L:Transposase and inactivated derivatives	0.95	8	15
N-CN1	2.182	COG2894D:Septum formation inhibitor-activating ATPase	0.95	4	5
N-CN1	2.182	COG2124Q:Cytochrome P450	0.95	4	5
N-CN1	2.182	COG0150F:Phosphoribosylaminoimidazole (AIR) synthetase	0.95	4	5
N-CN1	2.182	COG0151F:Phosphoribosylamine-glycine ligase	0.95	4	5
N-CN1	2.182	COG1047O:FKBP-type peptidyl-prolyl cis-trans isomerases 2	0.95	4	5
N-CN1	2.182	COG0456R:Acetyltransferases	0.95	4	5
N-CN1	2.182	COG0807H:GTP cyclohydrolase II	0.95	4	5
N-CN1	2.172	COG0069E:Glutamate synthase domain 2	0.95	6	10
N-CN1	2.172	COG0046F:Phosphoribosylformylglycinamidine (FGAM) synthase, synthetase domain	0.95	6	10
N-CN1	2.14	COG3696P:Putative silver efflux pump	0.95	14	35
N-CN1	2.12	COG0458EF:Carbamoylphosphate synthase large subunit (split gene in MJ)	0.95	9	19
N-CN1	2.11	COG2931Q:RTX toxins and related Ca2+-binding proteins	0.95	25	75
N-CN1	2.049	COG1391OT:Glutamine synthetase adenylyltransferase	0.95	5	8
N-CN1	2.049	COG0288P:Carbonic anhydrase	0.95	5	8
N-CN1	1.997	COG2823R:Predicted periplasmic or secreted lipoprotein	0.95	6	11
N-CN1	1.997	COG0751J:Glycyl-tRNA synthetase, beta subunit	0.95	6	11
N-CN2	5.315	COG2931Q:RTX toxins and related Ca2+-binding proteins	0.99	25	19
N-CN2	5.236	COG3209M:Rhs family protein	0.95	14	5
N-CN2	3.919	COG0174E:Glutamine synthetase	0.99	12	8
N-CN2	3.705	COG0050J:GTPases - translation elongation factors	0.99	12	9
N-CN2	3.452	COG0840NT:Methyl-accepting chemotaxis protein	0.99	43	75
N-CN2	3.393	COG0112E:Glycine/serine hydroxymethyltransferase	0.95	9	6
N-CN2	3.393	COG0458EF:Carbamoylphosphate synthase large subunit (split gene in MJ)	0.95	9	6
N-CN2	3.388	COG0443O:Molecular chaperone	0.99	11	9
N-CN2	3.192	COG3264M:Small-conductance mechanosensitive channel	0.95	7	4
N-CN2	2.997	COG0057G:Glyceraldehyde-3-phosphate dehydrogenase/erythrose-4-phosphate dehydrogenase	0.99	11	11
N-CN2	2.932	COG4644L:Transposase and inactivated derivatives, TnpA family	0.99	13	15

N-CN2	2.77	COG3523S:Uncharacterized protein conserved in bacteria	0.95	6	4
N-CN2	2.77	COG0469G:Pyruvate kinase	0.95	6	4
N-CN2	2.77	COG1570L:Exonuclease VII, large subunit	0.95	6	4
N-CN2	2.695	COG2202T:FOG: PAS/PAC domain	0.99	18	27
N-CN2	2.633	COG2200T:FOG: EAL domain	0.99	13	17
N-CN2	2.632	COG0465O:ATP-dependent Zn proteases	0.95	7	6
N-CN2	2.476	COG1082G:Sugar phosphate isomerases/epimerases	0.95	6	5
N-CN2	2.476	COG2823R:Predicted periplasmic or secreted lipoprotein	0.95	6	5
N-CN2	2.39	COG0617J:tRNA nucleotidyltransferase/poly(A) polymerase	0.95	7	7
N-CN2	2.327	COG0330O:Membrane protease subunits, stomatin/prohibitin homologs	0.95	11	15
N-CN2	2.312	COG4569Q:Acetaldehyde dehydrogenase (acetylating)	0.95	5	4
N-CN2	2.312	COG0196H:FAD synthase	0.95	5	4
N-CN2	2.312	COG2766T:Putative Ser protein kinase	0.95	5	4
N-CN2	2.227	COG1538MU:Outer membrane protein	0.95	26	51
N-CN2	2.213	COG0621J:2-methylthioadenine synthetase	0.95	6	6
N-CN2	2.199	COG1020Q:Non-ribosomal peptide synthetase modules and related proteins	0.95	12	18
N-CN2	2.161	COG0531E:Amino acid transporters	0.95	10	14
N-CN2	2.152	COG0834ET:ABC-type amino acid transport/signal transduction systems, periplasmic component/domain	0.95	20	37
N-CN2	2.151	COG3436L:Transposase and inactivated derivatives	0.95	8	10
N-CN2	2.065	COG2204T:Response regulator containing CheY-like receiver, AAA-type ATPase, and DNA-binding domains	0.95	12	19
N-CN2	2.02	COG0612R:Predicted Zn-dependent peptidases	0.95	5	5
N-CN2	2.02	COG0288P:Carbonic anhydrase	0.95	5	5
N-CN2	2.013	COG3696P:Putative silver efflux pump	0.95	14	24
N-CN2	2.007	COG0543HC:2-polyprenylphenol hydroxylase and related flavodoxin oxidoreductases	0.95	10	15
N-CN2	1.974	COG4638PR:Phenylpropionate dioxygenase and related ring-hydroxylating dioxygenases, large terminal subunit	0.95	13	22
N-CN2	1.968	COG3344L:Retron-type reverse transcriptase	0.95	8	11
N-CN2	1.968	COG4977K:Transcriptional regulator containing an amidase domain and an AraC-type DNA-binding HTH domain	0.95	8	11

[1] z-scores were calculated as proposed by Li (2009). See **Supplementary Materials and Methods** for details. Confidence levels of 0.95 and 0.99 are defined by *z* values of 1.96 and 2.58, and minimum H (the total number of observations in metagenomes) values of 4 and 7, respectively (Li, 2009). Over-representation of COG functions was made by pair comparisons as indicated.

[2] COG categories included: A (RNA processing and modification), B (Chromatin structure and dynamics), C (Energy production and conversion), D (Cell cycle control and mitosis), E (Amino acid metabolism and transport), F (Nucleotide metabolism and transport), G (Carbohydrate metabolism and transport), H (Coenzyme metabolism), I (Lipid metabolism), J (Translation), K (Transcription), L (Replication and repair), M (Cell

wall/membrane/envelop biogenesis), N (Cell otility), O (Post-translational modification, protein turnover, chaperone functions), Q (Secondary structure), T (Signal transduction), U (Intracellular trafficking and secretion), Y (Nuclear structure), Z (Cytoskeleton), R (General function prediction only) and S (Function unknown).

[3] The number of genes identified per COG is specifically indicated as *Instances in set 1* for those found in the metagenome 1 and *Instances in set 2* for those found in the metagenome 2. The percentage of those genes in relation to the total number of genes in a given metagenome is shown in panel (B).

(B) Percentage of genes per COG found to be enriched (see (A)).

Class[1]	Percentage (%) of the nr. genes in N[2]	Percentage (%) of the nr. genes in Nbs[2]	Percentage (%) of the nr. genes in CN1[2]	Percentage (%) of the nr. genes in CN2[2]
COG0583K:Transcriptional regulator	1.743	0.768	1.401	2.136
COG0642T:Signal transduction histidine kinase	1.217	0.73	0.629	0.989
COG1028IQR:Dehydrogenases with different specificities (related to short-chain alcohol dehydrogenases)	0.907	0.73	0.836	1.082
COG0477GEPR:Permeases of the major facilitator superfamily	0.955	0.561	0.733	0.853
COG1629P:Outer membrane receptor proteins, mostly Fe transport	0.764	0.553	1.057	1.082
COG2801L:Transposase and inactivated derivatives	0.334	0.438	0.275	0.423
COG3436L:Transposase and inactivated derivatives	0.191	0.415	0.093	0.072
COG0841V:Cation/multidrug efflux pump	0.43	0.384	0.6	0.616
COG0745TK:Response regulators consisting of a CheY-like receiver domain and a winged-helix DNA-binding domain	0.501	0.361	0.423	0.545
COG0582L:Integrase	0.477	0.353	0.192	0.566
COG0845M:Membrane-fusion protein	0.167	0.346	0.472	0.43
COG0840NT:Methyl-accepting chemotaxis protein	1.026	0.33	0.492	0.538
COG1012C:NAD-dependent aldehyde dehydrogenases	0.931	0.307	0.59	0.667
COG1538MU:Outer membrane protein	0.621	0.3	0.354	0.366
COG0784T:FOG: CheY-like receiver	0.406	0.284	0.172	0.409
COG0553KL:Superfamily II DNA/RNA helicases, SNF2 family	0.239	0.277	0.088	0.272
COG0673R:Predicted dehydrogenases and related proteins	0.095	0.277	0.197	0.036
COG3181S:Uncharacterized protein conserved in bacteria	0.024	0.269	0.703	0.846
COG0438M:Glycosyltransferase	0.167	0.254	0.374	0.215
COG2963L:Transposase and inactivated derivatives	0.143	0.246	0.093	0.186
COG0596R:Predicted hydrolases or acyltransferases (alpha/beta hydrolase superfamily)	0.597	0.238	0.462	0.423
COG0683E:ABC-type branched-chain amino acid transport systems, periplasmic component	0.072	0.231	0.536	0.344
COG2207K:AraC-type DNA-binding domain-containing proteins	0.382	0.223	0.393	0.516
COG0318IQ:Acyl-CoA synthetases (AMP-forming)/AMP-acid ligases II	0.215	0.215	0.379	0.194
COG0526OC:Thiol-disulfide isomerase and thioredoxins	0.119	0.215	0.152	0.28
COG1373R:Predicted ATPase (AAA+ superfamily)	0	0.215	0.039	0
COG2199T:FOG: GGDEF domain	0.358	0.215	0.147	0.294
COG3250G:Beta-galactosidase/beta-glucuronidase	0	0.215	0.059	0.014
COG0789K:Predicted transcriptional regulators	0.143	0.207	0.088	0.215

COG1024I:Enoyl-CoA hydratase/carnithine racemase	0.215	0.207	0.349	0.358
COG3039L:Transposase and inactivated derivatives, IS5 family	0.119	0.207	0.034	0.136
COG0625O:Glutathione S-transferase	0.31	0.2	0.27	0.323
COG0793M:Periplasmic protease	0.024	0.2	0.088	0.065
COG1131V:ABC-type multidrug transport system, ATPase component	0.095	0.2	0.143	0.1
COG2204T:Response regulator containing CheY-like receiver, AAA-type ATPase, and DNA-binding domains	0.286	0.2	0.211	0.136
COG0183I:Acetyl-CoA acetyltransferase	0.191	0.192	0.344	0.165
COG0834ET:ABC-type amino acid transport/signal transduction systems, periplasmic component/domain	0.477	0.192	0.202	0.265
COG0610V:Type I site-specific restriction-modification system, R (restriction) subunit and related helicases	0.024	0.184	0.044	0.065
COG0776L:Bacterial nucleoid DNA-binding protein	0.143	0.184	0.084	0.093
COG0286V:Type I restriction-modification system methyltransferase subunit	0.095	0.177	0.054	0.072
COG2202T:FOG: PAS/PAC domain	0.43	0.177	0.113	0.194
COG4584L:Transposase and inactivated derivatives	0.191	0.177	0.074	0.093
COG0494LR:NTP pyrophosphohydrolases including oxidative damage repair enzymes	0.167	0.169	0.088	0.158
COG3696P:Putative silver efflux pump	0.334	0.169	0.172	0.172
COG0209F:Ribonucleotide reductase, alpha subunit	0.048	0.161	0.098	0.05
COG0664T:cAMP-binding proteins - catabolite gene activator and regulatory subunit of cAMP-dependent protein kinases	0.072	0.161	0.202	0.108
COG1484L:DNA replication protein	0.119	0.161	0.074	0.072
COG2010C:Cytochrome c, mono- and diheme variants	0.072	0.161	0.093	0.179
COG0210L:Superfamily I DNA and RNA helicases	0.286	0.154	0.093	0.158
COG0366G:Glycosidases	0.072	0.154	0.074	0.065
COG0463M:Glycosyltransferases involved in cell wall biogenesis	0.095	0.154	0.251	0.143
COG0515RTKL:Serine/threonine protein kinase	0.072	0.154	0.064	0.136
COG0050J:GTPases - translation elongation factors	0.286	0.146	0.084	0.065
COG0451MG:Nucleoside-diphosphate-sugar epimerases	0.191	0.146	0.251	0.129
COG0543HC:2-polyprenylphenol hydroxylase and related flavodoxin oxidoreductases	0.239	0.146	0.064	0.108
COG0732V:Restriction endonuclease S subunits	0.024	0.146	0.044	0.086
COG1136V:ABC-type antimicrobial peptide transport system, ATPase component	0.072	0.146	0.093	0.057
COG3464L:Transposase and inactivated derivatives	0.095	0.146	0.039	0.036
COG3547L:Transposase and inactivated derivatives	0.143	0.146	0.044	0.057
COG0277C:FAD/FMN-containing dehydrogenases	0.119	0.138	0.211	0.308
COG0643NT:Chemotaxis protein histidine kinase and related kinases	0.167	0.138	0.044	0.1
COG0330O:Membrane protease subunits, stomatin/prohibitin homologs	0.263	0.131	0.138	0.108
COG0559E:Branched-chain amino acid ABC-type transport system, permease components	0.048	0.131	0.315	0.194

COG0697GER:Permeases of the drug/metabolite transporter (DMT) superfamily	0.263	0.131	0.29	0.244
COG2197TK:Response regulator containing a CheY-like receiver domain and an HTH DNA-binding domain	0.143	0.131	0.374	0.251
COG3344L:Retron-type reverse transcriptase	0.191	0.131	0.015	0.079
COG3507G:Beta-xylosidase	0	0.131	0.034	0.029
COG4638PR:Phenylpropionate dioxygenase and related ring-hydroxylating dioxygenases, large terminal subunit	0.31	0.131	0.177	0.158
COG0057G:Glyceraldehyde-3-phosphate dehydrogenase/erythrose-4-phosphate dehydrogenase	0.263	0.123	0.049	0.079
COG0265O:Trypsin-like serine proteases, typically periplasmic, contain C-terminal PDZ domain	0.024	0.123	0.162	0.079
COG1002V:Type II restriction enzyme, methylase subunits	0.024	0.123	0.025	0.072
COG2003L:DNA repair proteins	0.048	0.123	0.064	0.129
COG3177S:Uncharacterized conserved protein	0.048	0.123	0.034	0.036
COG0624E:Acetylornithine deacetylase/Succinyl-diaminopimelate desuccinylase and related deacylases	0.143	0.115	0.241	0.151
COG0635H:Coproporphyrinogen III oxidase and related Fe-S oxidoreductases	0.095	0.115	0.049	0.086
COG0665E:Glycine/D-amino acid oxidases (deaminating)	0.215	0.115	0.236	0.272
COG0668M:Small-conductance mechanosensitive channel	0.119	0.115	0.029	0.115
COG0826O:Collagenase and related proteases	0.072	0.115	0.044	0.029
COG1846K:Transcriptional regulators	0.072	0.115	0.226	0.129
COG2826L:Transposase and inactivated derivatives, IS30 family	0.143	0.115	0.025	0.086
COG0334E:Glutamate dehydrogenase/leucine dehydrogenase	0.048	0.108	0.02	0.065
COG0443O:Molecular chaperone	0.263	0.108	0.162	0.065
COG0534V:Na+-driven multidrug efflux pump	0.095	0.108	0.044	0.043
COG0835NT:Chemotaxis signal transduction protein	0.191	0.108	0.044	0.1
COG1092R:Predicted SAM-dependent methyltransferases	0.095	0.108	0.034	0.05
COG1192D:ATPases involved in chromosome partitioning	0.048	0.108	0.152	0.222
COG3335L:Transposase and inactivated derivatives	0.095	0.108	0.044	0.007
COG4177E:ABC-type branched-chain amino acid transport system, permease component	0.024	0.108	0.241	0.186
COG0084L:Mg-dependent DNase	0.024	0.1	0.029	0.057
COG0085K:DNA-directed RNA polymerase, beta subunit/140 kD subunit	0.263	0.1	0.147	0.129
COG0348C:Polyferredoxin	0.072	0.1	0.029	0.079
COG0410E:ABC-type branched-chain amino acid transport systems, ATPase component	0.072	0.1	0.211	0.115
COG0458EF:Carbamoylphosphate synthase large subunit (split gene in MJ)	0.215	0.1	0.093	0.043
COG0621J:2-methylthioadenine synthetase	0.143	0.1	0.103	0.043
COG0810M:Periplasmic protein TonB, links inner and outer membranes	0.024	0.1	0.059	0.122
COG1752R:Predicted esterase of the alpha-beta hydrolase superfamily	0.024	0.1	0.074	0.029
COG4147R:Predicted symporter	0.048	0.1	0.034	0.065

COG4775M:Outer membrane protein/protective antigen OMA87	0.048	0.1	0.059	0.029
COG0028EH:Thiamine pyrophosphate-requiring enzymes [acetolactate synthase, pyruvate dehydrogenase (cytochrome), glyoxylate carboligase, phosphonopyruvate decarboxylase]	0.143	0.092	0.143	0.208
COG0060J:Isoleucyl-tRNA synthetase	0.095	0.092	0.074	0.029
COG0156H:7-keto-8-aminopelargonate synthetase and related enzymes	0.024	0.092	0.059	0.029
COG0174E:Glutamine synthetase	0.286	0.092	0.069	0.057
COG0249L:Mismatch repair ATPase (MutS family)	0	0.092	0.059	0.029
COG0550L:Topoisomerase IA	0.119	0.092	0.044	0.136
COG0629L:Single-stranded DNA-binding protein	0	0.092	0.049	0.129
COG1414K:Transcriptional regulator	0.119	0.092	0.192	0.251
COG1479S:Uncharacterized conserved protein	0.024	0.092	0.025	0.043
COG1804C:Predicted acyl-CoA transferases/carnitine dehydratase	0.048	0.092	0.241	0.165
COG2200T:FOG: EAL domain	0.31	0.092	0.059	0.122
COG2890J:Methylase of polypeptide chain release factors	0.095	0.092	0.034	0.108
COG4206H:Outer membrane cobalamin receptor protein	0.048	0.092	0.025	0.029
COG4644L:Transposase and inactivated derivatives, TnpA family	0.31	0.092	0.059	0.108
COG0402FR:Cytosine deaminase and related metal-dependent hydrolases	0.215	0.085	0.074	0.1
COG0612R:Predicted Zn-dependent peptidases	0.119	0.085	0.143	0.036
COG1082G:Sugar phosphate isomerases/epimerases	0.143	0.085	0.069	0.036
COG1475K:Predicted transcriptional regulators	0.072	0.085	0.167	0.201
COG1522K:Transcriptional regulators	0.191	0.085	0.172	0.115
COG1609K:Transcriptional regulators	0.048	0.085	0.241	0.108
COG0069E:Glutamate synthase domain 2	0.143	0.077	0.049	0.065
COG0082E:Chorismate synthase	0.119	0.077	0.025	0.043
COG0365I:Acyl-coenzyme A synthetases/AMP-(fatty) acid ligases	0.167	0.077	0.177	0.115
COG0569P:K+ transport systems, NAD-binding component	0.024	0.077	0.02	0.014
COG0715P:ABC-type nitrate/sulfonate/bicarbonate transport systems, periplasmic components	0.095	0.077	0.177	0.129
COG0798P:Arsenite efflux pump ACR3 and related permeases	0.048	0.077	0.025	0.036
COG1062C:Zn-dependent alcohol dehydrogenases, class III	0.072	0.077	0.039	0.108
COG1236J:Predicted exonuclease of the beta-lactamase fold involved in RNA processing	0.024	0.077	0.044	0.1
COG1289S:Predicted membrane protein	0.048	0.077	0.034	0.108
COG1555L:DNA uptake protein and related DNA-binding proteins	0.024	0.077	0.025	0.022
COG1974KT:SOS-response transcriptional repressors (RecA-mediated autopeptidases)	0	0.077	0.025	0.093
COG2378K:Predicted transcriptional regulator	0.048	0.077	0.044	0.1

COG2823R:Predicted periplasmic or secreted lipoprotein	0.143	0.077	0.054	0.036
COG0288P:Carbonic anhydrase	0.119	0.069	0.039	0.036
COG0465O:ATP-dependent Zn proteases	0.167	0.069	0.059	0.043
COG0469G:Pyruvate kinase	0.143	0.069	0.039	0.029
COG0520E:Selenocysteine lyase	0.143	0.069	0.044	0.079
COG0617J:tRNA nucleotidyltransferase/poly(A) polymerase	0.167	0.069	0.059	0.05
COG1047O:FKBP-type peptidyl-prolyl cis-trans isomerases 2	0.095	0.069	0.025	0.036
COG1826U:Sec-independent protein secretion pathway components	0.024	0.069	0.02	0.05
COG0005F:Purine nucleoside phosphorylase	0.024	0.061	0.02	0.014
COG0179Q:2-keto-4-pentenoate hydratase/2-oxohepta-3-ene-1,7-dioic acid hydratase (catechol pathway)	0.072	0.061	0.167	0.143
COG0329EM:Dihydrodipicolinate synthase/N-acetylneuraminate lyase	0.095	0.061	0.138	0.115
COG0392S:Predicted integral membrane protein	0.048	0.061	0.02	0.029
COG0503F:Adenine/guanine phosphoribosyltransferases and related PRPP-binding proteins	0	0.061	0.02	0.022
COG0763M:Lipid A disaccharide synthetase	0.024	0.061	0.02	0.029
COG1066O:Predicted ATP-dependent serine protease	0.048	0.061	0.02	0.043
COG1167KE:Transcriptional regulators containing a DNA-binding HTH domain and an aminotransferase domain (MocR family) and their eukaryotic orthologs	0.31	0.061	0.206	0.215
COG1197LK:Transcription-repair coupling factor (superfamily II helicase)	0.119	0.061	0.128	0.043
COG1230P:Co/Zn/Cd efflux system component	0.072	0.061	0.049	0.108
COG1250I:3-hydroxyacyl-CoA dehydrogenase	0.191	0.061	0.157	0.129
COG1352NT:Methylase of chemotaxis methyl-accepting proteins	0.167	0.061	0.049	0.093
COG1651O:Protein-disulfide isomerase	0.024	0.061	0.054	0.129
COG1825J:Ribosomal protein L25 (general stress protein Ctc)	0.048	0.061	0.02	0.022
COG3279KT:Response regulator of the LytR/AlgR family	0.024	0.061	0.202	0.029
COG3971Q:2-keto-4-pentenoate hydratase	0.048	0.061	0.025	0.086
COG4942D:Membrane-bound metallopeptidase	0.024	0.061	0.02	0.029
COG0046F:Phosphoribosylformylglycinamidine (FGAM) synthase, synthetase domain	0.143	0.054	0.049	0.057
COG0047F:Phosphoribosylformylglycinamidine (FGAM) synthase, glutamine amidotransferase domain	0.119	0.054	0.034	0.057
COG0150F:Phosphoribosylaminoimidazole (AIR) synthetase	0.095	0.054	0.025	0.057
COG0154J:Asp-tRNAAsn/Glu-tRNAGln amidotransferase A subunit and related amidases	0.119	0.054	0.162	0.179
COG0456R:Acetyltransferases	0.095	0.054	0.025	0.036
COG0460E:Homoserine dehydrogenase	0.024	0.054	0.02	0.065
COG0693R:Putative intracellular protease/amidase	0.024	0.054	0.064	0.129
COG0751J:Glycyl-tRNA synthetase, beta subunit	0.143	0.054	0.054	0.022

COG0753P:Catalase	0.167	0.054	0.054	0.1
COG1404O:Subtilisin-like serine proteases	0.072	0.054	0.084	0.165
COG1529C:Aerobic-type carbon monoxide dehydrogenase, large subunit CoxL/CutL homologs	0.095	0.054	0.123	0.057
COG1593G:TRAP-type C4-dicarboxylate transport system, large permease component	0.072	0.054	0.143	0.086
COG1734T:DnaK suppressor protein	0.048	0.054	0.034	0.093
COG1802K:Transcriptional regulators	0.167	0.054	0.187	0.129
COG1894C:NADH:ubiquinone oxidoreductase, NADH-binding (51 kD) subunit	0.048	0.054	0.034	0.086
COG2201NT:Chemotaxis response regulator containing a CheY-like receiver domain and a methylesterase domain	0.167	0.054	0.074	0.093
COG2303E:Choline dehydrogenase and related flavoproteins	0.048	0.054	0.108	0.129
COG3311K:Predicted transcriptional regulator	0.024	0.054	0.02	0.079
COG3706T:Response regulator containing a CheY-like receiver domain and a GGDEF domain	0.143	0.054	0.029	0.108
COG0112E:Glycine/serine hydroxymethyltransferase	0.215	0.046	0.069	0.043
COG0531E:Amino acid transporters	0.239	0.046	0.157	0.1
COG1020Q:Non-ribosomal peptide synthetase modules and related proteins	0.286	0.046	0.162	0.129
COG1049C:Aconitase B	0.048	0.046	0.034	0.122
COG1335Q:Amidases related to nicotinamidase	0.119	0.046	0.147	0.108
COG1570L:Exonuclease VII, large subunit	0.143	0.046	0.025	0.029
COG2373R:Large extracellular alpha-helical protein	0.119	0.046	0.133	0.057
COG3203M:Outer membrane protein (porin)	0	0.046	0.118	0.108
COG3264M:Small-conductance mechanosensitive channel	0.167	0.046	0.01	0.029
COG3451U:Type IV secretory pathway, VirB4 components	0.072	0.046	0.098	0.337
COG3481R:Predicted HD-superfamily hydrolase	0	0.046	0.015	0.136
COG0151F:Phosphoribosylamine-glycine ligase	0.095	0.038	0.025	0.029
COG0190H:5,10-methylene-tetrahydrofolate dehydrogenase/Methenyl tetrahydrofolate cyclohydrolase	0.095	0.038	0.02	0.029
COG0196H:FAD synthase	0.119	0.038	0.029	0.029
COG0251J:Putative translation initiation inhibitor, yjgF family	0.072	0.038	0.108	0.115
COG0399M:Predicted pyridoxal phosphate-dependent enzyme apparently involved in regulation of cell wall biogenesis	0.024	0.038	0.093	0.029
COG0471P:Di- and tricarboxylate transporters	0	0.038	0.02	0.065
COG0507L:ATP-dependent exoDNAse (exonuclease V), alpha subunit - helicase superfamily I member	0	0.038	0.103	0.086
COG0601EP:ABC-type dipeptide/oligopeptide/nickel transport systems, permease components	0.024	0.038	0.143	0.151
COG0747E:ABC-type dipeptide transport system, periplasmic component	0.095	0.038	0.246	0.222
COG1670J:Acetyltransferases, including N-acetylases of ribosomal proteins	0.072	0.038	0.113	0.079
COG1966T:Carbon starvation protein, predicted membrane protein	0.095	0.038	0.02	0.022
COG2165NU:Type II secretory pathway, pseudopilin PulG	0	0.038	0.025	0.072

COG3842E:ABC-type spermidine/putrescine transport systems, ATPase components	0.119	0.038	0.103	0.079
COG4569Q:Acetaldehyde dehydrogenase (acetylating)	0.119	0.038	0.015	0.029
COG4948MR:L-alanine-DL-glutamate epimerase and related enzymes of enolase superfamily	0.119	0.038	0.143	0.079
COG4977K:Transcriptional regulator containing an amidase domain and an AraC-type DNA-binding HTH domain	0.191	0.038	0.064	0.079
COG0308E:Aminopeptidase N	0.119	0.031	0.123	0.043
COG0439I:Biotin carboxylase	0.119	0.031	0.025	0.072
COG0765E:ABC-type amino acid transport system, permease component	0.119	0.031	0.202	0.136
COG0807H:GTP cyclohydrolase II	0.095	0.031	0.025	0.029
COG0843C:Heme/copper-type cytochrome/quinol oxidases, subunit 1	0.143	0.031	0.108	0.129
COG1171E:Threonine dehydratase	0.072	0.031	0.103	0.072
COG1304C:L-lactate dehydrogenase (FMN-dependent) and related alpha-hydroxy acid dehydrogenases	0	0.031	0.113	0.072
COG1305E:Transglutaminase-like enzymes, putative cysteine proteases	0.048	0.031	0.074	0.093
COG1322S:Uncharacterized protein conserved in bacteria	0.095	0.031	0.02	0.029
COG1391OT:Glutamine synthetase adenylyltransferase	0.119	0.031	0.039	0.057
COG1450NU:Type II secretory pathway, component PulD	0.024	0.031	0.02	0.072
COG1653G:ABC-type sugar transport system, periplasmic component	0	0.031	0.098	0.079
COG1737K:Transcriptional regulators	0.119	0.031	0.059	0.043
COG1793L:ATP-dependent DNA ligase	0.119	0.031	0.044	0.065
COG1807M:4-amino-4-deoxy-L-arabinose transferase and related glycosyltransferases of PMT family	0	0.031	0.088	0.065
COG2124Q:Cytochrome P450	0.095	0.031	0.025	0.022
COG2234R:Predicted aminopeptidases	0.119	0.031	0.069	0.072
COG2274V:ABC-type bacteriocin/lantibiotic exporters, contain an N-terminal double-glycine peptidase domain	0.048	0.031	0.098	0.022
COG3523S:Uncharacterized protein conserved in bacteria	0.143	0.031	0.098	0.029
COG4993G:Glucose dehydrogenase	0.048	0.031	0.059	0.093
COG0702MG:Predicted nucleoside-diphosphate-sugar epimerases	0.024	0.023	0.108	0.043
COG2894D:Septum formation inhibitor-activating ATPase	0.095	0.023	0.025	0.043
COG2931Q:RTX toxins and related Ca2+-binding proteins	0.597	0.023	0.369	0.136
COG3459G:Cellobiose phosphorylase	0.119	0.023	0.025	0.057
COG4585T:Signal transduction histidine kinase	0.024	0.023	0.138	0.065
COG5001T:Predicted signal transduction protein containing a membrane domain, an EAL and a GGDEF domain	0.191	0.023	0.025	0.086
COG0270L:Site-specific DNA methylase	0.048	0.015	0.034	0.086
COG1113E:Gamma-aminobutyrate permease and related permeases	0.072	0.015	0.02	0.079
COG2766T:Putative Ser protein kinase	0.119	0.015	0.02	0.029
COG3209M:Rhs family protein	0.334	0.015	0.197	0.036

COG3210U:Large exoproteins involved in heme utilization or adhesion	0.143	0.015	0.133	0.294
COG3419NU:Tfp pilus assembly protein, tip-associated adhesin PilY1	0.024	0.015	0.02	0.072
COG4773P:Outer membrane receptor for ferric coprogen and ferric-rhodotorulic acid	0.143	0.015	0.034	0.093
COG2113E:ABC-type proline/glycine betaine transport systems, periplasmic components	0.167	0.008	0.029	0.014
COG2831U:Hemolysin activation/secretion protein	0.143	0.008	0.044	0.072
COG3285L:Predicted eukaryotic-type DNA primase	0.095	0.008	0.02	0.036
COG3501S:Uncharacterized protein conserved in bacteria	0.263	0.008	0.069	0.136

[1] COG category in included: A (RNA processing and modification), B (Chromatin structure and dynamics), C (Energy production and conversion), D (Cell cycle control and mitosis), E (Amino acid metabolism and transport), F (Nucleotide metabolism and transport), G (Carbohydrate metabolism and transport), H (Coenzyme metabolism), I (Lipid metabolism), J (Translation), K (Transcription), L (Replication and repair), M (Cell wall/membrane/envelop biogenesis), N (Cell otility), O (Post-translational modification, protein turnover, chaperone functions), Q (Secondary structure), T (Signal transduction), U (Intracellular trafficking and secretion), Y (Nuclear structure), Z (Cytoskeleton), R (General function prediction only) and S (Function unknown).

[2] Percentage was calculated based on the total number of genes per COG for a given metagenome, as indicated in Pannel (A), and the total number of genes indicated in Table S3: 5211 (for N), 21130 (for Nbs), 27124 (for CN1) and 18391 (for CN2).

Table S5 List of enzymes found to contribute to the major aromatics aerobic degradation pathways via di- and trihydroxylated intermediates identified in the metagenomes of samples N, Nbs, CN1 and CN2.

(A) CN1 enzymes features

List of unique proteins[1]	**Phylogeny**[2]	**ORF**[1]	**Contig**[1]	**FAMILY**[3]	**SUBFAMILY**[3]	**BEST HITS**[4]
CN1-1 = CN2-1	*Betaproteobacteria (Achromobacter)*	CN1/10033	CN1contig06710	Rieske ferredoxin	Naphthalene dioxygenase U2 /Salicylate 5-hydroxylase	100% *Ralstonia* sp. U2 (AAD12609)
		CN1/10034	CN1contig06710	Rieske alpha	Naphthalene dioxygenase (*Proteobacteria*-like)	99% *Ralstonia* sp. U2 (AAD12610)
		CN1/21052	CN1contig15209	Rieske beta	Naphthalene dioxygenase	89% *Ralstonia* sp. U2 (AP12611)
		CN1/21053	CN1contig15209	Rieske alpha	Naphthalene dioxygenase (*Proteobacteria*-like)	98% *Ralstonia* sp. U2 (AP12610)
CN1-2	*Betaproteobacteria (Achromobacter)*	CN1/11872	CN1contig08008	Rieske ferredoxin	Naphthalene dioxygenase/salicylate 5-hydroxylase	100% *Ralstonia* sp. U2 (AAD12606)
CN1-3	*Alphaproteobacteria (Azospirillum)*	CN1/3872	CN1contig02509	Cupin	Gentisate dioxygenase	53% *Ralstonia* sp. U2 (AAD12619)
CN1-4 = CN2-19	*Betaproteobacteria (Achromobacter)*	CN1/3348	CN1contig02145	Cupin	Gentisate dioxygenase	96% *Achromobacter piechaudii* (ZP_06684540)
		CN1/12318	CN1contig08355	Cupin	Gentisate dioxygenase	58% *Bradyrhizobium* sp. BTAi1 (Bbta_7501; YP_001243259)
CN1-5 = CN2-18	*Betaproteobacteria (Comamonas)*	CN1/5183	CN1contig03376	Cupin	Gentisate dioxygenase	65% *Comamonas testosteroni* KF-1 (ZP_03542624)
	Betaproteobacteria (Comamonas)	CN1/21802	CN1contig15797	Cupin	Gentisate dioxygenase	86% *Comamonas testosteroni* KF-1 (ZP_03542624)
CN1-6 = CN2-17	*Betaproteobacteria (Achromobacter)*	CN1/5541	CN1contig03648	Cupin	Gentisate dioxygenase	100% *Ralstonia* sp. U2 (AAD12619)
		CN1/6354	CN1contig04174	Cupin	Gentisate dioxygenase	100% *Ralstonia* sp. U2 (AAD12619)
		CN1/23999	CN1contig17568	Cupin	Gentisate dioxygenase	100% *Ralstonia* sp. U2 (AAD12619)
CN1-7	*Betaproteobacteria (Achromobacter)*	CN1/0487	CN1contig00295	Cupin	Gentisate dioxygenase	58% *Achromobacter xylooxydans* (EGP45942)
CN1-8 = CN2-3	*Betaproteobacteria (Achromobacter)*	CN1/14978	CN1contig10341	EXDO	Dihydroxynaphthalene dioxygenase	100% *Ralstonia* sp. U2 (AP12614)
CN1-9	*Alphaproteobacteria (Azospirillum)*	CN1/0554	CN1contig00338	Other protein families (isomerase)	2-Hydroxychromene-2-carboxylate isomerase	89% NahD *Azospirillum* B510 Azl004590 (YP_003447641)
CN1-10	*Betaproteobacteria (Achromobacter)*	CN1/24001	CN1contig17568	Other protein families (isomerase)	2-Hydroxychromene-2-carboxylate isomerase	100% *Ralstonia* sp. U2 (AAD12617)
CN1-11	*Betaproteobacteria (Achromobacter)*	CN1/23998	CN1contig17567	Other protein families (aldolase)	trans-o-Hydroxybenzylidenepyruvate hydratase-aldolase	99% *Ralstonia* sp. U2 (AAD12166)
		CN1/24002	CN1contig17568	Other protein families (aldolase)	trans-o-Hydroxybenzylidenepyruvate hydratase-aldolase (NagE-like)	100% *Ralstonia* sp. U2 (AAD12616)

CN1-12	*Betaproteobacteria (Achromobacter)*	CN1/14979	CN1contig10341	Other protein families (dehydrogenase)	Salicylaldehyde dehydrogenase	98% *Ralstonia* sp. U2 (AAD12613)
CN1-13	*Betaproteobacteria (Achromobacter)*	CN1/5933	CN1contig03894	Rieske alpha	Salicylate 5-hydroxylase	96% *Ralstonia* sp. U2 (AAD12607)
		CN1/11871	CN1contig08008	Rieske alpha	Salicylate 5-hydroxylase	96% *Ralstonia* sp. U2 (AAD12607)
		CN1/20818	CN1contig15044	Rieske alpha	Salicylate 5-hydroxylase	100% *Ralstonia* sp. U2 (AAD12607)
		CN1/20821	CN1contig15047	Rieske alpha	Salicylate 5-hydroxylase	100% *Ralstonia* sp. U2 (AAD12607)
		CN1/26968	CN1contig20657	Rieske alpha	Salicylate 5-hydroxylase	98% *Ralstonia* sp. U2 (AAD12607)
		CN1/26969	CN1contig20658	Rieske alpha	Salicylate 5-hydroxylase	81% *Ralstonia* sp. U2 (AAD12607)
		CN1/27027	CN1contig20949	Rieske alpha	Salicylate 5-hydroxylase	91% *Ralstonia* sp. U2 (AAD12607)
		CN1/10032	CN1contig06710	Rieske beta	Salicylate 5-hydroxylase	100% *Ralstonia* sp. U2 (AAD12608)
		CN1/20824	CN1contig15049	Rieske beta	Salicylate 5-hydroxylase	78% *Ralstonia* sp. U2 (AP12608)
CN1-14	*Alphaproteobacteria (Azospirillum)*	CN1/5223	CN1contig03408	Rieske reductase	Salicylate 5-hydroxylase	71% *Azospirillum* B510 (Azl a9220)
		CN1/5224	CN1contig03408	Rieske ferredoxin	Salicylate 5-hydroxylase	71% *Azospirillum* B510 (Azl a9190)
CN1-15 = CN2-11	*Betaproteobacteria (Achromobacter)*	CN1/5860	CN1contig03845	Rieske alpha	Salicylate 5-hydroxylase	98% *Achromobacter piechaudii* (ZP_06686009)
		CN1/5861	CN1contig03845	Rieske beta	Salicylate 5-hydroxylase	95% *Achromobacter piechaudii* (ZP_06686008)
		CN1/5862	CN1contig03845	Rieske ferredoxin	Salicylate 5-hydroxylase	93% *Achromobacter piechaudii* (ZP_06686007)
		CN1/21472	CN1contig15525	Rieske alpha	Salicylate 5-hydroxylase	100% *Achromobacter piechaudii* (ZP_06686009)
CN1-16 = CN2-13	*Betaproteobacteria (Achromobacter)*	CN1/5934	CN1contig03894	Rieske beta	Salicylate 5-hydroxylase	73% *Polaromonas* sp. JS666 (Bpro_09849)
		CN1/5935	CN1contig03894	Rieske ferredoxin	Salicylate 5-hydroxylase	62% *Polaromonas* sp. JS666 (Bpro_0985)
		CN1/20819	CN1contig15045	Rieske alpha	Salicylate 5-hydroxylase	83% *Polaromonas* sp. JS666 (Bpro_0983)
CN1-17	*Betaproteobacteria (Achromobacter)*	CN1/20820	CN1contig15046	Rieske alpha	Salicylate 5-hydroxylase	77% *Ralstonia eutropha* JMP134 (ReutB4719)
CN1-18	*Alphaproteobacteria (Azospirillum)*	CN1/0489	CN1contig00295	Other protein families (hydrolase)	Fumarylpyruvate hydrolase	81% *Azospirillum* B510 (Azl a9250)
		CN1/3873	CN1contig02509	Other protein families (hydrolase)	Fumarylpyruvate hydrolase	
CN1-19	*Betaproteobacteria (Achromobacter)*	CN1/3347	CN1contig02145	Other protein families (hydrolase)	Fumarylpyruvate hydrolase	92% *Achromobacter piechaudii* (ZP_06684541)
CN1-20	*Betaproteobacteria (Comamonas)*	CN1/5182	CN1contig03376	Other protein families (hydrolase)	Fumarylpyruvate hydrolase	72% *Comamonas testosteroni* S44 (ZP_07045767)
CN1-21	*Betaproteobacteria (Comamonas)*	CN1/6353	CN1contig04174	Other protein families (hydrolase)	Fumarylpyruvate hydrolase	77% *Comamonas testosteroni* S44 (YP_0032788219)
CN1-22	*Betaproteobacteria (Comamonas)*	CN1/12017	CN1contig08111	Flavoprotein monooxygenase	3-Hydroxybenzoate 6-hydroxylase	100% *Comamonas testosteroni* KF-1 (ZP_03541756)
CN1-23	*Betaproteobacteria (Achromobacter)*	CN1/18365	CN1contig12835	EXDO	2,3-Dihydroxybiphenyl dioxygenase (*Bacillus*-like)	57% BphC *Bacillus* sp. JF8 (Q8GR45)
CN1-24 = CN2-29	*Betaproteobacteria (Achromobacter)*	CN1/1576	CN1contig01024	Rieske alpha	Function unknown (distantly related to benzoate dioxygenase)	99% *Achromobacter piechaudii* (ZP_06686030)
CN1-25 = CN2-28	*Betaproteobacteria (Achromobacter)*	CN1/22427	CN1contig16308	Rieske alpha	Function unknown (distantly related to benzoate dioxygenase)	99% *Achromobacter piechaudii* (ZP_06685665)
CN1-26	*Betaproteobacteria* (genus	CN1/6284	CN1contig04134	Other protein families	Phenol hydroxylase	77% *Alicycliphilus denitrificans* BC (Alide0326)

	unclear)					
CN1-27	*Alphaproteobacteria* (genus unclear)	CN1/6610	CN1contig04340	EXDO	Catechol 2,3-dioxygenase (*Sphingomonads*-like)	61% *Sphingomonas* sp. P2 (BAC65437)
		CN1/6777	CN1contig04464	EXDO	Catechol 2,3-dioxygenase	
CN1-28	*Betaproteobacteria* (genus unclear)	CN1/12486	CN1contig08481	Flavoprotein monooxygenase	4-Hydroxybenzoate 3-hydroxylase	59% *Burkholderia xenovorans* LB400 (Bxe_A2040)
CN1-29	*Betaproteobacteria (Comamonas)*	CN1/13178	CN1contig08971	Flavoprotein monooxygenase	4-Hydroxybenzoate 3-hydroxylase	100% *Comamonas testosteroni* KF-1 (ZP_03543395)
		CN1/24144	CN1contig17697	Flavoprotein monooxygenase	4-Hydroxybenzoate 3-hydroxylase	100% *Comamonas testosteroni* KF-1 (ZP_03543395)
CN1-30	*Alphaproteobacteria (Azospirillum)*	CN1/26430	CN1contig19887	Flavoprotein monooxygenase	4-Hydroxybenzoate 3-hydroxylase	91% *Azospirillum* B510 (YP_003452470)
CN1-31	*Betaproteobacteria (Comamonas)*	CN1/11127	CN1contig07470	Rieske alpha	Vanillate monooxygenase	99% *Comamonas testosteroni* KF-1 (ZP_03541171)
CN1-32 = CN2-70	*Betaproteobacteria (Achromobacter)*	CN1/18221	CN1contig12726	Rieske alpha	Vanillate monooxygenase	100% *Achromobacter xylosooxidans* A8 (AXYL02987)
CN1-33	*Alphaproteobacteria (Azospirillum)*	CN1/5532	CN1contig03643	INTRA	Protocatechuate 3,4-dioxygenase beta subunit	91% *Azospirillum* B510 (Azl d00470)
		CN1/21609	CN1contig15627	INTRA	Protocatechuate 3,4-dioxygenase alpha subunit	89% *Azospirillum* B510 (YP_003452416)
		CN1/23374	CN1contig17041	INTRA	Protocatechuate 3,4-dioxygenase alpha subunit	87% *Azospirillum* B510 (d00460)
CN1-34	*Betaproteobacteria (Achromobacter)*	CN1/4160	CN1contig02693	INTRA	Protocatechuate 3,4-dioxygenase beta subunit	81% *Ralstonia eutropha* JMP134 (ReutB5025)
		CN1/7682	CN1contig05094	INTRA	Protocatechuate 3,4-dioxygenase alpha subunit	78% *Ralstonia eutropha* JMP134 (ReutB5024)
CN1-35	*Betaproteobacteria (Comamonas)*	CN1/18367	CN1contig12837	LigB	Protocatechuate 4,5-dioxygenase beta subunit	100% *Comamonas testosteroni* KF-1 (ZP_03542666)
		CN1/18368	CN1contig12837	LigB	Protocatechuate 4,5-dioxygenase alpha subunit	100% *Comamonas testosteroni* KF-1 (ZP_03542665)
CN1-36	*Betaproteobacteria* (genus unclear)	CN1/2889	CN1contig01853	LigB	Protocatechuate 4,5-dioxygenase	64% *Azoarcus* sp. BH72 (YP_934026)
CN1-37	*Alphaproteobacteria (Azospirillum)*	CN1/7681	CN1contig05094	Other protein families (isomerase)	Carboxymuconate cycloisomerase	54% *Acinetobacter* sp. SH024 (ZP_06690230)
CN1-38	*Alphaproteobacteria (Azospirillum)*	CN1/11990	CN1contig08096	Other protein families (reductase)	Maleylacetate reductase	49% *Burkholderia pseudomallei* MSHR346 (YP_004330781)
CN1-39	*Bacteroidetes*	CN1/4680	CN1contig03032	Other protein families (non-heme dioxygenase)	4-Hydroxyphenylpyruvate dioxygenase	78% *Chitinophaga pinensis* (Cpin5420)
		CN1/14773	CN1contig10195	Other protein families (non-heme dioxygenase)	4-Hydroxyphenylpyruvate dioxygenase	73% *Chitinophaga pinensis* (Cpin5420)
CN1-40 = CN2-57	*Gammaproteobacteria* (genus unclear)	CN1/10673	CN1contig07168	Other protein families (non-heme dioxygenase)	4-Hydroxyphenylpyruvate dioxygenase	83% *Stenotrophomonas maltophilia* R551-3 (Saml3739)
CN1-41	*Betaproteobacteria (Comamonas)*	CN1/15465	CN1contig10702	Other protein families (non-heme dioxygenase)	4-Hydroxyphenylpyruvate dioxygenase	100% *Comamonas testosteroni* KF-1 (ZP_0354115)
CN1-42	*Gammaproteobacteria (Legionella)*	CN1/21045	CN1contig15205	Other protein families (non-heme dioxygenase)	4-Hydroxyphenylpyruvate dioxygenase	78% *Legionella drancourtii* LLAP12 (ZP_05110703)
CN1-43	*Alphaproteobacteria (Azospirillum)*	CN1/21429	CN1contig15498	Other protein families (non-heme dioxygenase)	4-Hydroxyphenylpyruvate dioxygenase	93% *Azospirillum* B510 (YP_003452613)

CN1-44	ORF too short for assigning phylogeny (homology for the 92 AA long fragment <50%)	CN1/24350	CN1contig17873	Other protein families (non-heme dioxygenase)	4-Hydroxyphenylpyruvate dioxygenase	49% *Mangifera indica* (ADO62712.1)
CN1-45	*Alphaproteobacteria (Azospirillum)*	CN1/0644	CN1contig00404	Cupin	Homogentisate dioxygenase	95% *Azospirillum* B510 (Azl024670)
		CN1/20618	CN1contig14896	Cupin	Homogentisate dioxygenase	94% *Azospirillum* B510 (YP_003449649)
CN1-46	*Alphaproteobacteria (Azospirillum)*	CN1/1252	CN1contig00814	Cupin	Homogentisate dioxygenase	95% *Azospirillum* B510 (Azl024670)
CN1-47	*Bacteroidetes*	CN1/3844	CN1contig02492	Cupin	Homogentisate dioxygenase	81% *Chitinophaga pinensis* (Cpin5427)
CN1-48	*Gammaproteobacteria (Pseudoxanthomonas)*	CN1/14821	CN1contig10230	Cupin	Homogentisate dioxygenase	81% *Pseudoxanthomonas suwonensis* 11-1 (Psesu2515
CN1-49	*Gammaproteobacteria (Legionella)*	CN1/19115	CN1contig13367	Cupin	Homogentisate dioxygenase	75% *Legionella drancourtii* LLAP12 (ZP_05109390)
CN1-50 = CN2-74	*Betaproteobacteria (Achromobacter)*	CN1/13083	CN1contig08896	LigB	Homoprotocatechuate dioxygenase	90% *Achromobacter piechaudii* (ZP_06690164)
		CN1/19314	CN1contig13509	LigB	Homoprotocatechuate dioxygenase	96% *Achromobacter piechaudii* (ZP_06690164)
CN1-51	*Betaproteobacteria (Achromobacter)*	CN1/0215	CN1contig00133	EXDO	2,3-Dihydroxy-p-cumate dioxygenase	58% *Achromombacter xylosooxidans* A8 (AXYL_02525)
CN1-52	*Betaproteobacteria* (genus unclear)	CN1/1319	CN1contig00858	LigB	Polycyclic arene diol extradiol dioxygenase (PhnC-like)	76% PhnC *Burkholderia sartisoli* (AAD09870)
		CN1/22639	CN1contig16483	LigB	Polycyclic arene diol extradiol dioxygenase (PhnC-like)	68% PhnC *Burkholderia sartisoli* (AAD09870)
CN1-53	*Betaproteobacteria* (genus unclear)	CN1/2698	CN1contig01726	Rieske alpha	Abietane diterpenoid dioxygenase	60% *Pseudomonas aeruginosa* 2192 (ZP_04934623)
		CN1/19902	CN1contig14229	Rieske alpha	Abietane diterpenoid dioxygenase	57% *Pseudomonas aeruginosa* 2192 (ZP_04934643)
		CN1/2699	CN1contig01726	Rieske beta	Abietane diterpenoid dioxygenase	57% *Burkholderia xenovorans* LB400 (Bxe_C0587)
		CN1/1963	CN1contig01249	Rieske beta	Abietane diterpenoid dioxygenase	57% *Burkholderia xenovorans* LB400 (Bxe_C0587)
CN1-54	*Betaproteobacteria* (genus unclear)	CN1/2697	CN1contig01726	EXDO	7-Oxo-11,12-dihydroxydehydroabietic acid dioxygenase	78% *Pseudomonas aeruginosa* 2192 (ZP_04934632)
		CN1/19780	CN1contig14044	EXDO	7-Oxo-11,12-dihydroxydehydroabietic acid dioxygenase	73% *Pseudomonas aeruginosa* 2192 (ZP_04934642)
		CN1/19781	CN1contig14045	EXDO	7-Oxo-11,12-dihydroxydehydroabietic acid dioxygenase	73% *Pseudomonas aeruginosa* 2192 (ZP_04934642)
		CN1/19782	CN1contig14046	EXDO	7-Oxo-11,12-dihydroxydehydroabietic acid dioxygenase	70% *Pseudomonas aeruginosa* 2192 (ZP_04934642)
		CN1/19813	CN1contig14070	EXDO	7-Oxo-11,12-dihydroxydehydroabietic acid dioxygenase	73% *Pseudomonas aeruginosa* 2192 (ZP_04934642)
CN1-55 = CN2-78	*Betaproteobacteria (Achromobacter)*	CN1/4264	CN1contig02763	LigB	2-Aminophenol 1,6-dioxygenase	89% *Burkholderia xenovorans* LB400 (Bxe_A1145)
CN1-56	*Betaproteobacteria (Comamonas)*	CN1/5605	CN1contig03686	LigB	2,3-Dihydroxyphenylpropionate dioxygenase	67% *Comamonas testosteroni* S44 (ZP_07047594)

CN1-57	*Betaproteobacteria (Comamonas)*	CN1/9860	CN1contig06591	Rieske alpha	Terephthalate dioxygenase	99% *Comamonas testosteroni* KF-1 (ZP_03542898)
CN1-58	*Betaproteobacteria (Comamonas)*	CN1/14740	CN1contig10167	Flavoprotein monooxygenase	6-Hydroxynicotinate 3-monooxygenase	99% *Comamonas testosteroni* KF-1 (ZP_03541517)
CN1-59	*Betaproteobacteria* (genus unclear)	CN1/19559	CN1contig13684	Rieske ferredoxin	Function unknown (possibly invoved in rhizopine degradation)	92% *Rhizobium etli* CIAT894 (ZP_03528797)
CN1-60	*Gammaproteobacteria*(genus unclear)	CN1/21845	CN1contig15836	Rieske alpha	Dioxygenase (possibly involved in indolacetic acid degradation)	80% LacC *Pseudomonas fluorescens* WH6 (ZP_07775391)
CN1-61	*Betaproteobacteria (Comamonas)*	CN1/6909	CN1contig04543	Rieske alpha	Phthalate 4,5-dioxygenase	100% *Comamonas testosteroni* CNB-2 (CtCNB1_1488)
		CN1/18929	CN1contig13244	Rieske alpha	Phthalate dioxygenase	100% *Comamonas testosteroni* CNB-2 (CtCNB1_1488)
CN1-62	*Betaproteobacteria (Comamonas)*	CN1/11422	CN1contig07683	Rieske alpha	Phthalate dioxygenases (Proteobacteria-like)	100% *Comamonas testosteroni* KF-1 (ZP_03542656)
CN1-63	*Betaproteobacteria (Achromobacter)*	CN1/15660	CN1contig10837	Rieske	Phthalate4,5- dioxygenase	71% *Polaromonas* sp. JS666(Bpro_19110)
		CN1/20322	CN1contig14660	Rieske alpha	Phthalate 4,5-dioxygenase	76% *Polaromonas* sp. JS666 (Bpro_19110)
CN1-64	*Betaproteobacteria (Achromobacter)*	CN1/20324	CN1contig14662	Rieske alpha	Phthalate 4,5-dioxygenase	71% *Polaromonas* sp. JS666 (Bpro_19110)
		CN1/20325	CN1contig14663	Rieske alpha	Phthalate 4,5-dioxygenase	72% *Polaromonas* sp. JS666 (Bpro_19110)
		CN1/20327	CN1contig14665	Rieske alpha	Phthalate 4,5-dioxygenase	58% *Polaromonas* sp. JS666 (Bpro_19110)
		CN1/20328	CN1contig14666	Rieske alpha	Phthalate 4,5-dioxygenase	73% *Polaromonas* sp. JS666 (Bpro_19110)
		CN1/20329	CN1contig14666	Rieske alpha	Phthalate 4,5-dioxygenase	72% *Polaromonas* sp. JS666 (Bpro_19110)
CN1-65	*Betaproteobacteria* (genus unclear)	CN1/6611	CN1contig04340	Rieske ferredoxin	XylT ferredoxin	58% *Alicycliphilus denitrificans* BC (Alide0329)
CN1-66	*Betaproteobacteria* - ORF too short for assigning phylogeny (homology for the 67 AA long fragment <50%)	CN1/11990	CN1contig08096	Other protein families (reductase)	Maleylacetate reductase	49% *Burkholderia pseudomallei* MSHR346 (YP_004330781.1)
CN1-67	*Betaproteobacteria (Achromobacter)*	CN1/24000	CN1contig17568	Other protein families (Glutathione S transferase)	Glutathione S transferase	100% *Ralstonia* sp. U2 (AAD12618)
CN1-68	*Betaproteobacteria* - homology lower than 50% for assigning phylogeny	CN1/22362	CN1contig16248	Rieske ferredoxin	Function unknown	49% *Burkholderia pseudomallei* MSHR346 (YP_002895688.1)
		CN1/20130	CN1contig14531	Rieske alpha	Function unknown	
CN1-69	*Betaproteobacteria (Achromobacter)*	CN1/2287	CN1contig01449	Flavoprotein monooxygenase	Function unknown	68% *Burkholderia xenovorans* LB400 (Bxe_C0213)
CN1-70	*Betaproteobacteria (Comamonas)*	CN1/9936	CN1contig06646	INTRA	Function unknown	99% *Comamonas testosteroni* KF-1 (ZP_03541192)
CN1-71	Homology lower than 51% for assigning phylogeny	CN1/10873	CN1contig07301	INTRA	Function unknown	51% *Leadbetterella byssophila* DSM 17132 (YP_003997051)
CN1-72	*Betaproteobacteria (Achromobacter)*	CN1/11790	CN1contig07951	Rieske alpha	Function unknown	99% *Achromobacter piechaudii* (ZP_06686669)
CN1-73	ORF too short for assigning phylogeny (homology for the 107 AA long fragment 55%)	CN1/14160	CN1contig09716	INTRA	Function unknown	55% *Planctomycesmaris* DSM 8797 (ZP_01852244)
CN1-74	*Betaproteobacteria* (genus unclear)	CN1/15436	CN1contig10674	Rieske alpha	Function unknown	69% *Burkholderia* sp. Ch1-1 (ZP_06845959)
CN1-75	*Bacteroidetes*	CN1/17365	CN1contig12083	INTRA	Function unknown	60% *Mucilaginibacter paludis* (ZP_07746550)
CN1-76	*Betaproteobacteria*	CN1/22362	CN1contig16248	Rieske ferredoxin	Function unknown	59% *Burkholderia* sp. Ch1-1 (ZP_09823127.1)

	(Achromobacter)					
CN1-77	*Betaproteobacteria (Comamonas)*	CN1/25534	CN1contig18962	EXDO	Function unknown	100% *Comamonas testosteroni* KF-1 (ZP_03546056)
CN1-78	*Alphaproteobacteria (Azospirillum)*	CN1/0553	CN1contig00338	Response regulator	Response regulator (LysR)	89% *Azospirillum* B510 (Azl004580)
CN1-79	*Betaproteobacteria (Achromobacter)*	CN1/1836	CN1contig01178	EXDO	Probably regulator (LysR)	56% *Polaromonas* sp. JS666 (Bpro_2710)
CN1-80	*Betaproteobacteria (Achromobacter)*	CN1/2307	CN1contig01467	EXDO	Probably regulator (LysR)	99% *Achromobacter piechaudii* (ZP_06690047)
CN1-81	*Betaproteobacteria (Achromobacter)*	CN1/2375	CN1contig01511	EXDO???	Probably regulator (LysR)	74% *Achromobacter xylosooxidans* (AXX-A EGP47715)
CN1-82	*Gammaproteobacteria (Pseudoxanthomonas)*	CN1/2582	CN1contig01649	EXDO???	Probably regulator (LysR)	67% *Pseudoxanthomonas suwonensis* 11-1 (Psesu1673)
CN1-83	*Betaproteobacteria (Comamonas)*	CN1/9644	CN1contig06457	LigB	Gallate dioxygenase	100% *Comamonas testosteroni* KF-1 (ZP_03542382)

[1] We might have overestimated the number of proteins in our metaproteomes by double counting of the fragmented open reading frames in the meta-sequences that were derived from the same gene/protein. In addition, some of the proteins (i.e. dioxygenases) herein investigated are formed by different subunits, that all together form a single protein. To take both issues into consideration the amino acid sequences of proteins with biochemical functions shown to be potentially involved in biodegradation, as suggested by in-house database-BLAST search, were manually BLAST and aligned to the best protein hit in the database to ensure identification of fragments most likely belonging to the same sequence/protein (see first column; "list of unique proteins"). As example, the protein named CN1-1 (see first column; "list of unique proteins") represents a single naphthalene dioxygenase that contain most likely alpha, beta and ferredoxin subunits of fragments CN1/10033, CN1/10034, CN1/21052 and CN1/21053 located in different short contigs. The protein ID (ORF) and contig where the gene was found within the pyro-sequences are specifically shown. BLAST alignment was performed between proteins contributing to the major aromatics aerobic degradation pathways found in the different metagenomes. Thus, proteins sharing the same "best hit" and amino acid sequence homology, were specifically indicated; as example, the protein named CN1-1 (see first column; "list of unique proteins") represent a single naphthalene dioxygenase most similar to CN2-1 naphthalene dioxygenase.

[2] For taxonomic assignment complete genome sequences were obtained from NCBI and used to map the contigs containing the gene coding the protein(s) with biochemical functions shown to be potentially involved in biodegradation.

[3] Functional assignment of the predicted genes involved in degradation was made on the basis of BLASTP analysis against an in-house database that contained gene sequences encoding enzymes that usually perform key catalysing steps in the aerobic degradation of pollutants via di- and trihydroxylated intermediates (**Table S8**). These enzymes included the following FAMILIES: Rieske non-heme iron oxygenases, type I extradiol dioxygenases of the vicinal oxygen chelate superfamily, the type II or LigB superfamily extradiol dioxygenases, and the type III extradiol dioxygenases belonging to the cupin superfamily as well as intradiol dioxygenases. SUBFAMILIES refer to the specific subfamily within FAMILIES assigned to the protein of interest.

[4] Accession number for the best in-house database/NCBI BLAST hit is specifically shown.

(B) CN2 enzymes features

List of unique proteins[1]	Phylogeny[2]	ORF[1]	Contig[1]	FAMILY[3]	SUBFAMILY[3]	BEST HITS[4]
CN2-1 = CN1-1	*Betaproteobacteria (Achromobacter)*	CN2/8665	CN2contig04944	Rieske ferredoxin	Naphthalene dioxygenase/salicylate 5-hydroxylase	94% NagAc *Ralstonia* sp. U2 (AAD12609)
		CN2/8666	CN2contig04944	Rieske alpha	Naphthalene dioxygenase (*Proteobacteria*-like)	95% NagAc *Ralstonia* sp. U2 (AAD12610)
		CN2/15793	CN2contig08915	Rieske beta	Naphthalene dioxygenase	96% NagAc *Ralstonia* sp. U2 (AAD12610)
		CN2/15794	CN2contig08915	Rieske alpha	Naphthalene dioxygenase (*Proteobacteria*-like)	97% NagAc *Ralstonia* sp. U2 (AAD12610)
CN2-2	*Gammaproteobacteria (Pseudomonas)*	CN2/11945	CN2contig07084	Rieske beta	Naphthalene dioxygenase (*Proteobacteria*-like)	99% NahAd *Pseudomonas* sp. ND6 (NP_943189) and 78% *Ralstonia* sp. U2 (AAD12611)
		CN2/11946	CN2contig07084	Rieske alpha	Naphthalene dioxygenase (*Proteobacteria*-like)	99% NahAc *Pseudomonas* sp. ND6 (NP_943188) and 89% *Ralstonia* sp. U2 (AAD12610)
		CN2/11947	CN2contig07084	Rieske ferredoxin	Naphthalene dioxygenase (*Proteobacteria*-like)	100% NahAb *Pseudomonas* sp. ND6 (NP_943187) and 80% *Ralstonia* sp. U2 (AAD12609)
		CN2/11948	CN2contig07084	Rieske reductase	Naphthalene dioxygenase (*Proteobacteria*-like)	NahAa 99% *Pseudomonas* sp. ND6 (NP_943186) and 67% *Ralstonia* sp. U2 (AAD12606)
CN2-3 = CN1-8	*Betaproteobacteria (Achromobacter)*	CN2/8417	CN2contig04760	EXDO	Dihydroxynaphthalene dioxygenase	93% *Ralstonia* sp. U2 (AAD12614)
CN2-4	*Gammaproteobacteria (Pseudomonas)*	CN2/11942	CN2contig07084	EXDO	Dihydroxynaphthalene dioxygenase	100% NahC *Pseudomonas* sp. ND6 (NP_943192) and 88% *Ralstonia* sp. U2 (AAD12614)
CN2-5	*Gammaproteobacteria (Pseudomonas)*	CN2/11944	CN2contig07084	Dihydrodiol dehydrogenase	Naphthalene dihydrodiol dehydrogenase	99% NahB *Pseudomonas* sp. ND6 (NP_943190) and 85% *Ralstonia* sp. U2 (AAD12612)
CN2-6	*Betaproteobacteria (Achromobacter)*	CN2/17233	CN2contig09709	Dihydrodiol dehydrogenase	Naphthalene dihydrodiol dehydrogenase	94% *Ralstonia* sp. U2 (AAD12612)
CN2-7	*Gammaproteobacteria (Pseudomonas)*	CN2/11939	CN2contig07084	Other protein families (isomerase)	2-Hydroxychromene-2-carboxylate isomerase	100% NahD *Pseudomonas* sp. ND6 (NP_943093) and 70% *Ralstonia* sp. U2 (AAD12617)
CN2-8	*Gammaproteobacteria (Pseudomonas)*	CN2/11940	CN2contig07084	Other protein families (aldolase)	trans-o-Hydroxybenzylidenepyruvate hydratase-aldolase	100% NahE *Pseudomonas* sp. ND6 (NP_943093) and 89% *Ralstonia* sp. U2 (AAD12616)
CN2-9	*Betaproteobacteria* - ORF too short (120 AA) for assigning phylogeny	CN2/6956	CN2contig03788	Salicylaldehyde dehydrogenase	Salicylaldehyde dehydrogenase	82% *Leptothrix cholodnii* SP-6 (Lcho0804; NC_010524)
CN2-10	*Gammaproteobacteria (Pseudomonas)*	CN2/11943	CN2contig07084	Other protein families (dehydrogenase)	Salicylaldehyde dehydrogenase	100% NahF *Pseudomonas* sp. ND6 (NP_943191) and 89% *Ralstonia* sp. U2 (AAD12613)
CN2-11 = CN1-15	*Betaproteobacteria (Achromobacter)*	CN2/0419	CN2contig00156	Rieske ferredoxin	Salicylate 5-hydroxylase	94% *Achromobacter piechaudii* (ZP_06686007)
		CN2/0420	CN2contig00156	Rieske beta	Salicylate 5-hydroxylase	96% *Achromobacter piechaudii* (ZP_06686008)
		CN2/0421	CN2contig00156	Rieske alpha	Salicylate 5-hydroxylase	98% *Achromobacter piechaudii* (ZP_06686009)
CN2-12	*Betaproteobacteria (Achromobacter)*	CN2/3310	CN2contig01635	Rieske alpha	Function unknown (distantly related to salicylate 5-hydroxylase)	96% *Achromobacter piechaudii* (ZP_06688288)
CN2-13 = CN1-	*Betaproteobacteria*	CN2/3755	CN2contig01872	Rieske beta	Salicylate 5-hydroxylase	71% *Polaromonas* sp. JS666 (Bpro0984)

16	*(Achromobacter)*	CN2/3756	CN2contig01872	Rieske alpha	Salicylate 5-hydroxylase	87% *Polaromonas* sp. JS666 (Bpro0983)
		CN2/6954	CN2contig03788	Rieske beta	Salicylate 5-hydroxylase	86% *Polaromonas* sp. JS666 (Bpro0984)
		CN2/6955	CN2contig03788	Rieske ferredoxin	Salicylate 5-hydroxylase	71% *Polaromonas* sp. JS666 (Bpro0985)
		CN2/7750	CN2contig04287	Rieske alpha	Salicylate 5-hydroxylase	88% *Polaromonas* sp. JS666 (Bpro0983)
CN2-14	*Gammaproteobacteria (Pseudomonas)*	CN2/16709	CN2contig09412	Rieske ferredoxin	Salicylate 5-hydroxylase	100% *Pseudomonas putida* AK5 (ACO92379)
		CN2/16710	CN2contig09412	Rieske beta	Salicylate 5-hydroxylase	100% *Pseudomonas putida* AK5 (ACO92378)
		CN2/16711	CN2contig09412	Rieske alpha	Salicylate 5-hydroxylase	100% *Pseudomonas putida* AK5 (ACO92377)
		CN2/16714	CN2contig09412	Rieske reductase	Salicylate 5-hydroxylase	100% *Pseudomonas putida* AK5 (ACO92374)
CN2-15	*Betaproteobacteria (Achromobacter)*	CN2/2008	CN2contig00900	Rieske alpha	Function unknown (related to salicylate 5-hydroxylase)	100% *Achromobacter piechaudii* (ZP_06687854)
		CN2/2009	CN2contig00900	Rieske ferredoxin	Function unknown (related to salicylate 5-hydroxylase)	94% *Achromobacter piechaudii* (ZP_06687853)
CN2-16	*Betaproteobacteria (Achromobacter)*	CN2/0906	CN2contig00386	Cupin	Gentisate dioxygenase	99% *Achromobacter piechaudii* (ZP_06685546)
CN2-17 = CN1-6	*Betaproteobacteria (Achromobacter)*	CN2/3390	CN2contig01687	Cupin	Gentisate dioxygenase	100% *Ralstonia* sp. U2 (AAD12619)
CN2-18 = CN1-5	*Betaproteobacteria* (genus unclear)	CN2/6053	CN2contig03225	Cupin	Gentisate dioxygenase	80% *Comamonas testosteroni* KF-2 (ZP_03542624)
		CN2/6059	CN2contig03229	Cupin	Gentisate dioxygenase	84% *Comamonas testosteroni* KF-2 (ZP_03542624)
CN2-19 = CN1-4	*Betaproteobacteria (Achromobacter)*	CN2/8906	CN2contig05108	Cupin	Gentisate dioxygenase	80% *Achromobacter piechaudii* (ZP_06684540)
		CN2/8907	CN2contig05108	Cupin	Gentisate dioxygenase	96% *Achromobacter piechaudii* (ZP_06684540)
CN2-20	*Gammaproteobacteria (Pseudomonas)*	CN2/16713	CN2contig09412	Cupin	Gentisate dioxygenase	100% *Pseudomonas putida* AK5 (ACO92375)
CN2-21	*Betaproteobacteria* (genus unclear)	CN2/3389	CN2contig01687	Other protein families (hydrolase)	Fumarylpyruvate hydrolase	82% *Leptothrix cholodnii* SP-6 Lcho3671 (ACB35925)
CN2-22	*Gammaproteobacteria (Pseudomonas)*	CN2/16708	CN2contig09412	Other protein families (hydrolase)	Fumarylpyruvate hydrolase	100% *Pseudomonas putida* AK5 (ACO92383)
CN2-23	*Gammaproteobacteria (Pseudomonas)*	CN2/16712	CN2contig09412	Other protein families (hydrolase)	Fumarylpyruvate hydrolase	100% *Pseudomonas putida* AK5 (ACO92376)
CN2-24	*Gammaproteobacteria (Pseudomonas)*	CN2/0050	CN2contig00016	Flavoprotein monooxygenase	Salicylate 1-hydroxylase	100% NahG *Pseudomonas* sp. ND6 (AAP4222)
		CN2/2365	CN2contig01098	Flavoprotein monooxygenase	Salicylate 1-hydroxylase	100% NahG *Pseudomonas* sp. ND6 (AAP4222)
		CN2/3285	CN2contig01620	Flavoprotein monooxygenase	Salicylate 1-hydroxylase	100% NahG *Pseudomonas* sp. ND6 (AAP4222)
		CN2/17374	CN2contig09758	Flavoprotein monooxygenase	Salicylate 1-hydroxylase	100% nahG *Pseudomonas* sp. ND6 (AAP44222)
		CN2/3796	CN2contig01901	Flavoprotein monooxygenase	Salicylate 1-hydroxylase	59% NahG *Pseudomonas* sp. ND6 (AAP44222)
CN2-25	*Gammaproteobacteria (Pseudomonas)*	CN2/16920	CN2contig09541	Flavoprotein monooxygenase	Salicylate 1-hydroxylase	100% NahW *Pseudomonas stutzeri* (AAD02157)
CN2-26	*Gammaproteobacteria (Pseudomonas)*	CN2/4832	CN2contig02517	EXDO	2,3-Dihydroxybiphenyl dioxygenase	84% PcbC *Pseudomonas* sp. DJ-12 (BAA07956)
CN2-27	*Betaproteobacteria (Achromobacter)*	CN2/7789	CN2contig04321	EXDO	2,3-Dihydroxybiphenyl dioxygenase	77% *Achromobacter xylosooxidans* A8 (pA81_033)
		CN2/13412	CN2contig07862	EXDO	2,3-Dihydroxybiphenyl	64% *Achromobacter xylosooxidans* A8 (pA81_033)

					dioxygenase	
CN2-28 = CN1-25	*Betaproteobacteria (Achromobacter)*	CN2/2038	CN2contig00922	Rieske alpha	Function unknown (distantly related to benzoate dioxygenase)	98% *Achromobacter piechaudii* (ZP_06685665)
CN2-29 = CN1-24	*Betaproteobacteria (Achromobacter)*	CN2/15271	CN2contig08673	Rieske alpha	Function unknown (distantly related to benzoate dioxygenase)	97% *Achromobacter piechaudii* (ZP_06686030)
		CN2/15272	CN2contig08673	Rieske beta	Function unknown (distantly related to benzoate dioxygenase)	98% *Achromobacter piechaudii* (ZP_06686029)
CN2-30	*Betaproteobacteria* (genus unclear)	CN2/6417	CN2contig03445	Rieske alpha	Benzoate dioxygenase	85% *Bulkhoderia multivorans* ATCC17616 Bmul3238 (ZP_03583404)
CN2-31	*Gammaproteobacteria (Pseudomonas)*	CN2/17820	CN2contig09876	Rieske ferredoxin	Benzoate dioxygenase	96% BenC *Pseudomonas stutzeri* A1501 (PST1669)
		CN2/17821	CN2contig09876	Rieske beta	Benzoate dioxygenase	96% BenB *Pseudomonas stutzeri* A1501 (PST1668)
		CN2/17822	CN2contig09876	Rieske alpha	Benzoate dioxygenase	97% BenA *Pseudomonas stutzeri* A1501 (PST1667)
CN2-32	*Gammaproteobacteria (Pseudomonas)*	CN2/13057	CN"contig07678	Dihydrodiol dehydrogenase	Benzoate dihydrodiol dehydrogenase	96% BenD *Pseudomonas stutzeri* A1501 (PST1670)
		CN2/17819	CN2contig09876	Dihydrodiol dehydrogenase	Benzoate dihydrodiol dehydrogenase	88% BenD *Pseudomonas stutzeri* A1501 (PST1670)
CN2-33	*Gammaproteobacteria (Pseudomonas)*	CN2/2673	CN2contig01289	Other protein families	Phenol hydroxylase subunit	86% LapP *Pseudomonas alkylphenolia* (AAP92394)
		CN2/2674	CN2contig01289	Other protein families	Phenol hydroxylase subunit	75% LapO *Pseudomonas alkylphenolia* (AAP92393)
		CN2/2675	CN2contig01289	Other protein families	Phenol hydroxylase subunit	86% LapN *Pseudomonas alkylphenolia* (AAP92392.1)
		CN2/2676	CN2contig01289	Other protein families	Phenol hydroxylase subunit	94% LapM *Pseudomonas alkylphenolia* (AAP92391)
		CN2/2677	CN2contig01289	Other protein families	Phenol hydroxylase subunit	81% LapL *Pseudomonas alkylphenolia* (AAP92390)
		CN2/2678	CN2contig01289	Other protein families	Phenol hydroxylase subunit	69% LapK *Pseudomonas al*kylphenolia (AAP92389)
CN2-34	*Betaproteobacteria (Achromobacter)*	CN2/9010	CN2contig05176	Flavoprotein monooxygenase	3-Hydroxybenzoate 6-hydroxylase	93% *Achromobacter piechaudii* (ZP_06684546)
CN2-35	*Gammaproteobacteria (Pseudomonas)*	CN2/5182	CN2contig02698	INTRA	Catechol 1,2-dioxygenase	79% *Pseudomonas denitrificans* (Pden1176)
CN2-36	*Betaproteobacteria* (genus unclear)	CN2/6108	CN2contig03261	INTRA	Catechol 1,2-dioxygenase	88% *Alicycliphilus denitrificans* (Alide_2650)
CN2-37	*Gammaproteobacteria (Pseudomonas)*	CN2/13061	CN2contig07678	INTRA	Catechol 1,2-dioxygenase	93% CatA *Pseudomonas stutzeri* A1501 (PST1674)
CN2-38	*Gammaproteobacteria (Pseudomonas)*	CN2/17371	CN2contig09758	EXDO	Catechol 2,3-dioxygenase (*Pseudomonads*-like)	100% NahH *Pseudomonas* sp. ND6 (AAP44220)
		CN2/18217	CN2contig10059	EXDO	Catechol 2,3-dioxygenase (*Pseudomonads*-like)	94% *Pseudomonas* sp. ND6 (AAP44220)
CN2-39	*Gammaproteobacteria (Pseudomonas)*	CN2/2679	CN2contig01289	EXDO	Catechol 2,3-dioxygenase	85% LapB *Pseudomonas* sp. Kl28 (3HPV_A)
CN2-40	*Betaproteobacteria* (genus unclear)	CN2/5184	CN2contig02698	Other protein families (isomerase)	Muconate cycloisomerase	78% *Bulkholderia multivorans* ATCC17616 (Bmul3235)
CN2-41	*Betaproteobacteria* (genus unclear)	CN2/6107	CN2contig03261	Other protein families (isomerase)	Muconate cycloisomerase	86% *Alicycliphilus denitrificans* (Alide_2651)
CN2-42	*Gammaproteobacteria (Pseudomonas)*	CN2/13059	CN2contig07678	Other protein families (isomerase)	Muconate cycloisomerase	95% CatB *Pseudomonas stutzeri* A1501 (PST1672)
CN2-43	*Gammaproteobacteria*	CN2/13060	CN2contig07678	Other protein families	Muconolactone isomerase	96% CatC *Pseudomonas stutzeri* A1501 (PST1673)

	(Pseudomonas)			(isomerase)		
CN2-44	*Betaproteobacteria* (genus unclear)	CN2/5183	CN2contig02698	Other protein families (isomerase)	Muconolactone isomerase	85% *Bulkholderia multivorans* ATCC17616 (Bmul3236)
CN2-45	*Gammaproteobacteria (Pseudomonas)*	CN2/2671	CN2contig01289	Other protein families (dehydrogenase)	2-Hydroxymuconic acid semialdehyde dehydrogenase	91% LapC *Pseudomonas* sp. Kl28
CN2-46	*Betaproteobacteria* (genus unclear)	CN2/5469	CN2contig02873	Other protein families (dehydrogenase)	2-Hydroxymuconic acid semialdehyde dehydrogenase	89% *Comamonas testosteroni* KF-1 (ZP_03546055)
CN2-47	*Gammaproteobacteria (Pseudomonas)*	CN2/17370	CN2contig09758	Other protein families (dehydrogenase)	2-Hydroxymuconic acid semialdehyde dehydrogenase	100% *Pseudomonas* sp. ND6 nahI (AAP44219)
		CN2/18216	CN2contig10059	Other protein families (dehydrogenase)	2-Hydroxymuconic acid semialdehyde dehydrogenase	96% *Pseudomonas* sp. ND6 nahI (AAP44219)
CN2-48	*Gammaproteobacteria (Pseudomonas)*	CN2/17369	CN2contig09758	Other protein families (hydrolase)	2-Hydroxymuconic acid semialdehyde hydrolase	100% *Pseudomonas* sp. ND6 nahN (AAP44218)
CN2-49	*Gammaproteobacteria (Pseudomonas)*	CN2/8952	CN2contig05137	Other protein families (lyase)	2-Oxopent-4-enoate hydratase	100% NahL *Pseudomonas putida* GJ31 (AAX38568)
CN2-50	*Gammaproteobacteria (Pseudomonas)*	CN2/2669	CN2contig01289	Other protein families (tautomerase)	Oxalocrotonate tautomerase	74% LapI *Pseudomonas* sp. Kl28
CN2-51	*Gammaproteobacteria (Pseudomonas)*	CN2/17365	CN2contig09758	Other protein families (tautomerase)	4-Oxalocrotonate tautomerase	100% NahJ *Pseudomonas* sp. ND6 (AAP44213)
CN2-52	*Gammaproteobacteria (Pseudomonas)*	CN2/17367	CN2contig09758	Other protein families (aldolase)	4-Hydroxy-2-oxovalerate aldolase	100% NahM *Pseudomonas* sp. ND6 (AAP44215)
CN2-53	*Gammaproteobacteria (Pseudomonas)*	CN2/2667	CN2contig01289	Other protein families (aldolase)	4-Hydroxy-2-oxovalerate aldolase	91% LapG *Pseudomonas alkylphenolia* KL28
CN2-54	*Gammaproteobacteria (Pseudomonas)*	CN2/2670	CN2contig01289	Other protein families (lyase)	2-Hydroxypent-2,4-dienoate hydratase/4-oxalocrotonate decarboxylase chimer	92% LapE (1-263) and 88% LapH (266-end) *Pseudomonas alkylphenolia* KL28
CN2-55	*Gammaproteobacteria (Pseudomonas)*	CN2/2668	CN2contig01289	Other protein families (dehydrogenase)	Acetaldehyde dehydrogenase	LapF 90% *Pseudomonas alkylphenolia* KL28
CN2-56	*Gammaproteobacteria (Pseudomonas)*	CN2/17368	CN2contig09758	Other protein families (dehydrogenase)	Acetaldehyde dehydrogenase	100% NahO *Pseudomonas* sp. ND6 (AAP44216)
CN2-57 = CN1-40	*Gammaproteobacteria* (genus unclear)	CN2/1892	CN2contig00828	Other protein families (non-heme dioxygenase)	4-Hydroxyphenylpyruvate dioxygenase	100% *Stenotrophomonas maltophilia* R551-3 (Smal_3739)
		CN2/4701	CN2contig02432	Other protein families (non-heme dioxygenase)	4-Hydroxyphenylpyruvate dioxygenase	96% *Stenotrophomonas maltophilia* R551-3 (Smal_3739)
CN2-58	*Gammaproteobacteria (Pseudomonas)*	CN2/3407	CN2contig01694	Other protein families (non-heme dioxygenase)	4-Hydroxyphenylpyruvate dioxygenase	94% *Pseudomonas stutzeri* A1501 (PST0869)
CN2-59	*Gammaproteobacteria (Pseudomonas)*	CN2/5593	CN2contig02955	Other protein families (non-heme dioxygenase)	4-Hydroxyphenylpyruvate dioxygenase	68% *Pseudomonas stutzeri* A1501 (PST_0200)
CN2-60	*Gammaproteobacteria* (genus unclear)	CN2/6542	CN2contig03521	Other protein families (non-heme dioxygenase)	4-Hydroxyphenylpyruvate dioxygenase	84*% Pseudomonas fluorescens* SBW25 PFL (YP_002872577)
CN2-61	*Betaproteobacteria (Achromobacter)*	CN2/13419	CN2contig07868	Other protein families (non-heme dioxygenase)	4-Hydroxyphenylpyruvate dioxygenase	100% *Achromobacter piechaudii* (ZP_06690133)
CN2-62	*Gammaproteobacteria* (genus unclear)	CN2/5775	CN2contig03059	Cupin	Homogentisate dioxygenase	77% *Pseudomonas putida* W619i (ACA71322) (PputW619_0817)
CN2-63	*Betaproteobacteria (Achromobacter)*	CN2/8307	CN2contig04687	Cupin	Homogentisate dioxygenase	99% *Achromobacter piechaudii* (ZP_06689958)
		CN2/12224	CN2contig07248	Cupin	Homogentisate dioxygenase	89% *Achromobacter piechaudii* (ZP_06689958)
CN2-64	*Gammaproteobacteria* (genus unclear)	CN2/1891	CN2contig00828	Cupin	Homogentisate dioxygenase	97% *Stenotrophomonas maltophilia* R551-3 (Smal_3738)
CN2-65	*Gammaproteobacteria*	CN2/5776	CN2contig03059	Other protein families	Fumarylacetoacetate hydrolase	70% *Pseudomonas putida* W619 (PputW619_0818)

	(genus unclear)			(hydrolase)		
CN2-66	*Betaproteobacteria (Achromobacter)*	CN2/12223	CN2contig07248	Other protein families (hydrolase)	Fumarylacetoacetate hydrolase	76% *Achromobacter piechaudii* (ZP_06689959)
CN2-67	*Betaproteobacteria* (genus unclear)	CN2/9561	CN2contig05551	Other protein families (lyase)	Benzylformate decarboxylase	59% *Bulkholderia phymatum* STM815 (YP_001862255)
CN2-68	*Gammaproteobacteria* (genus unclear)	CN2/5171	CN2contig02690	Flavoprotein monooxygenase	4-Hydroxybenzoate 3-hydroxylase	88% *Pseudomonas putida* W619 (W619_1946)
CN2-69	*Betaproteobacteria (Achromobacter)*	CN2/9560	CN2contig05551	Flavoprotein monooxygenase	4-Hydroxybenzoate 3-hydroxylase	75% *Achromobacter piechaudii* (ZP_06689206)
CN2-70 = CN1-32	*Betaproteobacteria (Achromobacter)*	CN2/2431	CN2contig01140	Rieske alpha	Vanillate monooxygenase	88% *Achromobacter xylosooxidans* A8 (AXYL_02987)
CN2-71	*Betaproteobacteria* (genus unclear)	CN2/6302	CN2contig03379	Rieske alpha	Vanillate monooxygenase	75% *Cupriavidus necator* JMP134 ReutB5370 (YP_297772)
CN2-72	*Betaproteobacteria (Achromobacter)*	CN2/6026	CN2contig03209	INTRA	Protocatechuate 3,4-dioxygenase beta subunit	87% *Achromobacter piechaudii* (ZP_06689204)
CN2-73	*Betaproteobacteria* (genus unclear)	CN2/5000	CN2contig02616	LigB	Homoprotocatechuate dioxygenase	89% *Burkholderia bronchiseptica* RB50 (NP_887287)
CN2-74 = CN1-50	*Betaproteobacteria (Achromobacter)*	CN2/8005	CN2contig04472	LigB	Homoprotocatechuate dioxygenase	93% *Achromobacter piechaudii* (ZP_06690164)
CN2-75	*Betaproteobacteria (Achromobacter)*	CN2/8006	CN2contig04472	Other protein families (dehydrogenase)	Carboxymethylmuconic semiladehyde dehydrogenase	99% *Achromobacter piechaudii* (ZP_06690163)
CN2-76	*Betaproteobacteria* (genus unclear)	CN2/5001	CN2contig02616	Other protein families (dehydrogenase)	Carboxymethylmuconic semiladehyde dehydrogenase	96% *Burkholderia bronchiseptica* RB50 (NP_887286)
CN2-77	*Betaproteobacteria (Achromobacter)*	CN2/2410	CN2contig01133	Rieske alpha	Ibuprofen-CoA dioxygenase (related to putative *Sphingomonas* sp. Ibu-2)	97% *Achromobacter piechaudii* (ZP_06686030)
		CN2/2409	CN2contig01133	Rieske reductase	Ibuprofen-CoA dioxygenase (related to putative *Sphingomonas* sp. Ibu-2)	92% *Achromobacter piechaudii* (ZP_06686031)
CN2-78 = CN1-55	*Betaproteobacteria* (genus unclear)	CN2/2603	CN2contig01250	LigB	2-Aminophenol 1,6-dioxygenase	88% *Burkholderia xenovorans* LB400 (Bxe_A1145)
CN2-79	*Betaproteobacteria (Achromobacter)*	CN2/5958	CN2contig03167	Flavoprotein monooxygenase	6-Hydroxynicotinate 3-monooxygenase	97% *Achromobacter piechaudii* (ZP_06684483)
CN2-80	*Gammaproteobacteria (Pseudomonas)*	CN2/17372	CN2contig09758	Rieske ferredoxin	XylT ferredoxin	100% NahT *Pseudomonas* sp. ND6 AAP44221 (NP_863104)
CN2-81	*Gammaproteobacteria (Pseudomonas)*	CN2/16715	CN2contig09412	Other protein families (regulator)	NahR transcriptional regulator	100% *Pseudomonas putida* AK5 (ACO92380)
CN2-82	*Gammaproteobacteria (Pseudomonas)*	CN2/2680	CN2contig01289	Other protein families (regulator)	LapR regulator	80% LapR *Pseudomonas alkylphenolia* KL28 (AAP92387)
CN2-83	*Gammaproteobacteria (Pseudomonas)*	CN2/5774	CN2contig03059	Other protein families (regulator)	LclR regulator	63% IclR *Pseudomonas syringae* pv.tabaci ATCC 11528 (ZP_05638791)
CN2-84	*Gammaproteobacteria (Pseudomonas)*	CN2/17823	CN2contig09876	Other protein families (regulator)	Regulator	93% BenR *Pseudomonas stutzeri* A1501 (PST1669)
CN2-85	*Betaproteobacteria (Achromobacter)*	CN2/2010	CN2contig00900	Flavoprotein monooxygenase	Function unknown	77% *Achromobacter piechaudii* (ZP_06687852)
CN2-86	*Betaproteobacteria (Achromobacter)*	CN2/2213	CN2contig01025	Rieske alpha	Function unknown	99% *Achromobacter piechaudii* (ZP_06686669)
CN2-87	*Betaproteobacteria (Achromobacter)*	CN2/2866	CN2contig01398	Rieske alpha	Function unknown	96% *Achromobacter piechaudii* (ZP_06688288)
		CN2/2867	CN2contig01398	Rieske beta	Function unknown	83% *Achromobacter piechaudii* (ZP_06688289)

CN2-88	*Betaproteobacteria* (genus unclear)	CN2/4017	CN2contig02029	Rieske beta	Function unknown	71% *Burkholderia bronchiseptica* RB50 (NP_887280)
		CN2/1050	CN2contig00452	Rieske alpha	Function unknown	92% *Burkholderia bronchiseptica* RB50 (NP_887278)
CN2-89	*Betaproteobacteria (Achromobacter)*	CN2/10832	CN2contig06384	Rieske alpha	Function unknown	78% *Achromobacter piechaudii* (ZP_06688288)
CN2-90	*Betaproteobacteria* (genus unclear)	CN2/1110	CN2contig00485	EXDO	Function unknown (LysR regulator)	54% *Burkholderia* sp. H160 (ZP_03269777)
CN2-91	*Betaproteobacteria* (genus unclear)	CN2/16975	CN2contig09603	Rieske alpha	Function unknown	71% *Comamonas testosteroni* CNB-2 (YP_003279568)
CN2-92	*Betaproteobacteria* (genus unclear)	CN2/1808	CN2contig00789	Rieske alpha	Function unknown	63% *Burkholderia xenovorans* LB400 Bxe_B1578 (ABE34386)
CN2-93	*Gammaproteobacteria (Pseudomonas)*	CN2/2672	CN2contig01289	Other protein families	Glutathione dependent enzyme	42% *Novosphingobium aromaticivorans* DSM 12444 (YP_497099)
CN2-94	*Gammaproteobacteria (Pseudomonas)*	CN2/2681	CN2contig01289	Other protein families	Putative methyl-accepting protein	67% *Pseudomonas alkylphenolia* KL28 (AAP92386)
CN2-95	*Gammaproteobacteria (Pseudomonas)*	CN2/2682	CN2contig01289	Other protein families (transporter)	Transporter	67% *Pseudomonas fluorescens* WH6 (ZP_07774655)
CN2-96	*Gammaproteobacteria (Pseudomonas)*	CN2/11941	CN2contig07084	Other protein families	Membrane protein	98% NahQ *Pseudomonas* sp. ND6 (AAP44191) and 76% *Ralstonia* sp. U2 (AAD12615)
CN2-97	*Gammaproteobacteria (Pseudomonas)*	CN2/13058	CN"contig07678	Other protein families (transporter)	Transport protein	90% BenK *Pseudomonas stutzeri* A1501 (PST1671)
CN2-98	*Gammaproteobacteria (Pseudomonas)*	CN2/13062	CN2contig07678	Other protein families (transporter)	Transport protein	90% BenE *Pseudomonas stutzeri* A1501 (PST1675)
CN2-99	*Gammaproteobacteria (Pseudomonas)*	CN2/13063	CN2contig07678	Other protein families (transporter)	Porin	75% BenD *Pseudomonas stutzeri* A1501 (PST1676)
CN2-100	*Gammaproteobacteria (Pseudomonas)*	CN2/11949	CN2contig07084	Other protein families (ligase)	Ligase	100% *Pseudomonas* sp. ND6 NP_943185.1 (not present in *Ralstonia* sp. U2)
CN2-101	*Gammaproteobacteria* (genus unclear)	CN2/17363	CN2contig09758	Other protein families (transposase)	Transposase	80% *Pseudomonas* sp. K23 (BAC56721.1)
CN2-102	*Gammaproteobacteria (Pseudomonas)*	CN2/17364	CN2contig09758	Other protein families	Chemotaxis	100% NahY *Pseudomonas* sp. ND6 (AAP44212)
CN2-13 = CN1-16	*Betaproteobacteria* (genus unclear)	CN2/3754	CN2contig01872	Rieske ferredoxin	Salicylate 5-hydroxylase	72% *Ralstonia eutropha* JMP134 ReutB4721 (YP_298913)

[1] We might have overestimated the number of proteins in our metaproteomes by double counting of the fragmented open reading frames in the meta-sequences that were derived from the same gene/protein. In addition, some of the proteins (i.e. dioxygenases) herein investigated are formed by different subunits, that all together form a single protein. To take both issues into consideration the amino acid sequences of proteins with biochemical functions shown to be potentially involved in biodegradation, as suggested by in-house database-BLAST search, were manually BLAST and aligned to the best protein hit in the database to ensure identification of fragments most likely belonging to the same sequence/protein (see first column; "list of unique proteins"). As example, the protein named CN2-1 (see first column; "list of unique proteins") represents a single naphthalene dioxygenase that contain most likely alpha, beta and ferredoxin subunits of fragments CN2/8665, CN28666, CN2/15793 and CN2/15794 located in different short contigs. The protein ID (ORF) and contig where the gene was found within the pyro-sequences are specifically shown. BLAST alignment was performed between proteins contributing to the major aromatics aerobic degradation pathways found in the different metagenomes. Thus, proteins sharing the same "best hit" and amino acid sequence homology, were specifically indicated; as example, the protein named CN2-1 (see first column; "list of unique proteins") represent a single naphthalene dioxygenase most similar to CN1-1 naphthalene dioxygenase.

[2] For taxonomic assignment complete genome sequences were obtained from NCBI and used to map the contigs containing the gene coding the protein(s) with biochemical functions shown to be potentially involved in biodegradation.

[3] Functional assignment of the predicted genes involved in degradation was made on the basis of BLASTP analysis against an in-house database that contained gene sequences encoding enzymes that usually perform key catalysing steps in the aerobic degradation of pollutants via di- and trihydroxylated intermediates (**Table S8**). These enzymes included the following FAMILIES: Rieske

non-heme iron oxygenases, type I extradiol dioxygenases of the vicinal oxygen chelate superfamily, the type II or LigB superfamily extradiol dioxygenases, and the type III extradiol dioxygenases belonging to the cupin superfamily as well as intradiol dioxygenases. SUBFAMILIES refer to the specific subfamily within FAMILIES assigned to the protein of interest.

[4] Accession number for the best in-house database/NCBI BLAST hit is specifically shown.

(C) N enzymes features

List of unique proteins[1]	**Phylogeny**[2]	**ORF**[1]	**Contig**[1]	**FAMILY**[3]	**SUBFAMILY**[3]	**BEST HITS**[4]
N-1	*Gammaproteobacteria (Pseudomonas)*	N/2334	Ncontig01812	EXDO	Dihydroxynaphthalene dioxygenase	99% NahC AN10 *Pseudomonas stutzeri* (AAD02140)
		N/3459	Ncontig02736	EXDO	Dihydroxynaphthalene dioxygenase	95% NahC AN10 *Pseudomonas stutzeri* (AAD02140)
N-2	*Gammaproteobacteria (Pseudomonas)*	N/3458	Ncontig02736	Other protein families (dehydrogenase)	Salicylaldehyde dehydrogenase	97% NahF AN10 *Pseudomonas* stutzeri (AAD02139)
N-3	*Gammaproteobacteria (Pseudomonas)*	N/1828	Ncontig01406	Flavoprotein monooxygenase	Salicylate 1-hydroxylase	100% NahG AN10 *Pseudomonas stutzeri* (AAD02416)
N-4	*Gammaproteobacteria (Pseudomonas)*	N/4256	Ncontig03508	Flavoprotein monooxygenase	Salicylate 1-hydroxylase	99% NahW AN10 *Pseudomonas stutzeri* (AAD02157)
N-5	*Gammaproteobacteria (Azotobacter)*	N/0017	Ncontig00015	EXDO	Catechol 2,3-dioxygenase (*Pseudomonads*-like)	91% XylE *Azotobacter vinelandii* DJ (also close to *Pseudomonas*)
N-6	*Gammaproteobacteria (Azotobacter)*	N/0616	Ncontig00459	EXDO	Catechol 2,3-dioxygenase (*Pseudomonads*-like)	99% NahH AN10 *Pseudomonas stutzeri* (AAD02148)
N-7	*Gammaproteobacteria (Pseudomonas)*	N/2199	Ncontig01706	INTRA	Catechol 1,2-dioxygenase	98% *Pseudomonas putida* KT24440 (NP_745846)
N-8	*Betaproteobacteria* - ORF too short (29 AA) for assigning phylogeny	N/4527	Ncontig03752	EXDO	Catechol 1,2-dioxygenase	100% *Pseudoxanthomonas spadix* BD-a59 (YP_004930521.1)
N-9	*Gammaproteobacteria (Pseudomonas)*	N/0290	Ncontig00216	Other protein families	Phenol hydroxylase subunit dmpN	93% Dmpn *Pseudomonas putida* H (CAA56743)
		N/0291	Ncontig00216	Other protein families	Phenol hydroxylase subunit dmpM	97% DmpM *Pseudomonas putida* H (CAA56723)
N-10	*Gammaproteobacteria (Pseudomonas)*	N/2511	Ncontig01963	Rieske alpha	Biphenyl dioxygenase	97% BphAa *Pseudomonas* sp. Cam-1 (AAK14781)
N-11	*Gammaproteobacteria (Pseudomonas)*	N/3972	Ncontig03244	Other protein families (dehydrogenase)	Biphenyl dihydrodiol dehydrogenase	97% BphB *Pseudomonas* sp. Cam-1 (AAK14786)
N.12	*Gammaproteobacteria (Pseudomonas)*	N/3973	Ncontig03244	EXDO	Dihydroxybiphenyl dioxygenase	88% BphC *Pseudomonas* sp.Cam-1 (AAK14786)
N-13	*Betaproteobacteria (Achromobacter)*	N/0414	Ncontig00302	Rieske alpha	Function unknown (distantly related to benzoate dioxygenase)	100% *Pseudomonas resinovorans* (NP_758556)
N-14	*Gammaproteobacteria (Pseudoxanthomonas)*	N/1923	Ncontig01484	Rieske beta	Benzoate dioxygenase	98% *Pseudoxanthomonas spadix* BD-a59 (YP_004931648.1)
		N/1924	Ncontig01484	Rieske alpha	Benzoate dioxygenase	
N-15	*Alphaproteobacteria (Novosphingobium)*	N/4369	Ncontig03607	Rieske beta	Benzoate dioxygenase	86% *Ralstonia pickettii* 12D (Rpic12D1429)
		N/4370	Ncontig03607	Rieske alpha	Benzoate dioxygenase	78% *Ralstonia pickettii* 12D (Rpic12D1428)
N-16	*Gammaproteobacteria*	N/4825	NNcontig04021	Flavoprotein	4-Hydroxybenzoate 3-	92% *Pseudomonas brassicacearum* subsp.

	(Pseudomonas)			monooxygenase	hydroxylase	*brassicacearum* NFM421 (PSEBR_a1192)
N-17	*Gammaproteobacteria (Azotobacter)*	N/0022	Ncontig00020	INTRA	Protocatechuate 3,4-dioxygenase alpha subunit	99% *Pseudomonas putida* NCIMB9869 (AAA25921)
N-18	*Gammaproteobacteria (Pseudomonas)*	N/4174	Ncontig03429	INTRA	Protocatechuate 3,4-dioxygenase beta subunit	92% *Pseudomonas brassicacearum* subsp. *brassicacearum* NFM421 (PSEBR_a1268)
		N/5143	Ncontig04332	INTRA	Protocatechuate 3,4-dioxygenase alpha subunit	92% *Pseudomonas brassicacearum* subsp. *brassicacearum* NFM421 (PSEBR_a1269)
N-19	*Gammaproteobacteria (Pseudomonas)*	N/1561	Ncontig01176	Rieske alpha	Vanillate monooxygenase	99% *Pseudomonas brassicacearum* subsp. *brassicacearum* NFM421 (YP_004353409)
N-20	*Gammaproteobacteria (Pseudomonas)*	N/4194	Ncontig03449	Rieske alpha	Vanillate monooxygenase	94% *Pseudomonas brassicacearum* subsp. *brassicacearum* NFM421 (PSEBR_a2987)
N-21	*Gammaproteobacteria (Pseudomonas)*	N/0696	Ncontig00520	Other protein families (non-heme dioxygenase)	4-Hydroxyphenylpyruvate dioxygenase	94% *Pseudomonas brassicacearum* subsp. *brassicacearum* NFM421 (PSEBR_a4910)
N-22	*Gammaproteobacteria (Pseudoxanthomonas)*	N/4944	Ncontig04138	Other protein families (non-heme dioxygenase)	4-Hydroxyphenylpyruvate dioxygenase	76% *Stenotrophomonas maltophilia* R551-3 (Smal_3739)
N-23	*Gammaproteobacteria (Pseudomonas)*	N/3453	Ncontig02731	LigB	2,3-Dihydroxyphenylpropionate dioxygenase	95% CbzE2 *Pseudomonas putida* GJ31 (AAX50134)
N-24	*Gammaproteobacteria (Pseudomonas)*	N/4168	Ncontig03424	Cupin	Homogentisate dioxygenase	84% *Pseudomonas brassicacearum* subsp. *brassicacearum* (PSEBR_a4696)
N-25	*Betaproteobacteria (Acidovorax)*	N/4526	Ncontig03752	Rieske ferredoxin	XylT ferredoxin	64% *C. metallidurans* CH34 (Rmet_1325)
N-26	*Gammaproteobacteria (Azotobacter)*	N/0617	Ncontig00459	Rieske ferredoxin	XylT ferredoxin	100% NahT *Pseudomonas stutzeri* AN10 (AAD02147)
N-27	*Gammaproteobacteria (Pseudomonas)*	N/0413	Ncontig00302	Rieske beta	Function unknown	100% *Pseudomonas resinovorans* (NP_758555)
		N/3099	Ncontig02436	Rieske beta	Function unknown	100% *Pseudomonas resinovorans* (NP_758555)
N-28	*Gammaproteobacteria (Pseudomonas)*	N/0898	Ncontig00670	Rieske alpha	Function unknown	94% *Pseudomonas brassicacearum* subsp. *brassicacearum* NFM421 (PSEBR_a5219)
N-29	*Alphaproteobacteria (Novosphingobium)*	N/4279	Ncontig03535	LigB	Function unknown	41% *Mycobacterium* KMS (Mkms5613)
N-30	*Gammaproteobacteria (Pseudomonas)*	N/4292	Ncontig03545	EXDO	Function unknown	94% *Pseudomonas brassicacearum* subsp. *brassicacearum* NFM421 (PSEBR_a2130)
		N/4293	Ncontig03546	EXDO	Function unknown	94% *Pseudomonas brassicacearum* subsp. *brassicacearum* NFM421 (PSEBR_a2130)
N-31	*Gammaproteobacteria (Pseudomonas)*	N/4379	Ncontig03618	Rieske alpha	Function unknown	96% *Pseudomonas putida* W619 (PputW619_4892)
N-32	*Gammaproteobacteria (Pseudomonas)*	N/5144	Ncontig04332	Other protein families (transporter)	Transport protein	98% *Pseudomonas brassicacearum* subsp. *brassicacearum* NFM421 (PSEBR_a1270)
N-33	*Gammaproteobacteria (Pseudomonas)*	N/2198	Ncontig01706	Other protein families (isomerase)	Muconolactone isomerase	100% *Pseudomonas putida* KT24440 (NP_745847)
N-34	*Gammaproteobacteria (Pseudomonas)*	N/1913	Ncontig01476	Rieske alpha	Carbazol dioxygenase	100% *Pseudomonas resinovorans* (NP_758567)
		N/2182	Ncontig01692	Rieske alpha	Carbazol dioxygenase	98% *Pseudomonas resinovorans* (NP_758567)
		N/4095	Ncontig03363	Rieske alpha	Carbazol dioxygenase	100% *Pseudomonas resinovorans* (NP_758567)
N-35	*Gammaproteobacteria (Pseudomonas)*	N/2319	Ncontig01797	LigB	2′-Aminobiphenyl-2,3-diol-1,2-dioxygenase	100% *Pseudomonas resinovorans* (NP_758568)

		N/2320	Ncontig01797	LigB	2′-Aminobiphenyl-2,3-diol-1,2-dioxygenase	100% *Pseudomonas resinovorans* (NP_758569)
N-36	*Gammaproteobacteria (Pseudomonas)*	N/1256	Ncontig00938	Rieske alpha	Anthranilate dioxygenase	100% *Pseudomonas resinovorans* (NP_758548)

[1] We might have overestimated the number of proteins in our metaproteomes by double counting of the fragmented open reading frames in the meta-sequences that were derived from the same gene/protein. In addition, some of the proteins (i.e. dioxygenases) herein investigated are formed by different subunits, that all together form a single protein. To take both issues into consideration the amino acid sequences of proteins with biochemical functions shown to be potentially involved in biodegradation, as suggested by in-house database-BLAST search, were manually BLAST and aligned to the best protein hit in the database to ensure identification of fragments most likely belonging to the same sequence/protein (see first column; "list of unique proteins"). As example, the protein named N-1 (see first column; "list of unique proteins") represents a single dihydroxynaphthalene dioxygenase that contain two fragments (N/2334 and N/3459) located in different short contigs. The protein ID (ORF) and contig where the gene was found within the pyro-sequences are specifically shown.

[2] For taxonomic assignment complete genome sequences were obtained from NCBI and used to map the contigs containing the gene coding the protein(s) with biochemical functions shown to be potentially involved in biodegradation.

[3] Functional assignment of the predicted genes involved in degradation was made on the basis of BLASTP analysis against an in-house database that contained gene sequences encoding enzymes that usually perform key catalysing steps in the aerobic degradation of pollutants via di- and trihydroxylated intermediates (**Table S8**). These enzymes included the following FAMILIES: Rieske non-heme iron oxygenases, type I extradiol dioxygenases of the vicinal oxygen chelate superfamily, the type II or LigB superfamily extradiol dioxygenases, and the type III extradiol dioxygenases belonging to the cupin superfamily as well as intradiol dioxygenases. SUBFAMILIES refer to the specific subfamily within FAMILIES assigned to the protein of interest.

[4] Accession number for the best in-house database/NCBI BLAST hit is specifically shown.

(D) Nbs enzymes features

List of unique proteins[1]	Phylogeny[2]	ORF[1]	Contig[1]	FAMILY[3]	SUBFAMILY[3]	BEST HITS[4]
Nbs-1	*Gammaproteobacteria (Pseudomonas)*	Nbs/0353	contig00020	Rieske ferredoxin	Naphthalene dioxygenase	99% *Pseudomonas stutzeri* AN10 (AF039533)
Nbs-2	*Betaproteobacteria (*genus unclear)	Nbs/4023	contig00679	Rieske ferredoxin	Naphthalene dioxygenase	100% DntAb *Burkholderia cepacia* (AAL50022.1)
Nbs-3	*Gammaproteobacteria (Pseudoxanthomonas)*	Nbs/5873	contig01300	Rieske ferredoxin	Naphthalene dioxygenase	99% *Pseudoxanthomonas spadix* BD-a59 (YP_004930516.1)
Nbs-4	*Betaproteobacteria (Comamonas)*	Nbs/12423	contig04930	Rieske ferredoxin	Naphthalene dioxygenase	100% *Comamonas testosteroni* (AAF72975.1)
Nbs-5	*Gammaproteobacteria (Pseudomonas)*	Nbs/14573	contig06585	Rieske alpha	Naphthalene dioxygenase	93% *Pseudomonas aeruginosa* 2192 (ZP_04934639)
Nbs-6	*Gammaproteobacteria (Pseudomonas)*	Nbs/15601	contig07460	Rieske alpha	Naphthalene dioxygenase	96% *Pseudomonas aeruginosa* 2192 (ZP_04934623)
Nbs-8	*Gammaproteobacteria (Pseudomonas)*	Nbs/0352	contig00020	Rieske alpha	Naphthalene dioxygenase (*Proteobacteria*-like)	96% *Pseudomonas stutzeri* AN10 (ADK11287)
Nbs-9	*Betaproteobacteria (Achromobacter)*	Nbs/4022	contig00679	Rieske alpha	Naphthalene dioxygenase (*Proteobacteria*-like)	96% *Ralstonia* sp. U2 (AAD12610.1)
Nbs-10	*Betaproteobacteria* (unclear genus)	Nbs/14853	contig06818	Rieske alpha	Naphthalene dioxygenase (*Proteobacteria*-like)	86% *Advenella kashmirensis* WT001(ZP_09478240.1)
Nbs-11	*Proteobacteria* - homology ≤ 50% for assigning phylogeny	Nbs/14967	contig06915	Rieske alpha	Naphthalene dioxygenase (*Proteobacteria*-like)	50% *Novosphingobium* sp. PP1Y (YP_004534053.1)
Nbs-12	*Gammaproteobacteria (Pseudoxanthomonas)*	Nbs/15072	contig06999	Rieske alpha	Naphthalene dioxygenase (*Proteobacteria*-like)	100% *Pseudoxanthomonas spadix* BD-a59 (YP_004929419.1)
Nbs-13	*Proteobacteria* - homology ≤ 50% for assigning phylogeny	Nbs/8937	contig02765	Rieske alpha	Naphthalene dioxygenase (*Proteobacteria*-like)	49% *Novosphingobium aromaticivorans* DSM 12444 (YP_001165925)
Nbs-14	*Gammaproteobacteria (Pseudoxanthomonas)*	Nbs/8944	contig02769	Rieske alpha	Naphthalene dioxygenase (*Proteobacteria*-like)	60% *Pseudoxanthomonas aromaticivorans* DSM 12444 (YP_001165876.1)
Nbs-15	*Betaproteobacteria (Achromobacter)*	Nbs/4021	contig00679	Rieske beta	Naphthalene dioxygenase (*Proteobacteria*-like)	99% *Polaromonas naphthalenivorans* CJ2 (YP_982710.1)
Nbs-16	*Gammaproteobacteria (Pseudomonas)*	Nbs/0348	contig00020	EXDO	Dihydroxynaphthalene dioxygenase	98% *Pseudomonas stutzeri* AN10 (AAD02140)
Nbs-17	*Betaproteobacteria (Achromobacter)*	Nbs/6698	contig01626	EXDO	Dihydroxynaphthalene dioxygenase	99% *Polaromonas naphthalenivorans* CJ2 (YP_982706)
Nbs-18	*Gammaproteobacteria (Pseudomonas)*	Nbs/0350	contig00020	Other protein families	cis-Naphthalene dihydrodiol dehydrogenase	99% *Pseudomonas stutzeri* (AAD02138.1)
Nbs-19	*Betaproteobacteria (Achromobacter)*	Nbs/16722	contig08489	Cupin	Gentisate dioxygenases (*Proteobacteria*-like)	99% *Ralstonia* sp U2 (AAD12619)
Nbs-20	*Betaproteobacteria* (genus unclear)	Nbs/20228	contig12968	Cupin	Gentisate dioxygenases (*Proteobacteria*-like)	78% *Delftia acidovorans* SPH-1 (YP_001564719.1)

Nbs-21	*Betaproteobacteria (Achromobacter)*	Nbs/4675	contig00879	Rieske alpha	Salicylate 5-hydroxylase	99% *Ralstonia* sp. U2 (YP_987295.1)
Nbs-22	*Gammaproteobacteria (Pseudomonas)*	Nbs/2420	contig00296	Rieske ferredoxin	Salicylate 5-hydroxylase	87% *Pseudomonas stutzeri* ATCC 14405 (EHY75968.1)
	Betaproteobacteria (genus unclear)	Nbs/18640	contig10655	Rieske alpha	Salicylate 5-hydroxylase	100% *Polaromonas naphthalenivorans* CJ2 (YP_982713.1)
Nbs-23	*Gammaproteobacteria (Pseudomonas)*	Nbs/10351	contig03578	EXDO	2,3-Dihydroxybiphenyl dioxygenase	87% *Pseudomonas putida* GB-1 (YP_001670926.1)
Nbs-24	*Gammaproteobacteria (Pseudomonas)*	Nbs/10637	contig03757	EXDO	2,3-Dihydroxybiphenyl dioxygenase	100% BphC *Pseudomonas putida* OU83 (CAA62978)
Nbs-25	*Betaproteobacteria* (genus unclear)	Nbs/4998	contig00982	Rieske alpha	Benzoate dioxygenase	83% *Lutiella nitroferrum* (ZP_03698426)
Nbs-26	*Betaproteobacteria (Achromobacter)*	Nbs/4996	contig00982	Rieske alpha	Benzoate dioxygenase	82% *Ralstonia eutropha* JMP134 (YP_298599)
Nbs-27	*Proteobacteria* (genus unclear)	Nbs/5872	contig01300	Rieske alpha	Benzoate dioxygenase	55% *Rhodococcus* sp. 19070 (AAK58903)
Nbs-28	*Gammaproteobacteria (Pseudomonas)*	Nbs/8642	contig02594	Rieske alpha	Benzoate dioxygenase	95% *Pseudomonas stutzeri* ATCC17588 (YP_004713948.1)
Nbs-29	*Gammaproteobacteria (Pseudomonas)*	Nbs/12117	contig04720	Rieske alpha	Benzoate dioxygenase	81% *Pseudomonas stutzeri* ATCC 17588 (YP_004713948.1)
Nbs-30	*Gammaproteobacteria (Pseudoxanthomonas)*	Nbs/14388	contig06426	Rieske alpha	Benzoate dioxygenase	95% *Pseudoxanthomonas spadix* BD-a59 (YP_004931670.1)
Nbs-31	*Gammaproteobacteria (Pseudomonas)*	Nbs/9486	contig03072	Rieske alpha	Benzoate dioxygenase	92% *Pseudomonas putida* F1 (YP_001268212)
Nbs-32	*Gammaproteobacteria* (genus unclear)	Nbs/13022	contig05352	Rieske alpha	Benzoate dioxygenase	76% *Marinobacter manganoxydans* MnI7-9 (ZP_09159757.1)
		Nbs/13021	contig05352	Rieske alpha	Benzoate dioxygenase	73% *Marinobacter manganoxydans* MnI7-9 (ZP_09159757.1)
Nbs-33	*Betaproteobacteria* (genus unclear)	Nbs/4996	contig00982	Rieske ferredoxin	Benzoate dioxygenase	94% *Burkholderia pseudomallei* DM98 (ZP_02408055.1)
Nbs-34	*Proteobacteria* - ORF too short (149 AA) for assigning phylogeny (similar homology for the 149 AA long fragment to different proteobacterial proteins)	Nbs/5871	contig01300	Rieske beta	Benzoate dioxygenase	67% *Neptuniibacter* sp. CAR-SF (BAG30836.1)
Nbs-35	*Betaproteobacteria (Achromobacter)*	Nbs/8011	contig02255	Other protein families	Phenol hydroxylase alpha subunit	81% LapN *Ralstonia pickettii* (AAB67108.1)
Nbs-36	*Gammaproteobacteria (Pseudomonas)*	Nbs/7718	contig02108	Other protein families	Phenol hydroxylase alpha subunit	94% LapM *Pseudomonas alkylphenolia* (AAP92391.1)
Nbs-37	*Gammaproteobacteria (Pseudomonas)*	Nbs/7719	contig02108	Other protein families	Phenol hydroxylase alpha subunit	90% LapN *Pseudomonas putida* (BAC75404.1)
Nbs-38	*Gammaproteobacteria (Pseudomonas)*	Nbs/2202	contig00257	EXDO	Catechol 2,3-dioxygenase (Pseudomonads-like)	100% *Pseudomonas* sp. 1YB2 (CAD67837.2)
Nbs-39	*Gammaproteobacteria (Pseudomonas)*	Nbs/6395	contig01501	EXDO	Catechol 2,3-dioxygenase (Pseudomonads-like)	Lapb P*seudomonas* sp. Kl28 (PDB 3HPV_A)

Nbs-40	Homology lower than 50% for assigning phylogeny	Nbs/7244	contig01875	EXDO	Catechol 2,3-dioxygenase	32% *Streptomyces bingchenggensis*_BCW-1 (ADI04647 and YP_004959778.1)
Nbs-41	*Firmicutes*	Nbs/7811	contig02153	EXDO	Catechol 2,3-dioxygenase (Bacillus-like)	87% *Bacillus tusciae* DSM2912 (YP_003590157)
Nbs-42	*Gammaproteobacteria (Pseudomonas)*	Nbs/16157	contig07952	EXDO	Catechol 2,3-dioxygenase of Proteobacteria EXDO_D (EXDO 1.2C according to Eltis)	77% *Pseudomonas putida* (CAA42452)
Nbs-43	*Gammaproteobacteria (Pseudomonas)*	Nbs/17776	contig09609	EXDO	Catechol 2,3-dioxygenase (Pseudomonads-like)	92% *Pseudomonas fluorescens* (AAW81687.1)
Nbs-44	*Gammaproteobacteria (Azotobacter)*	Nbs/18042	contig09918	EXDO	Catechol 2,3-dioxygenase (Pseudomonads-like)	79% LapB *Azotobacter vinelandii* DJ (YP_002800217)
Nbs-45	*Gammaproteobacteria (Pseudomonas)*	Nbs/17777	contig07576	EXDO	Catechol 2,3-dioxygenase of Proteobacteria EXDO_D (EXDO 1.2C according to Eltis)	72% *Pseudomonas fluorescens* (AAW81687.1)
Nbs-46	*Betaproteobacteria* (genus unclear)	Nbs/17781	contig09615	EXDO	Catechol 2,3-dioxygenases of *Proteobacteria* EXDO A, B and C (EXDO 1.2A and B according to Eltis	78% *Azotobacter vinelandii* DJ (YP_002798088.1)
Nbs-47	*Gammaproteobacteria (Pseudomonas)*	Nbs/7685	contig02092	INTRA	Catechol 1,2-dioxygenase	88% *Pseudomonas stutzeri* ATCC 17588 (YP_004713954.1)
Nbs-48	*Betaproteobacteria* (genus unclear)	Nbs/8920	contig02755	INTRA	Catechol 1,2-dioxygenase	66% *Burkholderia gladioli* BSR3 (YP_004350183.1)
Nbs-49	*Gammaproteobacteria (Pseudomonas)*	Nbs/18051	contig09929	INTRA	Catechol 1,2-dioxygenase	84% *Pseudomonas stutzeri* A1501 (YP_001172200)
Nbs-50	*Gammaproteobacteria (Pseudomonas)*	Nbs/18099	contig09988	INTRA	Catechol 1,2-dioxygenase	86% *Pseudomonas stutzeri* ATCC 17588 (YP_004713954.1)
Nbs-51	*Alphaproteobacteria (Rhizobium/Azospirillum)*	Nbs/12122	contig04724	INTRA	Catechol 1,2-dioxygenase	73% *Rhizobium leguminosarum* bv. viciae 3841 (YP_764725)
Nbs-52	*Betaproteobacteria* (genus unclear)	Nbs/3467	contig00528	LigB	Protocatechuate 4,5-dioxygenase subunit alpha	87% *Azoarcus* sp. BH72 (YP_934042.1)
Nbs-53	*Betaproteobacteria* (genus unclear)	Nbs/3466	contig00528	LigB	Protocatechuate 4,5-dioxygenase	78% *Delftia* sp. Cs1-4 (YP_004487058.1)
Nbs-54	*Actinobacteria/Alphaproteobacteria* - ORF too short (123 AA) for assigning phylogeny (similar homology for the 149 AA long fragment to different actinobacterial and proteobacterial proteins)	Nbs/16581	contig08349	LigB	Protocatechuate 4,5-dioxygenase	38% *Mycobacterium* sp. KMS (YP_937705; YP_937705; YP_935611)
Nbs-55	*Betaproteobacteria* (genus unclear)	Nbs/12309	contig04851	LigB	Homoprotocatechuate dioxygenase (unknown function)	69% *Anaerolinea thermophila* UNI-1 (YP_004173079.1)
Nbs-56	*Gammaproteobacteria (Pseudomonas)*	Nbs/14384	contig06425	Rieske	Carbazol dioxygenase	69% *Pseudomonas stutzeri* ATCC 17588 (YP_004715686.1)
Nbs-57	*Gammaproteobacteria (Pseudomonas)*	Nbs/5721	contig01243	Flavoprotein monooxygenase	4-Hydroxybenzoate 3-hydroxylase	69% *Pseudomonas fluorescens* (1BGN_A)
Nbs-58	*Gammaproteobacteria (Pseudomonas)*	Nbs/1332	contig00115	Regulator	LysR family transcriptional regulator	91% *Pseudomonas stutzeri* ATCC 14405 (EHY79459.1)
Nbs-59	*Gammaproteobacteria (Pseudomonas)*	Nbs/13567	contig00124	Regulator	LysR family transcriptional regulator	85% *Pseudomonas stutzeri* ATCC 14405 (AEA84304.1)

Nbs-60	*Gammaproteobacteria (Pseudomonas)*	Nbs/2539	contig00319	Regulator	LysR family transcriptional regulator	52% *Pseudomonas syringae* pv. *phaseolicola* 1448A (YP_274501.1)
Nbs-61	*Gammaproteobacteria* (genus unclear)	Nbs/2616	contig00334	Regulator	LysR family transcriptional regulator	66% *Shewanella sediminis* HAW-EB3 (YP_001475964.1)
Nbs-62	*Gammaproteobacteria (Pseudomonas)*	Nbs/4668	contig00875	Regulator	LysR family transcriptional regulator	96% *Pseudomonas stutzeri* ATCC 14405 (EHY78223.1)
Nbs-63	*Betaproteobacteria (Achromobacter)*	Nbs/4999	contig00982	Regulator	LysR family transcriptional regulator	75% *Ralstonia solanacearum* Po82 (AEG69165.1)
Nbs-64	*Betaproteobacteria* (genus unclear)	Nbs/5152	contig01035	Regulator	LysR family transcriptional regulator	73% *Neisseria shayeganii* 871 (ZP_08886506.1)
Nbs-65	*Gammaproteobacteria (Pseudomonas)*	Nbs/5865	contig01298	Regulator	LysR family transcriptional regulator	99% *Pseudomonas stutzeri* ATCC 14405 (EHY75813.1)
Nbs-66	*Gammaproteobacteria (Pseudomonas)*	Nbs/6080	contig01575	Regulator	LysR family transcriptional regulator	72% *Pseudomonas stutzeri* ATCC 14405 (EHY79459.1)
Nbs-67	*Betaproteobacteria* (genus unclear)	Nbs/6635	contig01600	Regulator	LysR family transcriptional regulator	64% *Advenella kashmirensis* WT001 (ZP_09478423.1)
Nbs-68	*Betaproteobacteria* (genus unclear)	Nbs/7081	contig01805	Regulator	LysR family transcriptional regulator	83% *Advenella kashmirensis* WT001 (ZP_09478363.1)
Nbs-69	*Betaproteobacteria* (genus unclear)	Nbs/8417	contig02464	Regulator	LysR family transcriptional regulator	65% *Advenella kashmirensis* WT001 (ZP_09478104.1)
Nbs-70	*Gammaproteobacteria (Pseudomonas)*	Nbs/9116	contig02868	Regulator	LysR family transcriptional regulator	98% *Pseudomonas stutzeri* A1501 (YP_001173593.1)
Nbs-71	*Gammaproteobacteria (Pseudomonas)*	Nbs/10186	contig03475	Regulator	LysR family transcriptional regulator	72% *Pseudomonas fulva* 12-X (YP_004473340.1)
Nbs-72	*Betaproteobacteria* (genus unclear)	Nbs/11187	contig04110	Regulator	LysR family transcriptional regulator	72% *Burkholderia* sp. CCGE1003 (YP_003909685.1)
Nbs-73	*Gammaproteobacteria (Pseudomonas)*	Nbs/11486	contig04304	Regulator	LysR family transcriptional regulator	85% *Pseudomonas mendocina* ymp (YP_001187511.1)
Nbs-74	*Betaproteobacteria (Achromobacter)*	Nbs/11557	contig04352	Regulator	LysR family transcriptional regulator	79% *Achromobacter xylosoxidans* A8 (YP_003977787.1)
Nbs-75	*Gammaproteobacteria (Pseudomonas)*	Nbs/13451	contig05768	Regulator	LysR family transcriptional regulator	100% *Pseudomonas stutzeri* ATCC 17588 (YP_004716500.1)
Nbs-76	*Gammaproteobacteria (Pseudomonas)*	Nbs/14525	contig06548	Regulator	LysR family transcriptional regulator	82% *Pseudomonas stutzeri* ATCC 17588 (YP_004714404.1)
Nbs-77	*Betaproteobacteria* (genus unclear)	Nbs/14961	contig06910	Regulator	LysR family transcriptional regulator	61% *Pusillimonas* sp. T7-7 (YP_004416250.1)
Nbs-78	*Firmicutes (Clostridium)*	Nbs/15469	contig07138	Regulator	LysR family transcriptional regulator	54% *Clostridium* sp. D5 (ZP_08131246.1)
Nbs-79	*Gammaproteobacteria (Pseudomonas)*	Nbs/15469	contig07340	Regulator	LysR family transcriptional regulator	99% *Pseudomonas putida* S16 (YP_004700228.1)
Nbs-80	*Gammaproteobacteria (Azotobacter)*	Nbs/15868	contig07702	Regulator	LysR family transcriptional regulator	86% *Azotobacter vinelandii* DJ (YP_002801308.1)
Nbs-81	*Gammaproteobacteria (Pseudomonas)*	Nbs/18359	contig00233	Regulator	LysR family transcriptional regulator	57% *Pseudomonas aeruginosa* PA7 (YP_001348799.1)
Nbs-82	*Gammaproteobacteria (Pseudomonas)*	Nbs/9636	contig03158	Other protein families	Muconate cycloisomerase	96% *Pseudomonas stutzeri* ATCC 17588 (YP_004713952.1)

Nbs-83	*Gammaproteobacteria (Pseudomonas)*	Nbs/0346	contig00020	Other protein families (isomerase)	2-Hydroxychromene-2-carboxylate isomerase	100% *Pseudomonas stutzeri* (AAD02142.1)
Nbs-84	*Gammaproteobacteria (Pseudomonas)*	Nbs/12628	contig05075	Other protein families (isomerase)	2-Hydroxychromene-2-carboxylate isomerase	74% *Pseudomonas aeruginosa* PAb1 (ZP_06876104.1)
Nbs-85	*Betaproteobacteria (Achromobacter)*	Nbs/13166	contig05458	Other protein families (isomerase)	2-Hydroxychromene-2-carboxylate isomerase	72% *Ralstonia* sp. U2 (AAD12617.1)
Nbs-86	*Gammaproteobacteria (Pseudomonas)*	Nbs/0349	contig00020	Other protein families (dehydrogenase)	Salicylaldehyde dehydrogenase	99% *Pseudomonas stutzeri* (AAD02139.1)
Nbs-87	*Betaproteobacteria (Polaromonas)*	Nbs/6699	contig01626	Other protein families (dehydrogenase)	Salicylaldehyde dehydrogenase	99% *Polaromonas naphthalenivorans* CJ2 (YP_982707.1)
Nbs-88	*Bacteroidetes/Chlorobi (Bacteroides)*	Nbs/10177	contig03468	Other protein families (isomerase)	Muconate cycloisomerase	60% *Bacteroides* sp. 3_2_5 (ZP_04842137.1)
Nbs-89	*Bacteroidetes/Chlorobi (Bacteroides)*	Nbs/11141	contig04082	Other protein families (isomerase)	Muconate cycloisomerase	72% *Bacteroides fragilis* YCH46 (YP_100030.1)
	Gammaproteobacteria (Pseudomonas)	Nbs/19441	contig11724	Other protein families (isomerase)	Muconate cycloisomerase	89% *Pseudomonas stutzeri* ATCC 17588 (YP_004713952.1)
Nbs-90	*Betaproteobacteria* (genus unclear)	Nbs/13304	contig05566	Other protein families	Maley acetate reductase	57% *Polaromonas* sp JS666B (CP000316; YP_548891.1)
Nbs-91	*Betaproteobacteria (Polaromonas)*	Nbs/13304	contig05566	Other protein families (reductase)	Maleylacetate reductase	57% *Polaromonas* sp. JS666 (YP_548891.1)
Nbs-92	*Gammaproteobacteria (Pseudomonas)*	Nbs/6412	contig01507	Other protein families (non-heme dioxygenase)	4-Hydroxyphenylpyruvate dioxygenase	99% *Pseudomonas stutzeri* ATCC 14405 (EHY76970.1)
Nbs-93	*Betaproteobacteria (Achromobacter)*	Nbs/13523	contig05736	Other protein families (non-heme dioxygenase)	4-Hydroxyphenylpyruvate dioxygenase	85% *Achromobacter arsenitoxydans* SY8 (ZP_09298025.1)
Nbs-94	*Betaproteobacteria (Achromobacter)*	Nbs/14476	contig06506	Other protein families (non-heme dioxygenase)	4-Hydroxyphenylpyruvate dioxygenase	85% *Achromobacter xylosoxidans* C54 (EFV83067.1)
Nbs-95	*Firmicutes* (genus unclear)	Nbs/3821	contig00621	Other protein families (lyase)	2-Oxopent-4-enoate hydratase/fumarylacetoacetate hydrolase	60% *Anaerolinea thermophila* UNI-1 (YP_004173999.1)
Nbs-96	*Gammaproteobacteria (Pseudomonas)*	Nbs/13150	contig05447	Other protein families (lyase)	2-Oxopent-4-enoate hydratase/fumarylacetoacetate hydrolase	70% *Pseudomonas putida* W619 (YP_001748852.1)
Nbs-97	*Gammaproteobacteria (Pseudomonas)*	Nbs/15955	contig07774	Other protein families (lyase)	2-Oxopent-4-enoate hydratase/fumarylacetoacetate hydrolase	93% *Pseudomonas alkylphenolia* (AAP92396.1)
Nbs-98	*Gammaproteobacteria (Pseudomonas)*	Nbs/10105	contig03427	Other protein families (tautomerase)	4-Oxalocrotonate tautomerase	98% *Pseudomonas putida* (AAQ89682.1)
Nbs-99	*Gammaproteobacteria (Pseudoxanthomonas)*	Nbs/10847	contig03888	Other protein families (tautomerase)	4-Oxalocrotonate tautomerase	100% *Pseudoxanthomonas spadix* BD-a59 (YP_004930525.1)
Nbs-100	*Gammaproteobacteria (Pseudomonas)*	Nbs/6273	contig01450	Other protein families (aldolase)	4-Hydroxy 2-oxovalerate aldolase	86% *Pseudomonas putida* (NP_542861.1)
Nbs-101	*Gammaproteobacteria (Pseudomonas)*	Nbs/9038	contig02820	Other protein families (aldolase)	4-Hydroxy 2-oxovalerate aldolase	91% *Pseudomonas putida* (BAD19055.1)
Nbs-102	*Gammaproteobacteria (Pseudomonas)*	Nbs/13881	contig06025	Other protein families (aldolase)	4-Hydroxy 2-oxovalerate aldolase	95% *Pseudomonas putida* F1 (YP_001268202.1)
Nbs-103	*Gammaproteobacteria (Pseudoxanthomonas)*	Nbs/16383	contig08171	Other protein families (aldolase)	4-Hydroxy 2-oxovalerate aldolase	100% *Pseudoxanthomonas spadix* BD-a59

						(YP_004930564.1)
Nbs-104	*Gammaproteobacteria (Pseudomonas)*	Nbs/4424	contig00798	Other protein families (aldolase)	4-Hydroxy 2-oxovalerate aldolase	100% *Pseudomonas putida* (Q9ZI56.1)
Nbs-105	*Gammaproteobacteria (Pseudomonas)*	Nbs/8785	contig02681	Other protein families (lyase)	4-Oxalocrotonate decarboxylase/fumarylacetoacetate hydrolase	88% *Pseudomonas alkylphenolia* (AAP92397.1)
Nbs-106	*Gammaproteobacteria (Pseudoxanthomonas)*	Nbs/10846	contig03888	Other protein families (lyase)	4-Oxalocrotonate decarboxylase/fumarylacetoacetate hydrolase	99% *Pseudoxanthomonas spadix* BD-a59 (YP_004930524.1)
Nbs-107	*Gammaproteobacteria (Pseudomonas)*	Nbs/13828	contig05983	Other protein families (lyase)	4-Oxalocrotonate decarboxylase/fumarylacetoacetate hydrolase	99% *Pseudomonas putida* (AAQ89681.1)
Nbs-108	*Gammaproteobacteria (Pseudoxanthomonas)*	Nbs/13949	contig06074	Other protein families (lyase)	4-Oxalocrotonate decarboxylase/fumarylacetoacetate hydrolase	100% *Pseudoxanthomonas spadix* BD-a59 (YP_004930523.1)
Nbs-109	*Gammaproteobacteria (Pseudomonas)*	Nbs/14206	contig06280	Other protein families (lyase)	4-Oxalocrotonate decarboxylase/fumarylacetoacetate hydrolase	98% *Pseudomonas putida* (AAQ89682.1) (YP_709353.1)
Nbs-110	*Gammaproteobacteria (Pseudomonas)*	Nbs/4424	contig00798	Other protein families (lyase)	4-Oxalocrotonate decarboxylase/fumarylacetoacetate hydrolase	100% *Pseudomonas stutzeri* (AAD02151.1)
Nbs-111	*Gammaproteobacteria (Pseudomonas)*	Nbs/4425	contig02681	Other protein families (dehydrogenase)	Acetaldehyde dehydrogenase	91% *Pseudomonas alkylphenolia* (AAP92399.1)
Nbs-112	*Gammaproteobacteria (Pseudomonas)*	Nbs/9039	contig02820	Other protein families (dehydrogenase)	Acetaldehyde dehydrogenase	91% *Azotobacter vinelandii* DJ ()YP_002800198.1)
Nbs-113	*Gammaproteobacteria (Pseudomonas)*	Nbs/13882	contig06025	Other protein families (dehydrogenase)	Acetaldehyde dehydrogenase	95% *Pseudomonas putida* F1 (YP_001268203.1)
Nbs-114	*Gammaproteobacteria (Pseudomonas)*	Nbs/4425	contig00798	Other protein families (dehydrogenase)	Acetaldehyde dehydrogenase	100% *Pseudomonas stutzeri* (Q9ZI57.1)
Nbs-115	*Gammaproteobacteria (Pseudomonas)*	Nbs/6272	contig01450	Other protein families (dehydrogenase)	Acetaldehyde dehydrogenase	90% *Pseudomonas putida* (YP_709351.1)

[1] We might have overestimated the number of proteins in our metaproteomes by double counting of the fragmented open reading frames in the meta-sequences that were derived from the same gene/protein. In addition, some of the proteins (i.e. dioxygenases) herein investigated are formed by different subunits, that all together form a single protein. To take both issues into consideration the amino acid sequences of proteins with biochemical functions shown to be potentially involved in biodegradation, as suggested by in-house database-BLAST search, were manually BLAST and aligned to the best protein hit in the database to ensure identification of fragments most likely belonging to the same sequence/protein (see first column; "list of unique proteins"). As example, the protein named Nbs-31 (see first column; "list of unique proteins") represents a single benzoate dioxygenase that contain two fragments (Nbs/11815 and Nbs/11816) located in same contig but having a frame shift. The protein ID (ORF) and contig where the gene was found within the pyro-sequences are specifically shown.

[2] For taxonomic assignment complete genome sequences were obtained from NCBI and used to map the contigs containing the gene coding the protein(s) with biochemical functions shown to be potentially involved in biodegradation.

[3] Functional assignment of the predicted genes involved in degradation was made on the basis of BLASTP analysis against an in-house database that contained gene sequences encoding enzymes that usually perform key catalysing steps in the aerobic degradation of pollutants via di- and trihydroxylated intermediates (**Table S8**). These enzymes included the following FAMILIES: Rieske non-heme iron oxygenases, type I extradiol dioxygenases of the vicinal oxygen chelate superfamily, the type II or LigB superfamily extradiol dioxygenases, and the type III extradiol dioxygenases belonging to the cupin superfamily as well as intradiol dioxygenases. SUBFAMILIES refer to the specific subfamily within FAMILIES assigned to the protein of interest.

[4] Accession number for the best in-house database/NCBI BLAST hit is specifically shown.

Table S6 Proteins of CN1 and CN2 communities identified and quantified by metaproteomic approaches. Protein annotation is specifically shown. The 816 most abundant proteins, among a total number of 1116, are listed below. For raw data, please contact authors directly.

Contig ID[1]	Protein ID[1]	Quantitative Value CN1[2]	Quantitative Value CN2[2]	Rel conc. CN1 (%)[2]	Rel conc. CN2 (%)[2]	Sequence Coverage [%][1]	Sequence Length[1]	PEP[3]
CN2/15081	pyridoxamine 5'-phosphate oxidase [EC:1.4.3.5]	0	1882812.51	0.0000	0.0587	39.1	215	3.1186E-20
CN2/17055	4-hydroxythreonine-4-phosphate dehydrogenase [EC:1.1.1.262]	0	2911516.89	0.0000	0.0908	6.4	329	0.00031214
CN2/3894	erythronate-4-phosphate dehydrogenase [EC:1.1.1.290]	0	167456.45	0.0000	0.0052	8.7	380	5.619E-08
CN1/27032	outer membrane protein OmpU	14407179.48	0	1.3173	0.0000	22.4	277	4.0763E-149
CN2/0960	3-hydroxyisobutyrate dehydrogenase [EC:1.1.1.31]	231897.30	3825444.01	0.0212	0.1193	40.8	299	2.4653E-147
CN2/18291	3-hydroxyisobutyrate dehydrogenase [EC:1.1.1.31]	0	710782.80	0.0000	0.0222	21.1	294	6.8887E-39
CN2/0869	3-isopropylmalate dehydrogenase [EC:1.1.1.85]	310579.92	1422731.95	0.0284	0.0444	31.4	360	6.1985E-67
CN1/20728	3-isopropylmalate dehydrogenase [EC:1.1.1.85]	1325725.07	0	0.1212	0.0000	10.9	349	2.4663E-07
CN2/0783	3-demethylubiquinone-9 3-methyltransferase [EC:2.1.1.- 2.1.1.64]	54570.04	2601783.80	0.0050	0.0812	55.6	232	3.4007E-74
CN2/17744	ubiquinone/menaquinone biosynthesis methyltransferase [EC:2.1.1.-]	0	618305.83	0.0000	0.0193	15.2	256	8.4864E-06
CN2/16713	Gentisate dioxygenase	131660.58	20882908.95	0.0120	0.6514	73.2	347	0
CN2/16712	Fumarylpyruvate hydrolase	2065760.52	17048900.21	0.1889	0.5318	49.4	231	8.3934E-100
CN1/14425	acylpyruvate hydrolase [EC:3.7.1.5]	950218.17	1071446.20	0.0869	0.0334	30	190	1.1181E-12
CN1/23999	Gentisate dioxygenase	21941841.03	14450193.18	2.0062	0.4507	63.6	55	2.2886E-23
CN1/6354	Gentisate dioxygenase	21005131.05	12753406.93	1.9205	0.3978	35.6	191	6.9311E-127
CN2/3389	Fumarylpyruvate hydrolase	5533982.89	0	0.5060	0.0000	46.7	90	1.8627E-43
CN1/3872	Gentisate dioxygenase	3608315.64	0	0.3299	0.0000	10.1	346	2.1153E-06
CN1/23316	C4-dicarboxylate-binding protein DctP	1331164.23	137698.96	0.1217	0.0043	41.9	124	1.7775E-17
CN1/25779	putative tricarboxylic transport membrane protein	2402357.22	1882130.83	0.2197	0.0587	23.4	175	2.0775E-51
CN1/20966	C4-dicarboxylate-binding protein DctP	13767763.74	126076.58	1.2588	0.0039	57.9	340	3.862E-108
CN1/23240	phosphoglycerate transport regulatory protein PgtC	4978799.30	489750.33	0.4552	0.0153	27.9	341	1.0556E-32
CN1/21648	C4-dicarboxylate-binding protein DctP	1104116.09	0	0.1010	0.0000	20	325	3.1874E-12
CN2/17891	C4-dicarboxylate-binding protein DctP	56941.18	17565303.35	0.0052	0.5479	70.1	331	0
CN1/9660	C4-dicarboxylate-binding protein DctP	2422204.74	98994.71	0.2215	0.0031	36.4	173	2.3756E-33
CN2/13571	cysteine desulfurase [EC:2.8.1.7]	36477.68	1053425.95	0.0033	0.0329	7.2	404	8.3251E-07
CN2/14529	1-deoxy-D-xylulose-5-phosphate synthase [EC:2.2.1.7]	0	998689.72	0.0000	0.0311	6.8	632	6.1754E-07
CN2/14147	sulfite reductase (NADPH) hemoprotein beta-component [EC:1.8.1.2]	515786.96	1551134.69	0.0472	0.0484	26.3	552	5.0207E-134
CN2/16708	Fumarylpyruvate hydrolase	40749.39	6930602.57	0.0037	0.2162	49.1	216	2.2104E-122
CN2/2666	1,4-alpha-glucan branching enzyme [EC:2.4.1.18]	13227.62	1735778.01	0.0012	0.0541	20.8	736	9.0884E-27
CN2/4159	alpha-amylase [EC:3.2.1.1]	0	1436607.87	0.0000	0.0448	12.7	667	8.2037E-14
CN2/4160	maltose alpha-D-glucosyltransferase [EC:5.4.99.16]	0	263478.19	0.0000	0.0082	4.2	1106	1.2942E-32
CN2/5230	cysteine desulfurase / selenocysteine lyase [EC:2.8.1.7 4.4.1.16]	0	609379.84	0.0000	0.0190	11.3	337	3.8697E-13
CN2/17124	RNA polymerase sigma-70 factor, ECF subfamily	0	3520443.17	0.0000	0.1098	21.2	193	1.1876E-101
CN1/9324	chaperonin GroEL	0	1613359.60	0.0000	0.0503	29.3	116	3.0529E-07
CN2/4031	chaperonin GroEL	0.00	3183710.81	0.0000	0.0993	30.1	209	2.0129E-21
CN1/13133	molecular chaperone DnaK	1844968.56	1193837.30	0.1687	0.0372	18.9	492	2.9103E-73
CN1/18429	chaperonin GroEL	3961.08	16628.43	0.0004	0.0005	17.8	348	2.1114E-35
CN2/14323	ATP-dependent RNA helicase DeaD	0	1058558.97	0.0000	0.0330	8.6	556	3.2325E-07
CN2/14484	host factor-I protein	0	2872118.05	0.0000	0.0896	47.6	84	7.0627E-20
CN2/12280	molecular chaperone DnaK	233076.10	20624368.01	0.0213	0.6433	61.5	637	0
CN2/18240	ribonuclease E [EC:3.1.26.12]	0	2068087.78	0.0000	0.0645	10.8	1008	2.4128E-38

CN2/17619	transcription termination factor Rho	84508.40	1126283.29	0.0077	0.0351	25.8	419	5.1957E-53
CN2/0474	chaperonin GroEL	372759.02	18806642.42	0.0341	0.5866	91.6	546	0
CN1/25232	chaperonin GroEL	909839.79	214510.92	0.0832	0.0067	22.9	279	7.6251E-16
CN1/26923	chaperonin GroEL	1041854.30	389610.81	0.0953	0.0122	39.6	53	0.0020169
CN1/24862	chaperonin GroEL	415368.45	1860106.82	0.0380	0.0580	36.3	245	7.2933E-59
CN1/18243	small subunit ribosomal protein S7	0	4810586.34	0.0000	0.1500	22.9	170	0.00019616
CN2/4130	large subunit ribosomal protein L1	852972.28	0	0.0780	0.0000	22.3	233	1.5129E-185
CN1/1687	large subunit ribosomal protein L7/L12	11601813.46	0	1.0608	0.0000	23.4	124	2.3961E-51
CN2/11678	large subunit ribosomal protein L7/L12	0	11056479.44	0.0000	0.3449	39.3	122	4.9795E-34
CN2/11679	large subunit ribosomal protein L10	0	7663592.46	0.0000	0.2390	32.5	166	6.6914E-39
CN2/11680	large subunit ribosomal protein L1	52740.69	9172087.12	0.0048	0.2861	50.2	231	1.8424E-138
CN2/11681	large subunit ribosomal protein L11	0	8633070.80	0.0000	0.2693	23.6	157	8.8221E-32
CN2/12035	large subunit ribosomal protein L20	0	1451217.90	0.0000	0.0453	16.1	118	0.00014532
CN2/12946	small subunit ribosomal protein S7	0	9170494.64	0.0000	0.2860	56.4	156	5.9266E-62
CN2/13076	large subunit ribosomal protein L25	96083.86	11088455.84	0.0088	0.3459	61.5	205	1.0862E-199
CN2/14178	small subunit ribosomal protein S1	213422.52	8035154.22	0.0195	0.2506	47.1	560	0
CN2/14474	small subunit ribosomal protein S18	0	1143944.79	0.0000	0.0357	31.6	76	4.7843E-07
CN2/14475	small subunit ribosomal protein S6	22439.25	7640464.01	0.0021	0.2383	53.6	138	8.302E-152
CN2/15993	large subunit ribosomal protein L21	0	1830288.47	0.0000	0.0571	33	103	2.6447E-10
CN2/17151	large subunit ribosomal protein L19	0	6964790.70	0.0000	0.2172	45.7	116	1.0096E-35
CN2/17313	small subunit ribosomal protein S4	39654.31	3846437.80	0.0036	0.1200	50.5	206	3.1511E-82
CN2/17314	small subunit ribosomal protein S11	0	3515999.77	0.0000	0.1097	31	129	1.4812E-11
CN2/17315	small subunit ribosomal protein S13	0	22642399.31	0.0000	0.7062	31.4	118	6.0912E-15
CN2/17317	large subunit ribosomal protein L15	0	4712716.90	0.0000	0.1470	29.2	144	6.0751E-119
CN2/17319	small subunit ribosomal protein S5	0	3614405.40	0.0000	0.1127	51.8	166	8.8409E-98
CN2/17320	large subunit ribosomal protein L18	0	1144953.73	0.0000	0.0357	30.2	116	5.0684E-16
CN2/17321	large subunit ribosomal protein L6	0	6772533.70	0.0000	0.2112	32.8	177	8.3105E-75
CN2/17324	large subunit ribosomal protein L5	27161.81	7627744.03	0.0025	0.2379	33.5	179	2.3422E-20
CN2/17325	large subunit ribosomal protein L24	0	5010192.49	0.0000	0.1563	56.7	104	1.1949E-14
CN2/17331	large subunit ribosomal protein L22	0	10636016.26	0.0000	0.3317	41.8	110	2.4614E-51
CN2/17335	large subunit ribosomal protein L4	0	13331872.96	0.0000	0.4158	58.5	200	1.4827E-37
CN2/17336	large subunit ribosomal protein L3	0	6284208.23	0.0000	0.1960	41.7	211	7.8231E-36
CN2/17337	small subunit ribosomal protein S10	0	9684142.22	0.0000	0.3021	35	103	8.0886E-18
CN2/17860	large subunit ribosomal protein L28	0	8763979.96	0.0000	0.2734	42.3	78	1.2032E-46
CN2/9219	small subunit ribosomal protein S9	0	3284673.86	0.0000	0.1025	46.2	130	3.2793E-18
CN2/9220	large subunit ribosomal protein L13	0	5482928.44	0.0000	0.1710	54.9	142	3.1717E-40
CN2/14472	large subunit ribosomal protein L9	0	7414002.56	0.0000	0.2312	35.1	148	1.545E-121
CN2/17329	large subunit ribosomal protein L16	0	8639835.11	0.0000	0.2695	35.3	136	1.2174E-31
CN2/17330	small subunit ribosomal protein S3	134478.93	5172753.97	0.0123	0.1613	25.4	228	3.8884E-19
CN2/17333	large subunit ribosomal protein L2	20287.08	7888341.27	0.0019	0.2460	51.3	273	1.4121E-138
CN2/15070	small subunit ribosomal protein S2	89737.71	3883499.76	0.0082	0.1211	66.4	241	7.4582E-206
CN2/17281	riboflavin synthase beta chain [EC:2.5.1.-]	0	5019985.00	0.0000	0.1566	78.5	158	1.4974E-77
CN2/17282	3,4-dihydroxy 2-butanone 4-phosphate synthase [EC:4.1.99.12]	0	1635216.27	0.0000	0.0510	44.3	359	7.8308E-123
CN2/12768	malate dehydrogenase (quinone) [EC:1.1.5.4]	264384.20	2091633.76	0.0242	0.0652	8.7	503	6.2615E-11
CN2/13721	2-isopropylmalate synthase [EC:2.3.3.13]	165198.68	5276712.57	0.0151	0.1646	36.8	514	3.384E-82
CN2/17520	pyruvate, water dikinase [EC:2.7.9.2]	228297.32	4654085.96	0.0209	0.1452	30.2	762	5.6568E-181
CN2/12080	dihydroorotase [EC:3.5.2.3]	118485.35	1759082.15	0.0108	0.0549	31.2	423	2.7079E-33
CN2/15718	dihydroorotate oxidase [EC:1.3.3.1]	718793.78	3213529.36	0.0657	0.1002	37.6	343	3.1619E-75
CN2/17144	dihydroorotase [EC:3.5.2.3]	0	944786.17	0.0000	0.0295	47.6	347	6.8889E-77

CN2/12359	DNA-directed RNA polymerase subunit beta' [EC:2.7.7.6]	0	6256704.14	0.0000	0.1952	42.9	63	2.4264E-09
CN2/12833	phosphoribosylamine--glycine ligase [EC:6.3.4.13]	0	3194123.80	0.0000	0.0996	27.9	430	1.1E-69
CN2/14389	ribonucleoside-diphosphate reductase alpha chain [EC:1.17.4.1]	0	774071.28	0.0000	0.0241	6.2	974	3.3978E-10
CN2/14558	adenylate kinase [EC:2.7.4.3]	107516.66	6593664.59	0.0098	0.2057	74	215	3.6315E-154
CN2/16347	5-(carboxyamino)imidazole ribonucleotide mutase [EC:5.4.99.18]	0	1220494.05	0.0000	0.0381	41.4	162	1.3202E-45
CN2/16348	5-(carboxyamino)imidazole ribonucleotide synthase [EC:6.3.4.18]	0	2927656.68	0.0000	0.0913	16.9	361	2.0016E-87
CN2/17657	phosphoribosylaminoimidazole-succinocarboxamide synthase	0	1642735.99	0.0000	0.0512	25.6	289	7.5157E-18
CN2/3446	DNA-directed RNA polymerase subunit beta' [EC:2.7.7.6]	19320.72	2125577.62	0.0018	0.0663	30.1	269	4.4014E-39
CN2/13564	nucleoside-diphosphate kinase [EC:2.7.4.6]	30588.53	11877603.95	0.0028	0.3705	70.6	143	2.9903E-104
CN2/13820	phosphoribosylformylglycinamidine synthase [EC:6.3.5.3]	47117.43	2241903.08	0.0043	0.0699	42.4	536	2.0107E-90
CN2/12242	phosphoribosylformylglycinamidine synthase [EC:6.3.5.3]	216343.94	1837585.24	0.0198	0.0573	27.7	762	1.518E-100
CN2/12944	DNA-directed RNA polymerase subunit beta' [EC:2.7.7.6]	157900.95	1741393.94	0.0144	0.0543	36.7	1066	8.1469E-148
CN2/17312	DNA-directed RNA polymerase subunit alpha [EC:2.7.7.6]	329157.35	5730385.91	0.0301	0.1787	40.5	333	1.7441E-106
CN2/0232	DNA-directed RNA polymerase subunit beta [EC:2.7.7.6]	265228.93	1922771.33	0.0243	0.0600	50.5	757	8.4161E-245
CN2/16758	polyribonucleotide nucleotidyltransferase [EC:2.7.7.8]	144892.36	4355526.28	0.0132	0.1359	52.4	701	0
CN2/11677	DNA-directed RNA polymerase subunit beta [EC:2.7.7.6]	0	2013904.15	0.0000	0.0628	39.4	513	5.8001E-89
CN1/19309	adenylate kinase [EC:2.7.4.3]	569922.42	73292.30	0.0521	0.0023	44.2	215	1.6732E-59
CN2/14003	propionyl-CoA carboxylase beta chain [EC:6.4.1.3]	0	843730.31	0.0000	0.0263	11.8	535	5.6385E-19
CN2/17524	2-methylcitrate synthase [EC:2.3.3.5]	92246.85	5359880.16	0.0084	0.1672	54.9	375	1.251E-134
CN2/17525	methylisocitrate lyase [EC:4.1.3.30]	0	5254574.80	0.0000	0.1639	33.6	295	1.7313E-86
CN2/17862	beta-alanine--pyruvate transaminase [EC:2.6.1.18]	0	947523.04	0.0000	0.0296	19	441	6.3094E-13
CN2/7240	acetate kinase [EC:2.7.2.1]	0	1258042.75	0.0000	0.0392	10.9	395	1.0129E-11
CN2/18205	adenosine deaminase [EC:3.5.4.4]	174552.15	1082610.80	0.0160	0.0338	38.6	316	3.5297E-57
CN2/13695	oxygen-independent coproporphyrinogen III oxidase [EC:1.3.99.22]	58887.12	1684150.35	0.0054	0.0525	28.1	462	1.6405E-34
CN2/16575	glutamate-1-semialdehyde 2,1-aminomutase [EC:5.4.3.8]	0	2540807.76	0.0000	0.0792	15.7	427	2.8837E-45
CN2/17615	porphobilinogen synthase [EC:4.2.1.24]	622482.40	5516858.28	0.0569	0.1721	37.7	337	3.6403E-88
CN2/1831	uroporphyrinogen decarboxylase [EC:4.1.1.37]	0	2779616.71	0.0000	0.0867	22.6	354	1.003E-26
CN2/2915	uroporphyrin-III C-methyltransferase [EC:2.1.1.107]	74464.40	2584162.24	0.0068	0.0806	24.9	385	3.22E-176
CN2/2917	hydroxymethylbilane synthase [EC:2.5.1.61]	0	1624017.32	0.0000	0.0507	21.5	312	1.8083E-39
CN2/5150	nicotinate-nucleotide--dimethylbenzimidazole	0	2449379.88	0.0000	0.0764	41.5	349	5.2596E-163
CN2/12888	uroporphyrin-III C-methyltransferase [EC:2.1.1.107]	15551.03	1402961.59	0.0014	0.0438	61.2	278	1.0001E-167
CN2/6835	glucose-1-phosphate thymidylyltransferase [EC:2.7.7.24]	0	2789117.38	0.0000	0.0870	16.2	925	8.2554E-93
CN1/9853	elongation factor EF-Tu [EC:3.6.5.3]	1884672.00	809734.38	0.1723	0.0253	41.3	298	3.5669E-68
CN1/22569	elongation factor EF-Tu [EC:3.6.5.3]	1715187.40	5260772.07	0.1568	0.1641	39.1	396	2.4168E-247
CN1/20602	elongation factor EF-Tu [EC:3.6.5.3]	2392303.35	118663.02	0.2187	0.0037	32.4	272	3.1502E-50
CN1/20438	elongation factor EF-Tu [EC:3.6.5.3]	1308057.27	4395080.42	0.1196	0.1371	19	105	3.9137E-07
CN2/11684	elongation factor EF-Tu [EC:3.6.5.3]	99146.20	51630868.31	0.0091	1.6104	80.1	397	0
CN2/16770	prephenate dehydratase [EC:4.2.1.51]	0	615817.95	0.0000	0.0192	16.4	365	1.2433E-20
CN2/17853	3-phosphoshikimate 1-carboxyvinyltransferase [EC:2.5.1.19]	0	1233111.15	0.0000	0.0385	9.9	616	2.4779E-11
CN2/17535	3-deoxy-7-phosphoheptulonate synthase [EC:2.5.1.54]	0	1413615.39	0.0000	0.0441	33.8	358	5.8262E-51
CN2/16710	Salicylate 5-hydroxylase	25115.50	14120100.34	0.0023	0.4404	59.6	156	4.5169E-182
CN2/16711	Salicylate 5-hydroxylase	80770.62	7837361.01	0.0074	0.2445	48.6	418	4.7232E-148
CN1/10034	Naphthalene dioxygenase (alpha)	19070325.78	11367157.77	1.7436	0.3546	63.3	180	2.0697E-135
CN2/11944	Naphthalene dihydrodiol dehydrogenase	142283.58	25349616.89	0.0130	0.7907	68	259	1.9574E-268
CN2/11946	Naphthalene dioxygenase (alpha)	1072816.79	17951113.23	0.0981	0.5599	74.8	449	0
CN2/8666	Naphthalene dioxygenase (alpha)	68772.39	3527985.67	0.0063	0.1100	76.9	65	1.1592E-106
CN1/20819	Salicylate 5-hydroxylase	0	794378.12	0.0000	0.0248	6.1	392	0.00029541
CN2/12753	UDP-N-acetylglucosamine--N-acetylmuramyl-(pentapeptide)	0	5747093.31	0.0000	0.1793	23	356	1.7921E-15
CN2/17999	D-alanyl-D-alanine carboxypeptidase (penicillin-binding protein 5/6)	0	1847765.18	0.0000	0.0576	17.9	375	8.4586E-19

CN2/13077	ribose-phosphate pyrophosphokinase [EC:2.7.6.1]	0	1418612.54	0.0000	0.0442	7.9	329	9.7449E-06
CN2/4438	phosphogluconate dehydratase [EC:4.2.1.12]	1178131.57	3613124.62	0.1077	0.1127	21.8	609	6.2026E-30
CN1/20692	transaldolase [EC:2.2.1.2]	1010030.63	0	0.0923	0.0000	30.8	302	1.0991E-202
CN2/4462	beta-lactamase [EC:3.5.2.6]	110401.15	337227.14	0.0101	0.0105	16	288	3.1537E-13
CN1/23302	ketol-acid reductoisomerase [EC:1.1.1.86]	1726033.97	70985.85	0.1578	0.0022	15.7	223	3.4843E-07
CN2/18166	branched-chain amino acid aminotransferase [EC:2.6.1.42]	45614.13	1864739.15	0.0042	0.0582	12.1	307	2.0411E-08
CN1/20942	branched-chain amino acid aminotransferase [EC:2.6.1.42]	1242118.39	0	0.1136	0.0000	35	117	3.9651E-85
CN2/15458	pantoate--beta-alanine ligase [EC:6.3.2.1]	0	3183720.16	0.0000	0.0993	14.7	285	1.8252E-09
CN2/17486	ketol-acid reductoisomerase [EC:1.1.1.86]	131634.39	10479385.10	0.0120	0.3269	57.6	349	1.1558E-221
CN2/15629	dihydroxy-acid dehydratase [EC:4.2.1.9]	275367.66	5806439.35	0.0252	0.1811	32.7	542	2.1539E-121
CN1/14088	ketol-acid reductoisomerase [EC:1.1.1.86]	850245.70	296389.67	0.0777	0.0092	32.4	71	2.9196E-12
CN1/5217	F-type H+-transporting ATPase subunit beta [EC:3.6.3.14]	2076100.68	186223.25	0.1898	0.0058	45.7	199	4.4332E-55
CN1/8708	cytochrome c oxidase subunit II [EC:1.9.3.1]	1811789.29	0	0.1657	0.0000	6.8	294	0.00018547
CN1/19273	F-type H+-transporting ATPase subunit alpha [EC:3.6.3.14]	0	421878.62	0.0000	0.0132	16.3	283	5.2724E-20
CN1/5087	F-type H+-transporting ATPase subunit alpha [EC:3.6.3.14]	2228790.74	0	0.2038	0.0000	21.4	126	5.7956E-11
CN1/3429	F-type H+-transporting ATPase subunit beta [EC:3.6.3.14]	1796124.84	37136.07	0.1642	0.0012	38.3	300	7.061E-45
CN2/0987	cytochrome bd-I oxidase subunit I [EC:1.10.3.-]	0	4393969.67	0.0000	0.1371	7.6	527	1.1155E-09
CN2/14505	inorganic pyrophosphatase [EC:3.6.1.1]	0	1502503.02	0.0000	0.0469	20.6	175	1.5116E-29
CN2/15264	NADH dehydrogenase I subunit G [EC:1.6.5.3]	27432.08	1782756.76	0.0025	0.0556	30.9	909	2.5564E-172
CN2/16386	F-type H+-transporting ATPase subunit delta [EC:3.6.3.14]	0	5544872.09	0.0000	0.1729	52.2	178	2.462E-179
CN2/16387	F-type H+-transporting ATPase subunit b [EC:3.6.3.14]	64391.13	4234284.63	0.0059	0.1321	41	156	2.2271E-222
CN2/16411	cb-type cytochrome c oxidase subunit III [EC:1.9.3.1]	0	2222467.19	0.0000	0.0693	9.7	1481	2.2922E-46
CN2/17245	F-type H+-transporting ATPase subunit beta [EC:3.6.3.14]	36867.65	7127972.20	0.0034	0.2223	66.2	160	3.4413E-251
CN2/17246	F-type H+-transporting ATPase subunit gamma [EC:3.6.3.14]	32609.47	1903108.97	0.0030	0.0594	64.9	288	3.3483E-195
CN2/17247	F-type H+-transporting ATPase subunit alpha [EC:3.6.3.14]	137916.57	4933446.68	0.0126	0.1539	36.9	225	2.3297E-243
CN2/2277	cytochrome c oxidase subunit II [EC:1.9.3.1]	60294.95	2859601.16	0.0055	0.0892	28.1	310	2.8397E-41
CN2/2465	NADH dehydrogenase I subunit C/D [EC:1.6.5.3]	0	749310.98	0.0000	0.0234	16.1	466	9.4759E-21
CN2/2467	NADH dehydrogenase I subunit F [EC:1.6.5.3]	0	1225416.58	0.0000	0.0382	56.8	148	3.2676E-41
CN2/9218	ubiquinol-cytochrome c reductase iron-sulfur subunit [EC:1.10.2.2]	31962.23	0	0.0029	0.0000	16.8	197	8.4645E-10
CN2/15095	NADH dehydrogenase I subunit C/D [EC:1.6.5.3]	0	0	0.0000	0.0000	31.8	88	0.000035901
CN2/18058	F-type H+-transporting ATPase subunit beta [EC:3.6.3.14]	298822.94	9635496.63	0.0273	0.3005	74.2	298	3.0008E-258
CN2/9216	ubiquinol-cytochrome c reductase cytochrome c1 subunit [EC:1.10.2.2]	0	2680730.69	0.0000	0.0836	29	259	1.5804E-22
CN2/16385	F-type H+-transporting ATPase subunit alpha [EC:3.6.3.14]	144858.70	16336163.19	0.0132	0.5095	53.5	288	7.1856E-140
CN1/19272	ubiquinol-cytochrome c reductase cytochrome c1 subunit [EC:1.10.2.2]	2029865.27	0	0.1856	0.0000	33.6	223	1.8994E-15
CN1/19677	inorganic pyrophosphatase [EC:3.6.1.1]	587803.37	127480.07	0.0537	0.0040	21.6	185	1.004E-08
CN1/2866	cytochrome c oxidase subunit II [EC:1.9.3.1]	746910.47	400112.67	0.0683	0.0125	19.9	267	4.2526E-10
CN1/3054	F-type H+-transporting ATPase subunit alpha [EC:3.6.3.14]	3040705.15	427818.87	0.2780	0.0133	41.7	276	1.0999E-39
CN2/12832	phosphoribosylaminoimidazolecarboxamide formyltransferase / IMP	0	731389.80	0.0000	0.0228	24.9	538	5.0017E-141
CN2/16494	nitric-oxide reductase, cytochrome c-containing subunit II	0	8000691.41	0.0000	0.2495	24.2	289	3.131E-54
CN1/7633	periplasmic nitrate reductase NapA [EC:1.7.99.4]	830299.58	0	0.0759	0.0000	34.7	196	8.4377E-23
CN2/0033	carbonic anhydrase [EC:4.2.1.1]	0	3280854.85	0.0000	0.1023	32.7	214	1.2603E-63
CN2/1030	nitrate reductase 1, delta subunit [EC:1.7.99.4]	0	306706.33	0.0000	0.0096	18.5	248	6.6929E-27
CN2/13702	nitrous-oxide reductase [EC:1.7.99.6]	195631.05	9550605.71	0.0179	0.2979	49.5	637	0
CN2/16257	periplasmic nitrate reductase NapA [EC:1.7.99.4]	188973.04	1999608.04	0.0173	0.0624	19.2	834	1.3016E-43
CN2/15634	nitrate reductase 1, alpha subunit [EC:1.7.99.4]	365045.79	1751972.76	0.0334	0.0546	28.6	1251	4.8337E-152
CN2/1029	nitrate reductase 1, beta subunit [EC:1.7.99.4]	30668.14	6487692.13	0.0028	0.2024	20.6	501	2.3055E-92
CN2/13825	quinolinate synthase [EC:2.5.1.72]	0	755252.26	0.0000	0.0236	20.9	278	5.8841E-25
CN2/18256	NAD(P) transhydrogenase [EC:1.6.1.1]	0	415905.35	0.0000	0.0130	15.5	464	1.3858E-22
CN1/23998	Hydroxybenzylidenepyruvate hydratase-aldolase	631015.59	483001.62	0.0577	0.0151	50.4	226	7.14E-65

CN2/11942	Dihydroxynaphthalene dioxygenase	1234442.15	23507311.92	0.1129	0.7332	96	302	0
CN2/11940	Hydroxybenzylidenepyruvate hydratase-aldolase	241540.89	13599322.46	0.0221	0.4242	75.4	334	3.2435E-291
CN1/14979	Salicylaldehyde dehydrogenase	8155730.72	55435534.67	0.7457	1.7291	39	123	1.158E-113
CN2/11943	Salicylaldehyde dehydrogenase	2996104.28	31803689.65	0.2739	0.9920	69.4	483	0
CN1/18517	catalase/peroxidase [EC:1.11.1.6]	1010900.71	0	0.0924	0.0000	18.4	141	4.2938E-23
CN2/13685	peroxidase [EC:1.11.1.7]	14812.49	4660424.32	0.0014	0.1454	60.8	212	2.0364E-112
CN2/1480	methylenetetrahydrofolate reductase (NADPH) [EC:1.5.1.20]	0	721940.89	0.0000	0.0225	30	283	1.5485E-23
CN1/22889	carbon-monoxide dehydrogenase large subunit [EC:1.2.99.2]	222354.73	0	0.0203	0.0000	6.8	798	6.23E-11
CN2/12278	dihydrodipicolinate reductase [EC:1.3.1.26]	0	1985943.48	0.0000	0.0619	22.5	262	5.3237E-69
CN2/17435	succinyl-diaminopimelate desuccinylase [EC:3.5.1.18]	0	810478.74	0.0000	0.0253	13.2	393	5.0894E-06
CN2/17659	dihydrodipicolinate synthase [EC:4.2.1.52]	0	5630905.15	0.0000	0.1756	38.7	292	3.3187E-199
CN2/5124	2,3,4,5-tetrahydropyridine-2-carboxylate N-succinyltransferase	0	3641882.51	0.0000	0.1136	28.2	344	9.9913E-178
CN2/17458	UDP-N-acetylmuramoylalanyl-D-glutamyl-2,6-diaminopimelate--D-alanyl-	800507.53	2628485.85	0.0732	0.0820	57.5	457	8.6266E-113
CN2/17459	UDP-N-acetylmuramoylalanyl-D-glutamate--2,6-diaminopimelate ligase	265345.73	2373952.90	0.0243	0.0740	35.4	486	2.048E-66
CN1/23338	dihydrodipicolinate synthase [EC:4.2.1.52]	3462026.11	32381.51	0.3165	0.0010	14.8	291	5.3634E-12
CN2/0815	2-dehydro-3-deoxyphosphooctonate aldolase (KDO 8-P synthase)	44918.26	3929553.07	0.0041	0.1226	38.4	281	5.9795E-76
CN2/17728	starch phosphorylase [EC:2.4.1.1]	219693.06	1013998.05	0.0201	0.0316	17.7	817	8.1252E-62
CN2/0958	methylmalonate-semialdehyde dehydrogenase [EC:1.2.1.27]	212526.15	4489251.66	0.0194	0.1400	29.9	492	3.2189E-53
CN1/2863	methylmalonate-semialdehyde dehydrogenase [EC:1.2.1.27]	1701459.21	112143.00	0.1556	0.0035	42.5	355	1.8181E-30
CN1/3000	superoxide dismutase, Fe-Mn family [EC:1.15.1.1]	5531969.88	767791.03	0.5058	0.0239	30.7	163	2.5954E-301
CN2/0460	superoxide dismutase, Fe-Mn family [EC:1.15.1.1]	110037.02	35468312.70	0.0101	1.1063	49.2	193	5.2745E-59
CN2/16234	cyclase HisF [EC:4.1.3.-]	0	1288439.23	0.0000	0.0402	24.9	265	1.468E-23
CN2/18232	phosphoribosylformimino-5-aminoimidazole carboxamide ribotide	0	1668698.56	0.0000	0.0520	24	100	2.09E-08
CN2/4324	histidinol-phosphate aminotransferase [EC:2.6.1.9]	0	845639.29	0.0000	0.0264	8.8	354	0.00090427
CN2/4325	histidinol dehydrogenase [EC:1.1.1.23]	165841.00	1415793.46	0.0152	0.0442	27.8	435	6.402E-30
CN1/22034	histidinol dehydrogenase [EC:1.1.1.23]	587232.72	198057.19	0.0537	0.0062	21.9	434	2.1855E-60
CN1/16907	malate synthase [EC:2.3.3.9]	724969.91	25723.40	0.0663	0.0008	17.6	205	2.396E-12
CN1/19397	formate dehydrogenase, alpha subunit [EC:1.2.1.2]	236993.69	0	0.0217	0.0000	12.8	484	2.6944E-21
CN1/4315	malate synthase [EC:2.3.3.9]	591668.00	8057.51	0.0541	0.0003	21.5	540	5.9799E-188
CN2/12290	glycolate oxidase [EC:1.1.3.15]	0	1498060.54	0.0000	0.0467	7.7	300	1.4352E-06
CN2/14920	isocitrate lyase [EC:4.1.3.1]	21322.94	3114796.31	0.0019	0.0972	51.3	226	3.101E-122
CN2/16997	methylenetetrahydrofolate dehydrogenase (NADP+) /	0	1938005.74	0.0000	0.0604	15.8	284	6.601E-48
CN2/12329	malate synthase [EC:2.3.3.9]	30976.28	1047643.15	0.0028	0.0327	26.7	727	1.1904E-167
CN2/14951	isocitrate lyase [EC:4.1.3.1]	0	8459447.62	0.0000	0.2639	42.4	304	3.5833E-114
CN2/15317	malate synthase [EC:2.3.3.9]	88694.10	3573543.31	0.0081	0.1115	51.8	726	0
CN1/25584	acetyl-CoA synthetase [EC:6.2.1.1]	318221.02	15247812.11	0.0291	0.4756	15.3	176	4.0646E-28
CN2/11887	glyceraldehyde 3-phosphate dehydrogenase [EC:1.2.1.12]	224695.43	2316612.19	0.0205	0.0723	37.5	483	4.698E-164
CN2/14229	glyceraldehyde 3-phosphate dehydrogenase [EC:1.2.1.12]	0	3300209.78	0.0000	0.1029	15	320	1.8951E-09
CN2/16227	phosphoglycerate mutase [EC:5.4.2.1]	99214.77	935788.96	0.0091	0.0292	30.6	516	3.007E-111
CN2/2100	acetyl-CoA synthetase [EC:6.2.1.1]	0	7279863.58	0.0000	0.2271	15.4	651	9.4808E-55
CN2/0814	enolase [EC:4.2.1.11]	33831.96	5451376.01	0.0031	0.1700	48.3	429	0
CN1/3997	acetyl-CoA synthetase [EC:6.2.1.1]	5529707.07	288775.10	0.5056	0.0090	24.8	270	1.4373E-142
CN1/4744	acetyl-CoA synthetase [EC:6.2.1.1]	3203367.39	127810.61	0.2929	0.0040	51	210	4.5835E-41
CN1/19801	glyceraldehyde 3-phosphate dehydrogenase [EC:1.2.1.12]	7575395.36	796136.00	0.6926	0.0248	39.1	335	3.6604E-59
CN1/19928	acetyl-CoA synthetase [EC:6.2.1.1]	5575471.11	57023.22	0.5098	0.0018	28.8	66	1.6917E-09
CN1/6997	glyceraldehyde 3-phosphate dehydrogenase [EC:1.2.1.12]	85192.90	0	0.0078	0.0000	13	307	0.000020785
CN2/3095	aminomethyltransferase [EC:2.1.2.10]	0	2638268.62	0.0000	0.0823	37.7	371	4.8005E-163
CN2/14159	D-3-phosphoglycerate dehydrogenase [EC:1.1.1.95]	74151.73	11683297.10	0.0068	0.3644	53	411	2.955E-134
CN2/16771	phosphoserine aminotransferase [EC:2.6.1.52]	300694.16	901725.47	0.0275	0.0281	31.6	361	5.6174E-29

CN2/17146	threonine synthase [EC:4.2.3.1]	49003.53	3140973.85	0.0045	0.0980	51.1	474	7.3908E-126
CN2/2297	tryptophan synthase alpha chain [EC:4.2.1.20]	0	960762.46	0.0000	0.0300	42.9	273	3.7466E-125
CN2/3091	glycine dehydrogenase [EC:1.4.4.2]	539960.21	3136941.16	0.0494	0.0978	24.5	952	2.6088E-81
CN1/6765	aminomethyltransferase [EC:2.1.2.10]	701360.19	0	0.0641	0.0000	17.9	179	7.7675E-07
CN1/7239	aminomethyltransferase [EC:2.1.2.10]	802481.21	0	0.0734	0.0000	16.5	133	1.9588E-07
CN2/16776	aminopeptidase N [EC:3.4.11.2]	136261.36	1475133.35	0.0125	0.0460	16.9	886	2.8141E-40
CN1/12719	glutamate--cysteine ligase [EC:6.3.2.2]	1423926.97	0	0.1302	0.0000	17.8	236	5.1907E-24
CN1/3766	mannitol 2-dehydrogenase [EC:1.1.1.67]	561995.01	0	0.0514	0.0000	4.1	466	0.0018069
CN2/17013	acyl-CoA dehydrogenase [EC:1.3.99.-]	69315.30	5054304.10	0.0063	0.1576	31.5	815	3.6795E-190
CN2/11654	enoyl-[acyl-carrier protein] reductase I [EC:1.3.1.9]	1398778.22	5209627.58	0.1279	0.1625	39.8	264	1.0493E-129
CN2/17447	3-oxoacyl-[acyl-carrier-protein] synthase I [EC:2.3.1.41]	0	3709211.30	0.0000	0.1157	42	405	1.0749E-118
CN2/17503	acetyl-CoA carboxylase biotin carboxyl carrier protein	0	5954655.35	0.0000	0.1857	35.8	151	8.1056E-21
CN2/5127	3-oxoacyl-[acyl-carrier-protein] synthase III [EC:2.3.1.180]	0	1718176.40	0.0000	0.0536	30.8	373	2.6135E-84
CN2/5241	[acyl-carrier-protein] S-malonyltransferase [EC:2.3.1.39]	90799.70	3866203.90	0.0083	0.1206	25.3	312	1.2918E-90
CN2/0818	acetyl-CoA carboxylase carboxyl transferase subunit alpha	0	1811062.20	0.0000	0.0565	36.4	316	7.4442E-61
CN2/17504	acetyl-CoA carboxylase, biotin carboxylase subunit [EC:6.4.1.2]	299692.27	9060534.60	0.0274	0.2826	27.4	449	6.7482E-57
CN1/25901	dihydropyrimidinase [EC:3.5.2.2]	831763.98	0	0.0760	0.0000	25.1	167	7.9071E-08
CN2/14326	IMP dehydrogenase [EC:1.1.1.205]	293196.19	6440649.35	0.0268	0.2009	72	489	0
CN2/14327	GMP synthase (glutamine-hydrolysing) [EC:6.3.5.2]	0	2217723.59	0.0000	0.0692	34.2	526	3.1611E-111
CN1/16766	glutathione S-transferase [EC:2.5.1.18]	1308399.49	101123.69	0.1196	0.0032	36.3	157	6.7309E-37
CN1/24000	Glutathione S transferase (NagJ-like)	2203721.06	3557166.62	0.2015	0.1110	30.3	201	1.0159E-30
CN2/17347	Glutathione S transferase (NagJ-like)	52937.20	11297726.33	0.0048	0.3524	48.8	205	2.0009E-61
CN2/13275	Glutathione S transferase (NagJ-like)	0	1958371.59	0.0000	0.0611	16	225	0.00013461
CN2/1343	Glutathione S transferase (NagJ-like)	0	2953568.82	0.0000	0.0921	31.7	312	6.1375E-93
CN2/1590	Glutathione S transferase (NagJ-like)	0	2397152.76	0.0000	0.0748	17.3	225	2.9753E-07
CN2/16380	DNA polymerase III subunit beta [EC:2.7.7.7]	322468.33	5276987.18	0.0295	0.1646	26.4	367	1.1874E-20
CN2/17307	single-strand DNA-binding protein	0	6237766.48	0.0000	0.1946	29.3	164	1.5984E-10
CN1/3021	DNA polymerase III subunit beta [EC:2.7.7.7]	3897433.08	77593.52	0.3563	0.0024	45.2	270	6.0605E-66
CN2/12752	UDP-N-acetylmuramate--alanine ligase [EC:6.3.2.8]	0	2735569.60	0.0000	0.0853	5.6	486	1.7743E-13
CN2/12755	UDP-N-acetylmuramoylalanine--D-glutamate ligase [EC:6.3.2.9]	79657.48	1731342.77	0.0073	0.0540	33.6	447	4.7176E-75
CN2/12751	D-alanine-D-alanine ligase [EC:6.3.2.4]	0	2667039.39	0.0000	0.0832	22.3	314	2.7915E-65
CN1/8111	adenosylhomocysteinase [EC:3.3.1.1]	1593520.12	127924.27	0.1457	0.0040	5.9	353	0.00023466
CN1/10612	cysteine synthase A [EC:2.5.1.47]	4936466.17	89020.34	0.4513	0.0028	55.9	177	2.845E-97
CN2/0870	aspartate-semialdehyde dehydrogenase [EC:1.2.1.11]	0	1446018.98	0.0000	0.0451	53.2	370	6.9606E-209
CN2/0883	O-succinylhomoserine sulfhydrylase [EC:2.5.1.-]	0	2122617.57	0.0000	0.0662	14.6	403	9.969E-19
CN2/13271	cysteine synthase A [EC:2.5.1.47]	0	4502590.91	0.0000	0.1404	62.3	324	3.2661E-170
CN2/1337	5-methyltetrahydropteroyltriglutamate--homocysteine	0	2966420.32	0.0000	0.0925	18.3	778	4.2895E-75
CN2/14137	5-methyltetrahydrofolate--homocysteine methyltransferase	0	1559036.08	0.0000	0.0486	11	1080	2.9247E-43
CN2/1481	adenosylhomocysteinase [EC:3.3.1.1]	1665862.65	4940854.48	0.1523	0.1541	29.4	473	3.3927E-84
CN2/16284	aspartate kinase [EC:2.7.2.4]	0	1386296.42	0.0000	0.0432	9.5	412	5.2011E-83
CN2/0915	5-methyltetrahydrofolate--homocysteine methyltransferase	0	948392.95	0.0000	0.0296	29.7	101	5.2187E-06
CN2/11771	aromatic-amino-acid transaminase [EC:2.6.1.57]	52882.41	5187958.14	0.0048	0.1618	38.7	390	8.0017E-84
CN2/12482	cysteine synthase B [EC:2.5.1.47]	5943575.01	0	0.5434	0.0000	8.7	300	0.00007926
CN2/13259	serine O-acetyltransferase [EC:2.3.1.30]	0	2915187.86	0.0000	0.0909	30.9	259	8.6319E-19
CN2/1490	S-adenosylmethionine synthetase [EC:2.5.1.6]	222563.50	2785557.81	0.0203	0.0869	25	396	8.6772E-21
CN2/1839	O-acetylhomoserine (thiol)-lyase [EC:2.5.1.49]	138744.27	7111508.50	0.0127	0.2218	41.2	425	2.6841E-57
CN1/1790	O-acetylhomoserine (thiol)-lyase [EC:2.5.1.49]	1556036.55	38034.47	0.1423	0.0012	54.8	343	1.5274E-136
CN1/18701	adenosylhomocysteinase [EC:3.3.1.1]	904409.96	0	0.0827	0.0000	21.6	153	2.4545E-06
CN1/21714	S-adenosylmethionine synthetase [EC:2.5.1.6]	1961374.06	0	0.1793	0.0000	17.4	282	6.2901E-28

CN1/7796	cysteine synthase A [EC:2.5.1.47]	3264495.11	182339.24	0.2985	0.0057	39.2	319	1.8588E-69
CN1/11469	S-adenosylmethionine synthetase [EC:2.5.1.6]	1649069.64	0	0.1508	0.0000	20	105	0.000003604
CN2/0654	glycine hydroxymethyltransferase [EC:2.1.2.1]	72003.03	8757772.55	0.0066	0.2732	46.8	417	7.0514E-267
CN2/10886	beta-glucosidase [EC:3.2.1.21]	0	1693327.78	0.0000	0.0528	2.8	757	0.000059933
CN2/3093	glycine hydroxymethyltransferase [EC:2.1.2.1]	8615.14	2434353.03	0.0008	0.0759	23.7	417	5.7618E-25
CN1/11923	glycine hydroxymethyltransferase [EC:2.1.2.1]	452884.24	0	0.0414	0.0000	16.1	217	6.7212E-08
CN2/14904	aconitate hydratase 1 [EC:4.2.1.3]	0	1413641.61	0.0000	0.0441	12.3	454	1.4094E-13
CN1/19403	dihydrolipoamide dehydrogenase [EC:1.8.1.4]	1092848.02	130356.37	0.0999	0.0041	48.5	66	2.4542E-10
CN1/13053	isocitrate dehydrogenase [EC:1.1.1.42]	4093605.83	221650.98	0.3743	0.0069	38.7	155	1.1925E-67
CN1/21366	dihydrolipoamide dehydrogenase [EC:1.8.1.4]	494438.98	0	0.0452	0.0000	17.1	216	8.1783E-06
CN1/13051	isocitrate dehydrogenase [EC:1.1.1.42]	2611354.97	1237163.72	0.2388	0.0386	37.2	282	1.26E-25
CN1/1802	dihydrolipoamide dehydrogenase [EC:1.8.1.4]	1834810.91	30027.69	0.1678	0.0009	20.3	399	9.4131E-18
CN2/0972	2-oxoglutarate dehydrogenase E1 component [EC:1.2.4.2]	161062.75	8884504.43	0.0147	0.2771	45.1	943	0
CN2/17404	aconitate hydratase 1 [EC:4.2.1.3]	72510.31	3040951.62	0.0066	0.0948	29.8	312	2.693E-96
CN2/17513	aconitate hydratase 2 [EC:4.2.1.3]	10624717.02	18010515.82	0.9714	0.5618	46.4	323	2.7298E-135
CN2/17817	fumarate hydratase, class II [EC:4.2.1.2]	0	14943375.10	0.0000	0.4661	66	150	8.7237E-33
CN2/2192	aconitate hydratase 1 [EC:4.2.1.3]	53126.73	5281130.74	0.0049	0.1647	48.1	351	4.1008E-264
CN2/7221	fumarate hydratase, class I [EC:4.2.1.2]	93689.27	1980595.42	0.0086	0.0618	34.3	507	1.4495E-127
CN2/0970	dihydrolipoamide dehydrogenase [EC:1.8.1.4]	148043.82	9362640.85	0.0135	0.2920	61.7	478	8.9434E-280
CN2/0971	2-oxoglutarate dehydrogenase E2 component (dihydrolipoamide	67457.66	18048128.27	0.0062	0.5629	64.6	404	0
CN2/0977	citrate synthase [EC:2.3.3.1]	105537.65	11735594.84	0.0096	0.3660	44.7	423	1.0127E-185
CN2/10478	pyruvate dehydrogenase E2 component (dihydrolipoamide	346944.68	19113646.79	0.0317	0.5962	42.2	552	0
CN2/16406	aconitate hydratase 1 [EC:4.2.1.3]	10394.38	3315984.69	0.0010	0.1034	74.3	226	8.3163E-173
CN2/17007	aconitate hydratase 2 [EC:4.2.1.3]	117847.15	6844721.39	0.0108	0.2135	70.2	262	1.3876E-233
CN2/17464	fumarate hydratase, class II [EC:4.2.1.2]	0	4455059.26	0.0000	0.1390	32	153	1.7191E-19
CN2/14871	pyruvate carboxylase subunit B [EC:6.4.1.1]	172892.01	1281210.62	0.0158	0.0400	40.1	272	1.9055E-57
CN2/14929	isocitrate dehydrogenase [EC:1.1.1.42]	49113.12	4821914.35	0.0045	0.1504	37.7	741	3.2818E-230
CN2/14930	isocitrate dehydrogenase [EC:1.1.1.42]	28826.30	1789463.05	0.0026	0.0558	23	418	4.0557E-30
CN2/17523	aconitate hydratase 1 [EC:4.2.1.3]	618471.07	3290253.41	0.0565	0.1026	47.5	869	1.019E-207
CN2/16590	aconitate hydratase 2 [EC:4.2.1.3]	101179.79	8380180.91	0.0093	0.2614	44.9	256	5.6258E-68
CN1/0968	aconitate hydratase 1 [EC:4.2.1.3]	705712.94	63467.53	0.0645	0.0020	21.8	229	8.4027E-24
CN1/13809	fumarate hydratase, class II [EC:4.2.1.2]	1095279.05	0	0.1001	0.0000	11.3	335	2.3334E-14
CN1/18275	aconitate hydratase 1 [EC:4.2.1.3]	3773649.29	206403.13	0.3450	0.0064	22.3	461	7.3453E-76
CN1/13955	aconitate hydratase 1 [EC:4.2.1.3]	2636314.40	333221.49	0.2410	0.0104	59.7	149	2.5704E-14
CN1/25093	malate dehydrogenase [EC:1.1.1.37]	3776862.43	189581.76	0.3453	0.0059	53.5	71	2.2497E-46
CN1/18959	fructose-bisphosphate aldolase, class II [EC:4.1.2.13]	0	0	0.0000	0.0000	14.9	175	1.2422E-06
CN1/21183	malate dehydrogenase [EC:1.1.1.37]	3143848.62	1579234.03	0.2874	0.0493	27.8	108	3.8974E-21
CN1/24031	malate dehydrogenase [EC:1.1.1.37]	14386933.41	105677.96	1.3154	0.0033	32.5	154	1.0219E-44
CN2/12319	ribose 5-phosphate isomerase A [EC:5.3.1.6]	0	0	0.0000	0.0000	30.9	223	1.2577E-78
CN2/14560	phosphoenolpyruvate carboxylase [EC:4.1.1.31]	720921.10	1413507.46	0.0659	0.0441	21.6	878	1.5226E-71
CN2/17711	phosphoenolpyruvate carboxykinase (ATP) [EC:4.1.1.49]	198600.71	2141019.08	0.0182	0.0668	48.1	514	4.4167E-303
CN2/17729	fructose-1,6-bisphosphatase I [EC:3.1.3.11]	0	690976.80	0.0000	0.0216	16.7	336	8.1952E-16
CN2/18112	phosphoglycerate kinase [EC:2.7.2.3]	0	7387651.30	0.0000	0.2304	36.2	387	5.8044E-112
CN2/18115	fructose-bisphosphate aldolase, class II [EC:4.1.2.13]	423646.38	7390508.22	0.0387	0.2305	51.4	354	3.3727E-162
CN2/17627	malate dehydrogenase (oxaloacetate-decarboxylating)(NADP+)	32020.44	9240644.95	0.0029	0.2882	64	422	3.4936E-271
CN2/7218	pyruvate kinase [EC:2.7.1.40]	0	4522699.71	0.0000	0.1411	51.7	484	4.2745E-231
CN1/12831	triosephosphate isomerase (TIM) [EC:5.3.1.1]	283524.07	0	0.0259	0.0000	6.6	211	2.7903E-07
CN1/14462	malate dehydrogenase (oxaloacetate-decarboxylating)(NADP+)	1712860.53	122943.90	0.1566	0.0038	38.4	770	3.5893E-196
CN1/24033	malate dehydrogenase [EC:1.1.1.37]	3078558.71	0	0.2815	0.0000	58.8	136	4.9693E-66

CN2/13553	alcohol dehydrogenase (NADP+) [EC:1.1.1.2]	0	4613253.35	0.0000	0.1439	27.2	323	3.5023E-24
CN1/25264	succinyl-CoA synthetase beta subunit [EC:6.2.1.5]	3494198.79	0	0.3195	0.0000	51.2	43	2.0881E-07
CN1/21663	succinyl-CoA synthetase beta subunit [EC:6.2.1.5]	6118193.98	142890.80	0.5594	0.0045	22.3	319	4.999E-33
CN2/0868	3-isopropylmalate/(R)-2-methylmalate dehydratase small subunit	0	2191259.32	0.0000	0.0683	29.3	215	4.0804E-15
CN2/0968	succinyl-CoA synthetase alpha subunit [EC:6.2.1.5]	161248.55	13803927.93	0.0147	0.4306	40	295	5.4795E-80
CN2/0969	succinyl-CoA synthetase beta subunit [EC:6.2.1.5]	207786.98	7227720.60	0.0190	0.2254	59.5	388	2.4083E-291
CN2/10479	pyruvate dehydrogenase E1 component [EC:1.2.4.1]	88378.33	5337308.23	0.0081	0.1665	35.9	896	1.2555E-280
CN1/3474	3-hydroxybutyrate dehydrogenase [EC:1.1.1.30]	707168.69	0	0.0647	0.0000	14.9	168	0.000048379
CN2/13999	3-oxoacid CoA-transferase subunit A [EC:2.8.3.5]	0	1402092.25	0.0000	0.0437	13.9	231	1.4079E-06
CN2/17488	acetolactate synthase I/II/III large subunit [EC:2.2.1.6]	0	535346.44	0.0000	0.0167	11.1	577	2.7468E-21
CN2/17368	Acetaldehyde dehydrogenase	89738.28	9410098.11	0.0082	0.2935	59	307	1.2421E-169
CN1/4225	formate C-acetyltransferase [EC:2.3.1.54]	255681.09	0	0.0234	0.0000	15.8	461	1.2572E-08
CN1/6621	acetoacetyl-CoA reductase [EC:1.1.1.36]	8386309.35	534880.24	0.7668	0.0167	49.2	240	4.8195E-181
CN1/7282	pyruvate dehydrogenase E1 component subunit alpha [EC:1.2.4.1]	0	80271.83	0.0000	0.0025	7.2	277	5.3259E-07
CN2/1650	aryl-alcohol dehydrogenase [EC:1.1.1.90]	3098233.53	0	0.2833	0.0000	7.1	184	2.3868E-11
CN1/14978	Dihydroxynaphthalene dioxygenase	23338452.23	15439374.45	2.1339	0.4816	65.7	134	2.9911E-263
CN2/8417	Dihydroxynaphthalene dioxygenase	6074888.68	3612052.85	0.5554	0.1127	58.1	155	1.4899E-112
CN2/3403	Unknown function	31453.38	1519951.94	0.0029	0.0474	55.1	227	5.0311E-50
CN2/12416	3-hydroxyacyl-CoA dehydrogenase / enoyl-CoA hydratase /	279698.22	4048051.36	0.0256	0.1263	20.6	204	1.4674E-24
CN2/15643	acyl-CoA thioesterase II [EC:3.1.2.-]	0	671545.78	0.0000	0.0209	38	255	5.6471E-31
CN2/15982	3-hydroxyacyl-CoA dehydrogenase / enoyl-CoA hydratase /	31257.82	3044033.65	0.0029	0.0949	36.9	244	4.1192E-111
CN2/5240	3-oxoacyl-[acyl-carrier protein] reductase [EC:1.1.1.100]	0	6429008.48	0.0000	0.2005	36.3	248	7.7308E-127
CN2/14184	isochorismatase [EC:3.3.2.1]	0	9261490.92	0.0000	0.2889	17.7	198	2.1001E-33
CN2/1495	transketolase [EC:2.2.1.1]	71750.83	2193592.63	0.0066	0.0684	26.9	665	4.2729E-57
CN2/8952	2-Oxopent-4-enoate hydratase	0	3096663.18	0.0000	0.0966	17.6	159	7.1903E-14
CN2/12417	acetyl-CoA acyltransferase [EC:2.3.1.16]	0	4883582.81	0.0000	0.1523	28.9	391	1.2339E-109
CN2/17598	3-oxoadipate enol-lactonase [EC:3.1.1.24]	0	1424591.65	0.0000	0.0444	8.3	264	1.576E-10
CN1/4142	enoyl-CoA hydratase [EC:4.2.1.17]	3265155.38	32780.13	0.2985	0.0010	57.1	126	5.0842E-20
CN2/0959	enoyl-CoA hydratase [EC:4.2.1.17]	0	1551348.53	0.0000	0.0484	26.5	257	4.5243E-36
CN2/0973	succinate dehydrogenase iron-sulfur protein [EC:1.3.99.1]	109456.85	20327885.92	0.0100	0.6340	61.5	234	6.0427E-54
CN2/0975	succinate dehydrogenase hydrophobic membrane anchor protein	0	438761.94	0.0000	0.0137	15.6	122	0.00023688
CN2/11642	enoyl-CoA hydratase [EC:4.2.1.17]	614185.66	1455472.80	0.0562	0.0454	13	285	8.5222E-121
CN2/13997	acetyl-CoA C-acetyltransferase [EC:2.3.1.9]	0	4136944.38	0.0000	0.1290	18.5	406	3.1757E-30
CN2/17132	enoyl-CoA hydratase [EC:4.2.1.17]	0	1898609.72	0.0000	0.0592	12.4	282	3.1936E-08
CN2/0974	succinate dehydrogenase flavoprotein subunit [EC:1.3.99.1]	79282.28	14746303.40	0.0072	0.4599	41.2	590	4.6101E-268
CN1/4220	acetyl-CoA C-acetyltransferase [EC:2.3.1.9]	2195120.70	0	0.2007	0.0000	17.8	315	5.0676E-26
CN1/8918	enoyl-CoA hydratase [EC:4.2.1.17]	3180440.50	0	0.2908	0.0000	25.6	121	1.3861E-08
CN2/17714	succinate dehydrogenase iron-sulfur protein [EC:1.3.99.1]	0	9449057.33	0.0000	0.2947	59.1	235	1.7183E-50
CN2/15239	exodeoxyribonuclease III [EC:3.1.11.2]	0	652870.89	0.0000	0.0204	14.4	270	6.8878E-06
CN1/19418	type VI secretion system secreted protein VgrG	78382.17	0	0.0072	0.0000	15	180	1.017E-17
CN1/22816	type VI secretion system secreted protein Hcp	12726625.62	0	1.1636	0.0000	15.9	164	1.862E-48
CN2/3581	outer membrane channel protein TolC	0	755337.14	0.0000	0.0236	39	482	3.9874E-229
CN2/16142	preprotein translocase subunit SecA	0	663389.90	0.0000	0.0207	8	892	2.8071E-62
CN1/4558	outer membrane channel protein TolC	2133645.51	786793.67	0.1951	0.0245	31.3	166	6.7679E-56
CN1/5344	outer membrane channel protein TolC	1869566.36	0	0.1709	0.0000	19.6	158	1.4844E-12
CN2/16707	methyl-accepting chemotaxis protein	242317.67	537751.83	0.0222	0.0168	28.4	676	2.5796E-181
CN2/16143	glutamate N-acetyltransferase / amino-acid N-acetyltransferase	0	6149150.88	0.0000	0.1918	23	405	3.0302E-73
CN2/17733	proline iminopeptidase [EC:3.4.11.5]	0	1164993.48	0.0000	0.0363	33	324	1.8471E-52
CN2/17863	4-guanidinobutyraldehyde dehydrogenase / NAD-dependent aldehyde	0	816391.98	0.0000	0.0255	17.2	151	1.9083E-07

CN2/17864	4-guanidinobutyraldehyde dehydrogenase / NAD-dependent aldehyde	0	5849430.99	0.0000	0.1824	44.9	345	2.2371E-116
CN2/17877	N-carbamoylputrescine amidase [EC:3.5.1.53]	0	6402712.66	0.0000	0.1997	36.5	293	2.8206E-40
CN2/17941	arginine deiminase [EC:3.5.3.6]	0	1070447.04	0.0000	0.0334	22.7	418	7.9467E-46
CN2/18007	glutamate-5-semialdehyde dehydrogenase [EC:1.2.1.41]	52069.98	2906341.11	0.0048	0.0907	18.6	457	8.9763E-33
CN2/2110	acetylornithine/N-succinyldiaminopimelate aminotransferase	0	1044078.79	0.0000	0.0326	29.5	308	6.0275E-16
CN2/3423	ornithine decarboxylase [EC:4.1.1.17]	0	2394362.50	0.0000	0.0747	20.4	387	1.1099E-64
CN2/3434	arginine decarboxylase [EC:4.1.1.19]	0	1684943.28	0.0000	0.0526	3.5	636	0.00089611
CN2/3591	acetylglutamate kinase [EC:2.7.2.8]	0	3180196.10	0.0000	0.0992	21	300	1.1444E-10
CN2/5119	acetylornithine/N-succinyldiaminopimelate aminotransferase	0	503174.31	0.0000	0.0157	14.1	398	7.0051E-13
CN2/5598	oxaloacetate decarboxylase, alpha subunit [EC:4.1.1.3]	0	1370414.91	0.0000	0.0427	44.2	330	5.7427E-183
CN2/16840	proline dehydrogenase [EC:1.5.99.8]	0	1933399.47	0.0000	0.0603	9.7	1053	2.3235E-30
CN2/17213	N-acetyl-gamma-glutamyl-phosphate reductase [EC:1.2.1.38]	0	4497218.50	0.0000	0.1403	52.6	365	1.0955E-77
CN1/1577	glutamate N-acetyltransferase / amino-acid N-acetyltransferase	630882.76	1554502.96	0.0577	0.0485	25	412	3.2996E-75
CN2/0963	molecular chaperone HtpG	379962.14	2084161.02	0.0347	0.0650	39.9	634	3.2239E-194
CN2/17310	catalase [EC:1.11.1.6]	0	598770.83	0.0000	0.0187	29.8	486	1.8746E-155
CN2/1390	Cu/Zn superoxide dismutase [EC:1.15.1.1]	0	4648075.06	0.0000	0.1450	50.9	175	8.5031E-66
CN2/15531	catalase [EC:1.11.1.6]	142603.37	2289376.68	0.0130	0.0714	27.9	616	6.4017E-37
CN2/17268	glycyl-tRNA synthetase alpha chain [EC:6.1.1.14]	0	1884555.97	0.0000	0.0588	11.4	308	8.7585E-61
CN2/13325	lysyl-tRNA synthetase, class II [EC:6.1.1.6]	0	1318207.94	0.0000	0.0411	47.6	500	3.6739E-262
CN2/12037	phenylalanyl-tRNA synthetase beta chain [EC:6.1.1.20]	0	664790.74	0.0000	0.0207	13.4	792	1.8818E-23
CN2/13958	aspartyl-tRNA(Asn)/glutamyl-tRNA (Gln) amidotransferase subunit B	64962.55	1705040.47	0.0059	0.0532	23.4	482	3.1387E-54
CN2/13959	aspartyl-tRNA(Asn)/glutamyl-tRNA (Gln) amidotransferase subunit A	18930.31	2734473.62	0.0017	0.0853	36	483	1.7335E-110
CN2/16001	isoleucyl-tRNA synthetase [EC:6.1.1.5]	15408.29	880210.69	0.0014	0.0275	24.5	943	3.3595E-153
CN2/16523	prolyl-tRNA synthetase [EC:6.1.1.15]	20271.26	3357764.58	0.0019	0.1047	26.8	571	5.1439E-212
CN2/16998	cysteinyl-tRNA synthetase [EC:6.1.1.16]	0	3079219.15	0.0000	0.0960	28.9	461	1.1908E-60
CN2/16999	glutaminyl-tRNA synthetase [EC:6.1.1.18]	0	1260515.12	0.0000	0.0393	25.7	556	1.2709E-131
CN2/17990	leucyl-tRNA synthetase [EC:6.1.1.4]	405371.96	1790393.53	0.0371	0.0558	24.1	868	1.9649E-115
CN2/18104	methionyl-tRNA formyltransferase [EC:2.1.2.9]	0	2266379.30	0.0000	0.0707	26.9	334	1.5027E-30
CN2/1354	valyl-tRNA synthetase [EC:6.1.1.9]	69895.12	1670667.95	0.0064	0.0521	15.6	788	1.1532E-109
CN2/14943	seryl-tRNA synthetase [EC:6.1.1.11]	0	1187629.30	0.0000	0.0370	25.4	426	3.9056E-30
CN2/16517	aspartyl-tRNA synthetase [EC:6.1.1.12]	2247190.27	3081152.10	0.2055	0.0961	15.6	591	7.8707E-49
CN2/17217	tyrosyl-tRNA synthetase [EC:6.1.1.1]	0	1044043.45	0.0000	0.0326	5.1	392	0.00040943
CN2/16283	alanyl-tRNA synthetase [EC:6.1.1.7]	47485.57	3047214.85	0.0043	0.0950	40.3	874	2.415E-196
CN1/22592	aspartyl-tRNA synthetase [EC:6.1.1.12]	1770451.77	0	0.1619	0.0000	29.6	81	2.4858E-08
CN1/25868	aspartyl-tRNA synthetase [EC:6.1.1.12]	3023714.70	351339.49	0.2765	0.0110	14.1	163	1.4172E-17
CN2/11948	Naphthalene dioxygenase (reductase)	601598.83	3134127.08	0.0550	0.0978	51.8	328	1.1705E-186
CN2/18056	bifunctional protein GlmU [EC:2.7.7.23 2.3.1.157]	89218.43	910555.89	0.0082	0.0284	20	454	2.2388E-41
CN2/3592	phosphomannomutase [EC:5.4.2.8]	0	797705.12	0.0000	0.0249	3.9	852	0.000018105
CN2/18130	UTP--glucose-1-phosphate uridylyltransferase [EC:2.7.7.9]	0	2356763.34	0.0000	0.0735	49.6	125	2.2115E-72
CN1/20621	glucose-6-phosphate isomerase [EC:5.3.1.9]	941013.95	0	0.0860	0.0000	15.3	550	1.3567E-38
CN1/24245	cytochrome c	13367895.24	267197.59	1.2222	0.0083	53.2	126	2.4094E-76
CN1/2015	cytochrome c	556785.11	0	0.0509	0.0000	27.9	179	7.6828E-34
CN1/14736	glutamin-(asparagin-)ase [EC:3.5.1.38]	6836704.36	147946.15	0.6251	0.0046	26.2	225	2.0445E-47
CN1/21047	glutamine synthetase [EC:6.3.1.2]	6098389.77	4296700.01	0.5576	0.1340	29.9	174	2.0003E-50
CN1/21428	carbamoyl-phosphate synthase large subunit [EC:6.3.5.5]	219936.95	0	0.0201	0.0000	15.4	169	1.3983E-06
CN2/12079	aspartate carbamoyltransferase catalytic subunit [EC:2.1.3.2]	0	2421959.42	0.0000	0.0755	25.7	338	1.7712E-44
CN2/12277	carbamoyl-phosphate synthase small subunit [EC:6.3.5.5]	0	1851297.86	0.0000	0.0577	29.1	382	1.0796E-74
CN2/1281	glutamate dehydrogenase (NADP+) [EC:1.4.1.4]	0	1887303.27	0.0000	0.0589	21.1	446	2.2659E-30
CN2/12908	glutamine synthetase [EC:6.3.1.2]	219052.07	17423400.73	0.0200	0.5435	66	468	0

CN2/17142	argininosuccinate synthase [EC:6.3.4.5]	94661.31	3483043.74	0.0087	0.1086	56.4	406	2.8946E-230
CN2/1830	glutamate synthase (NADPH/NADH) small chain [EC:1.4.1.13 1.4.1.14]	65987.04	2709832.84	0.0060	0.0845	35.8	472	1.2672E-81
CN2/2920	argininosuccinate lyase [EC:4.3.2.1]	35038.25	1408955.72	0.0032	0.0439	22.8	464	6.6623E-49
CN2/14925	adenylosuccinate lyase [EC:4.3.2.2]	131772.41	4116467.54	0.0120	0.1284	44.5	456	3.9914E-164
CN2/3977	succinate-semialdehyde dehydrogenase (NADP+) [EC:1.2.1.16]	0	2653360.14	0.0000	0.0828	51.9	489	0
CN2/0882	amidophosphoribosyltransferase [EC:2.4.2.14]	144563.20	2890462.10	0.0132	0.0902	19.8	501	1.2099E-82
CN2/15721	glutamate dehydrogenase [EC:1.4.1.2]	365709.46	757680.09	0.0334	0.0236	20.1	1624	3.3872E-173
CN2/1829	glutamate synthase (NADPH/NADH) large chain [EC:1.4.1.13 1.4.1.14]	150445.12	4743798.76	0.0138	0.1480	41.3	1481	0
CN2/14479	adenylosuccinate synthase [EC:6.3.4.4]	225640.61	7275481.36	0.0206	0.2269	41.5	431	1.4418E-255
CN2/12275	carbamoyl-phosphate synthase large subunit [EC:6.3.5.5]	130652.13	1371993.73	0.0119	0.0428	29.9	1073	4.709E-147
CN1/21702	argininosuccinate synthase [EC:6.3.4.5]	2137245.50	614466.90	0.1954	0.0192	42.5	261	1.8297E-32
CN1/3065	argininosuccinate synthase [EC:6.3.4.5]	1412729.59	0	0.1292	0.0000	22.7	150	2.135E-98
CN1/4637	glutamate synthase (NADPH/NADH) small chain [EC:1.4.1.13 1.4.1.14]	343666.71	0	0.0314	0.0000	11.3	302	6.9936E-06
CN2/14982	long-chain acyl-CoA synthetase [EC:6.2.1.3]	0	542631.58	0.0000	0.0169	5	1203	5.2686E-08
CN1/22984	D-xylose transport system substrate-binding protein	8064629.65	33618.27	0.7374	0.0010	74.2	93	1.1582E-49
CN1/10124	branched-chain amino acid transport system substrate-binding protein	18191840.47	597482.53	1.6633	0.0186	59.1	215	5.0407E-188
CN1/18625	branched-chain amino acid transport system substrate-binding protein	12387757.31	343543.66	1.1326	0.0107	48.8	172	6.0629E-76
CN1/2180	branched-chain amino acid transport system substrate-binding protein	3277079.49	0	0.2996	0.0000	25.8	163	1.8144E-07
CN1/2242	sulfonate/nitrate/taurine transport system substrate-binding protein	4451439.71	41424.15	0.4070	0.0013	50	342	2.2555E-85
CN1/22743	sn-glycerol 3-phosphate transport system substrate-binding protein	3379423.13	146157.62	0.3090	0.0046	43.8	208	9.1332E-28
CN1/26125	peptide/nickel transport system substrate-binding protein	740758.37	0	0.0677	0.0000	10.3	243	0.00001215
CN1/6691	D-methionine transport system substrate-binding protein	2522189.55	151984.59	0.2306	0.0047	34.4	180	5.8735E-15
CN1/0120	sulfonate/nitrate/taurine transport system substrate-binding protein	2379103.89	0	0.2175	0.0000	21.5	181	2.1886E-28
CN1/9988	general L-amino acid transport system substrate-binding protein	7957550.62	0	0.7276	0.0000	30.4	115	6.567E-09
CN1/0141	peptide/nickel transport system substrate-binding protein	454535.40	0	0.0416	0.0000	5.3	359	0.000010824
CN1/16861	ribose transport system substrate-binding protein	9734707.81	326439.48	0.8901	0.0102	42.6	331	1.5317E-108
CN1/9446	branched-chain amino acid transport system substrate-binding protein	540094.61	0	0.0494	0.0000	18.5	378	1.4068E-67
CN1/12166	D-methionine transport system substrate-binding protein	2640753.47	371537.32	0.2414	0.0116	45.6	261	1.8467E-65
CN1/13135	glutamate/aspartate transport system substrate-binding protein	8900969.61	1518272.26	0.8138	0.0474	57.1	156	1.6174E-62
CN1/13433	branched-chain amino acid transport system substrate-binding protein	1404191.46	0	0.1284	0.0000	24.6	350	5.617E-171
CN1/4757	branched-chain amino acid transport system substrate-binding protein	7370127.85	236592.24	0.6739	0.0074	51.7	116	2.2709E-206
CN1/5921	branched-chain amino acid transport system substrate-binding protein	402846.47	0	0.0368	0.0000	9.8	204	0.00033504
CN1/21409	branched-chain amino acid transport system substrate-binding protein	4262096.96	276316.03	0.3897	0.0086	33.8	364	4.6512E-35
CN1/5845	branched-chain amino acid transport system substrate-binding protein	1359366.90	73956.21	0.1243	0.0023	35.1	299	6.7327E-70
CN1/11527	octopine/nopaline transport system substrate-binding protein	4471216.15	0	0.4088	0.0000	39.9	163	1.7955E-36
CN1/22156	peptide/nickel transport system substrate-binding protein	24890059.22	0	2.2757	0.0000	28.3	106	6.4672E-07
CN1/23313	peptide/nickel transport system substrate-binding protein	765022.02	0	0.0699	0.0000	17.3	150	1.7816E-13
CN1/14263	branched-chain amino acid transport system substrate-binding protein	2650037.38	0	0.2423	0.0000	21.4	308	3.3666E-14
CN1/24315	dipeptide transport system substrate-binding protein	2083212.52	136907.07	0.1905	0.0043	42.1	534	4.2949E-135
CN1/5273	branched-chain amino acid transport system substrate-binding protein	6754691.31	145481.59	0.6176	0.0045	32.8	174	9.0559E-64
CN1/0701	putative tungstate transport system substrate-binding protein	11004122.75	345593.59	1.0061	0.0108	69.2	237	5.9011E-230
CN2/11650	peptide/nickel transport system substrate-binding protein	0	544222.27	0.0000	0.0170	4.5	1211	2.1956E-07
CN2/12653	branched-chain amino acid transport system substrate-binding protein	0	16742148.72	0.0000	0.5222	36.8	182	6.6148E-76
CN2/15527	D-methionine transport system substrate-binding protein	0	12366281.33	0.0000	0.3857	76.3	257	0
CN2/16212	iron(III) transport system substrate-binding protein	68801.45	2840813.28	0.0063	0.0886	48.3	333	1.2714E-79
CN2/17881	putrescine transport system substrate-binding protein	103519.95	12505848.36	0.0095	0.3901	50.4	802	0
CN2/18286	branched-chain amino acid transport system substrate-binding protein	837069.69	49773010.46	0.0765	1.5525	71.2	372	0
CN2/18290	phosphonate transport system substrate-binding protein	70320.20	6969933.08	0.0064	0.2174	73.5	283	1.2315E-220
CN2/18340	phosphate transport system substrate-binding protein	0	3053573.38	0.0000	0.0952	36.4	321	2.4876E-39

CN2/5143	molybdate transport system substrate-binding protein	1138578.17	2451328.07	0.1041	0.0765	54.4	250	1.8659E-87
CN2/15829	branched-chain amino acid transport system substrate-binding protein	101328.77	18996109.78	0.0093	0.5925	48.9	219	1.9643E-194
CN2/17678	general L-amino acid transport system substrate-binding protein	176385.45	7846623.33	0.0161	0.2447	48.2	342	4.1009E-212
CN2/3889	sulfonate/nitrate/taurine transport system substrate-binding protein	0	1620799.88	0.0000	0.0506	23.4	329	1.0643E-58
CN1/1342	molybdate transport system substrate-binding protein	11134496.27	1172180.74	1.0180	0.0366	57.9	273	1.5461E-112
CN1/1589	peptide/nickel transport system substrate-binding protein	1179519.30	0	0.1078	0.0000	23.4	474	8.4064E-39
CN1/20852	branched-chain amino acid transport system substrate-binding protein	3361638.82	74522.21	0.3074	0.0023	19.3	394	1.6454E-16
CN1/4572	D-xylose transport system substrate-binding protein	11013425.44	3239988.14	1.0070	0.1011	57.1	205	5.1247E-104
CN1/0472	peptide/nickel transport system substrate-binding protein	711038.04	0	0.0650	0.0000	16.1	397	2.9649E-24
CN1/18589	putrescine transport system substrate-binding protein	19070026.46	0	1.7436	0.0000	30.9	94	8.1756E-23
CN1/20751	general L-amino acid transport system substrate-binding protein	1350498.01	0	0.1235	0.0000	40	110	4.7088E-20
CN1/2830	putrescine transport system substrate-binding protein	8866116.49	926542.16	0.8106	0.0289	43.5	214	3.0373E-160
CN1/7821	branched-chain amino acid transport system substrate-binding protein	4775812.11	231394.14	0.4367	0.0072	34.1	264	5.6506E-32
CN1/20323	peptide/nickel transport system substrate-binding protein	817603.27	0	0.0748	0.0000	13.2	538	4.9969E-192
CN1/11488	iron(III) transport system substrate-binding protein	1009222.95	0	0.0923	0.0000	7.5	332	2.3277E-06
CN1/8815	putative multiple sugar transport system substrate-binding protein	569301.82	1418919.12	0.0521	0.0443	13.4	366	1.0231E-14
CN2/17371	Catechol 2,3-dioxygenase	0	2237812.86	0.0000	0.0698	11.7	307	4.4192E-19
CN1/25671	carboxymethylenebutenolidase [EC:3.1.1.45]	2951096.49	71052.13	0.2698	0.0022	33.8	231	3.8947E-58
CN1/24538	carboxymethylenebutenolidase [EC:3.1.1.45]	1207741.59	0	0.1104	0.0000	39.5	76	1.0935E-07
CN2/17367	4-Hydroxy-2-oxovalerate aldolase	28999.24	17964462.87	0.0027	0.5603	44.2	344	3.7332E-134
CN2/1894	carboxymethylenebutenolidase [EC:3.1.1.45]	0	3559015.19	0.0000	0.1110	50.7	268	6.8874E-67
CN1/0954	carboxymethylenebutenolidase [EC:3.1.1.45]	5480025.36	79879.07	0.5010	0.0025	50.9	285	4.4969E-175
CN2/14496	aldehyde dehydrogenase (NAD+) [EC:1.2.1.3]	79441.68	12615144.30	0.0073	0.3935	60.5	506	0
CN2/17957	alcohol dehydrogenase (acceptor) [EC:1.1.99.8]	115632.31	4710539.71	0.0106	0.1469	41.2	590	3.0315E-185
CN2/13496	aldehyde dehydrogenase (NAD+) [EC:1.2.1.3]	0	2371610.87	0.0000	0.0740	75	88	1.8897E-145
CN2/0266	aldehyde dehydrogenase (NAD+) [EC:1.2.1.3]	5903.87	455819.23	0.0005	0.0142	44.9	321	6.315E-266
CN1/21617	Unknown function	2325564.63	225870.52	0.2126	0.0070	35.8	193	9.759E-12
CN2/16569	Unknown function	105446.45	9165046.53	0.0096	0.2859	69	503	0
CN2/0810	S-(hydroxymethyl)glutathione dehydrogenase / alcohol dehydrogenase	0	1451409.44	0.0000	0.0453	30.5	371	4.9043E-39
CN2/0829	alcohol dehydrogenase [EC:1.1.1.1]	0	1415676.12	0.0000	0.0442	44.2	425	4.602E-56
CN2/13226	alcohol dehydrogenase [EC:1.1.1.1]	164027.30	3701291.98	0.0150	0.1154	44.7	340	2.6024E-115
CN2/14077	alcohol dehydrogenase [EC:1.1.1.1]	0	2853599.21	0.0000	0.0890	37.8	357	1.5754E-42
CN2/14149	alcohol dehydrogenase [EC:1.1.1.1]	0	2121582.14	0.0000	0.0662	52.7	330	2.6007E-107
CN2/14228	alcohol dehydrogenase [EC:1.1.1.1]	219116.24	5539168.82	0.0200	0.1728	62.6	388	0
CN2/17718	Unknown function	114300.55	2468423.09	0.0105	0.0770	22.3	596	2.7381E-98
CN2/17719	alcohol dehydrogenase [EC:1.1.1.1]	0	3576707.68	0.0000	0.1116	27	385	1.8458E-144
CN2/15838	Unknown function	119819.15	488900.71	0.0110	0.0152	10.7	1259	9.3414E-34
CN2/16351	alcohol dehydrogenase [EC:1.1.1.1]	287443.87	4438562.88	0.0263	0.1384	34.3	335	2.0892E-23
CN1/0477	OmpA-OmpF porin, OOP family	2959758.36	0	0.2706	0.0000	51.2	86	1.2172E-72
CN1/21052	Naphthalene dioxygenase (beta)	20653305.70	15632289.36	1.8884	0.4876	79.7	79	6.8527E-172
CN2/11667	elongation factor EF-G [EC:3.6.5.3]	0	2028250.47	0.0000	0.0633	28.6	63	0.0049088
CN1/14685	Azospirillum B510 uid46085	2026955.55	0	0.1853	0.0000	24.7	170	1.9445E-24
CN1/2182	Unknown function	2127619.69	0	0.1945	0.0000	47.3	184	1.142E-86
CN2/15067	ribosome recycling factor	0	2858578.27	0.0000	0.0892	50.3	185	1.2287E-128
CN2/15515	NADPH2:quinone reductase [EC:1.6.5.5]	37889.32	2809307.94	0.0035	0.0876	51.1	325	7.2248E-110
CN2/17901	Membrane lipoprotein lipid attachment site containing protein USSDB6	0	1496805.19	0.0000	0.0467	21.4	206	2.2349E-27
CN2/17339	elongation factor EF-G [EC:3.6.5.3]	62324.29	17160351.48	0.0057	0.5352	73.1	495	0
CN1/22376	Uncharacterized protein conserved in bacteria	972659.66	0	0.0889	0.0000	13.7	322	2.9955E-06
CN2/12385	Outer membrane protein (porin)	0	756088.63	0.0000	0.0236	22.4	228	7.322E-25

CN2/12463	Outer membrane protein (porin)	194131.66	13343466.37	0.0177	0.4162	68.9	225	0
CN2/11528	TRAP-type mannitol/chloroaromatic compound transport system, periplasmic component	10549341.82	622436.88	0.9645	0.0194	15.9	378	2.6264E-74
CN1/21660	putative spermidine/putrescine transport system substrate-binding	7275296.32	237607.96	0.6652	0.0074	22.4	340	2.0561E-57
CN1/24218	Unknown function	10517896.06	0	0.9617	0.0000	39.8	88	3.1841E-22
CN1/6117	UDP-3-keto-D-GlcNAcA aminotransferase [EC:2.6.1.-]	277281.10	0	0.0254	0.0000	5.3	741	1.3415E-22
CN1/5537	Outer membrane protein (porin)	38874269.68	578179.82	3.5543	0.0180	52.5	335	0
CN2/17358	3-oxoacyl-[acyl-carrier protein] reductase (EC 1.1.1.100)	0	14609397.11	0.0000	0.4557	68.1	257	0
CN1/25150	Adenosylmethionine-8-amino-7-oxononanoate aminotransferase	320296.97	0	0.0293	0.0000	6.6	483	0.000007003
CN1/25151	TonB-dependent receptor	1551607.71	0	0.1419	0.0000	12.4	573	8.0436E-22
CN1/11478	outer membrane protein	735372.79	0	0.0672	0.0000	28.6	70	0.00057539
CN1/11723	peroxiredoxin (alkyl hydroperoxide reductase subunit C)	13542504.70	16304918.78	1.2382	0.5086	61.8	102	9.3921E-99
CN1/11796	ABC-type uncharacterized transport system, auxiliary component	170309.01	0	0.0156	0.0000	36.2	160	3.8706E-28
CN1/12077	Uncharacterized conserved protein	324859.70	0	0.0297	0.0000	18.3	191	1.7945E-06
CN1/13821	FKBP-type peptidyl-prolyl cis-trans isomerase FkpA [EC:5.2.1.8]	2947786.63	0	0.2695	0.0000	15.9	107	0.000058041
CN1/14794	Uncharacterized protein conserved in bacteria	2513430.51	0	0.2298	0.0000	34.9	195	2.0668E-17
CN1/15192	iron complex outermembrane recepter protein	365216.49	39574.75	0.0334	0.0012	13.7	358	6.7155E-43
CN1/16478	iron complex outermembrane recepter protein	289834.31	0	0.0265	0.0000	14.4	181	2.1603E-09
CN1/16841	Unknown function	18599734.87	0	1.7006	0.0000	38.4	73	9.9295E-09
CN1/19027	Unknown function	1102373.60	0	0.1008	0.0000	30.4	102	4.216E-08
CN1/24974	simple sugar transport system substrate-binding protein	1975651.37	659629.26	0.1806	0.0206	36.6	238	7.2373E-29
CN1/25732	Unknown function	352049.31	0	0.0322	0.0000	20.4	167	4.9963E-07
CN1/26001	OmpA-OmpF porin, OOP family	10819898.77	242300.91	0.9893	0.0076	63.4	142	5.4658E-113
CN1/26178	OmpA-OmpF porin, OOP family	2194198.89	0	0.2006	0.0000	57.8	102	3.3944E-158
CN1/26408	Outer membrane protein (porin)	6129628.10	412043.38	0.5604	0.0129	57.2	215	2.0084E-203
CN1/26758	thiol:disulfide interchange protein DsbA	2414134.82	0	0.2207	0.0000	45.6	103	3.4859E-25
CN1/3177	polar amino acid transport system substrate-binding protein	921903.07	0	0.0843	0.0000	27.4	208	1.6788E-62
CN1/5178	peroxiredoxin (alkyl hydroperoxide reductase subunit C)	1314924.92	4281.71	0.1202	0.0001	52.8	89	2.6057E-61
CN1/8884	peptidyl-prolyl cis-trans isomerase C [EC:5.2.1.8]	198720.44	320227.86	0.0182	0.0100	21.4	201	1.9789E-26
CN1/9884	iron complex outermembrane recepter protein	709898.28	0	0.0649	0.0000	30.3	165	1.011E-26
CN1/5146	copper resistance protein B	1104504.31	327738.98	0.1010	0.0102	12.7	150	0.00050301
CN1/7710	putative ABC transport system substrate-binding protein	8021361.34	0	0.7334	0.0000	45.9	157	2.2616E-16
CN1/12257	Outer membrane protein (porin)	12344733.71	2377619.65	1.1287	0.0742	30	237	6.385E-70
CN2/0121	Zn-dependent hydrolases, including glyoxylases	43317.62	22724029.06	0.0040	0.7088	68.1	326	1.9874E-211
CN1/20858	putative 3-hydroxyphenylpropionic acid porine	1554796.96	26403024.68	0.1422	0.8235	15.9	151	6.374E-37
CN1/0246	Unknown function	658310.02	2138171.97	0.0602	0.0667	6.7	1090	2.5707E-13
CN1/19236	Uncharacterized protein related to plant photosystem II stability/assembly factor	1384284.33	25897.93	0.1266	0.0008	13.6	338	1.9772E-25
CN1/5522	Unknown function	1310952.83	0	0.1199	0.0000	8	464	7.6954E-12
CN1/4879	Tricarboxylate transport protein Tct	3328087.27	651655.68	0.3043	0.0203	28.6	280	5.0147E-26
CN1/20107	TonB-dependent receptor	309411.38	0	0.0283	0.0000	12.7	497	4.2126E-67
CN1/8487	multiple sugar transport system substrate-binding protein	8281758.07	0	0.7572	0.0000	53.9	76	6.5422E-16
CN1/5006	TolB protein	2005453.82	45262.25	0.1834	0.0014	19.9	282	2.0231E-62
CN1/5808	ammonium transporter, Amt family	2290649.80	0	0.2094	0.0000	42	69	1.985E-08
CN1/6424	TRAP-type mannitol/chloroaromatic compound transport system	18345199.05	211679.38	1.6773	0.0066	30.2	242	3.7323E-96
CN1/6528	TRAP-type mannitol/chloroaromatic compound transport system	13393151.19	0	1.2246	0.0000	29.3	164	9.5148E-63
CN1/8007	molybdate transport system regulatory protein	2351921.03	0	0.2150	0.0000	32.4	68	2.3826E-10
CN1/8259	acyl carrier protein	4488969.75	0	0.4104	0.0000	16.5	79	2.076E-11
CN2/16717	Unknown function	20764.75	1972536.97	0.0019	0.0615	42.2	453	1.6018E-96
CN1/10473	Unknown function	217193.20	0	0.0199	0.0000	13.1	198	0.00013872

CN1/19990	Alkylphosphonate utilization operon protein Phn	669286.52	0	0.0612	0.0000	10.3	321	0.000025825
CN1/16278	chloride peroxidase [EC:1.11.1.10]	1352105.95	0	0.1236	0.0000	24.8	278	6.0006E-60
CN1/15691	isoquinoline 1-oxidoreductase, beta subunit [EC:1.3.99.16]	2603713.29	111049.00	0.2381	0.0035	30.9	320	3.6393E-74
CN1/5064	TRAP-type mannitol/chloroaromatic compound transport system	17171142.21	45995.86	1.5700	0.0014	23.6	123	1.6245E-34
CN1/6522	Unknown function	252733.16	166591.30	0.0231	0.0052	16.3	459	8.673E-20
CN1/9140	Trp repressor binding protein	1293099.55	827983.63	0.1182	0.0258	32.8	128	4.504E-17
CN1/27123	autoaggregation protein RapA/B/C	3051009.90	202458.22	0.2790	0.0063	36.3	256	1.6067E-161
CN2/17349	aryl-alcohol dehydrogenase (NADP+) [EC:1.1.1.91]	0	1013985.57	0.0000	0.0316	15.5	343	3.0498E-45
CN1/11117	polar amino acid transport system substrate-binding protein	3188188.85	175999.98	0.2915	0.0055	67	261	2.9604E-179
CN1/18498	TRAP-type C4-dicarboxylate transport system, periplasmic component	893288.27	0	0.0817	0.0000	8.6	303	6.3775E-33
CN1/16755	simple sugar transport system substrate-binding protein	1393812.70	10823.29	0.1274	0.0003	39.5	157	1.3529E-16
CN1/11755	Outer membrane protein/protective antigen OMA87	5316092.27	0	0.4861	0.0000	37.8	90	2.9362E-40
CN1/14868	Rubrerythrin	3980639.24	0	0.3640	0.0000	37.2	137	3.1479E-136
CN1/22033	simple sugar transport system substrate-binding protein	1527676.56	0	0.1397	0.0000	15.3	413	1.1702E-98
CN1/24001	2-Hydroxychromene-2-carboxylate isomerase	5585275.98	4475435.28	0.5107	0.1396	56.8	199	6.8289E-204
CN1/24003	outer membrane protein	11466397.21	13581614.11	1.0484	0.4236	55.6	153	1.9646E-149
CN2/11941	outer membrane protein	0	50188189.94	0.0000	1.5654	35	180	3.2096E-116
CN2/11945	Naphthalene dioxygenase (beta)	245103.68	32738920.68	0.0224	1.0212	63.9	194	0
CN2/15794	Naphthalene dioxygenase (alpha)	5717664.52	2312838.17	0.5228	0.0721	69.4	284	0
CN2/17233	Naphthalene dihydrodiol dehydrogenase	270367.83	3154740.66	0.0247	0.0984	91.9	62	0
CN1/27023	electron transfer flavoprotein alpha subunit	1148232.65	0	0.1050	0.0000	57.1	91	1.2525E-12
CN1/9298	peroxiredoxin (alkyl hydroperoxide reductase subunit C)	1801409.70	142461.74	0.1647	0.0044	51.3	187	6.6374E-43
CN2/0275	Xylulose-5-phosphate phosphoketolase (EC 4.1.2.9)	0	1090210.33	0.0000	0.0340	23.1	801	6.7064E-46
CN2/0456	outer membrane lipase/esterase	0	1540765.57	0.0000	0.0481	16.8	637	1.701E-53
CN2/0457	Unknown function	0	1389441.65	0.0000	0.0433	35.3	292	4.9516E-109
CN2/0475	chaperonin GroES	764448.49	17376547.98	0.0699	0.5420	80.4	97	5.6298E-176
CN2/0787	Unknown function	0	4934046.19	0.0000	0.1539	44.6	83	8.9817E-08
CN2/0797	Universal stress protein family	0	13838701.26	0.0000	0.4316	54.5	143	2.386E-38
CN2/0837	(1->4)-alpha-D-glucan 1-alpha-D-glucosylmutase [EC:5.4.99.15]	0	1886024.19	0.0000	0.0588	16.9	951	5.5492E-36
CN2/0839	maltooligosyltrehalose trehalohydrolase [EC:3.2.1.141]	0	995966.49	0.0000	0.0311	13.7	590	6.5453E-55
CN2/0875	uncharacterized membrane protein	0	499780.45	0.0000	0.0156	12.3	920	6.3103E-52
CN2/1035	molybdenum cofactor biosynthesis protein B	0	1544521.05	0.0000	0.0482	27.1	181	2.2342E-42
CN2/1036	molybdopterin biosynthesis protein MoeA	0	5199119.27	0.0000	0.1622	16.6	410	1.0411E-34
CN2/11629	trigger factor	47839.10	5459099.65	0.0044	0.1703	47.2	436	0
CN2/11630	cell division protein ZipA	0	7325839.12	0.0000	0.2285	9.5	273	5.1291E-16
CN2/11655	peptidyl-prolyl cis-trans isomerase D [EC:5.2.1.8]	0	465293.15	0.0000	0.0145	14.3	614	1.842E-10
CN2/11656	DNA-binding protein HU-beta	0	24075985.76	0.0000	0.7509	47.8	90	3.2948E-36
CN2/11768	FKBP-type peptidyl-prolyl cis-trans isomerase FklB [EC:5.2.1.8]	100015.06	6991568.43	0.0091	0.2181	49.2	252	0
CN2/11888	Na+-transporting NADH:ubiquinone oxidoreductase subunit A	0	2234632.38	0.0000	0.0697	13.7	445	2.1269E-06
CN2/12021	Uncharacterized conserved protein	605222.13	14699653.57	0.0553	0.4585	54.5	134	1.2205E-103
CN2/12022	Tricarboxylate porin Opd	163371.77	44113227.63	0.0149	1.3759	76	420	0
CN2/12033	translation initiation factor IF-3	0	1247485.68	0.0000	0.0389	22.6	177	7.3124E-25
CN2/12281	molecular chaperone GrpE	480318.43	3195527.31	0.0439	0.0997	32.6	187	3.7337E-42
CN2/12291	Predicted L-lactate dehydrogenase, Iron-sulfur cluster-binding subunit Ykg	0	566868.47	0.0000	0.0177	17.5	486	2.6123E-19
CN2/12320	hypothetical protein	0	9103235.27	0.0000	0.2839	61.7	115	1.5965E-62
CN2/12333	DNA-binding protein HU-alpha	0	4656308.93	0.0000	0.1452	32.2	90	3.4756E-25
CN2/12338	LemA protein	192462.28	6569621.08	0.0176	0.2049	32.2	202	1.9216E-139
CN2/12346	TdcF protein	0	4982579.32	0.0000	0.1554	69	126	7.4834E-47
CN2/12419	DNA topoisomerase I [EC:5.99.1.2]	0	328246.45	0.0000	0.0102	6.9	866	4.061E-10

CN2/12749	cell division protein FtsA	0	1210581.99	0.0000	0.0378	15.3	418	1.1667E-13
CN2/12760	Zn-dependent protease with chaperone function	0	1389391.32	0.0000	0.0433	24.8	270	6.2483E-24
CN2/12765	Unknown function	0	856066.65	0.0000	0.0267	16.6	453	1.4016E-17
CN2/12766	uncharacterized lipoprotein	0	6491535.90	0.0000	0.2025	53.6	194	6.6471E-221
CN2/12791	iron complex outermembrane recepter protein	55711.37	1528792.78	0.0051	0.0477	28.8	278	2.303E-30
CN2/1280	Unknown function	39079.41	1646636.56	0.0036	0.0514	19	564	1.9971E-55
CN2/1284	4-methyl-5(b-hydroxyethyl)-thiazole monophosphate biosynthesis	0	3483583.24	0.0000	0.1087	18.4	185	4.1566E-23
CN2/12849	cytosolic long-chain acyl-CoA thioester hydrolase family protein	130404.16	18528832.24	0.0119	0.5779	59.9	269	1.1941E-185
CN2/12910	GTP-binding protein	0	1793998.88	0.0000	0.0560	15	606	5.5277E-22
CN2/12930	oligopeptidase A [EC:3.4.24.70]	48731.56	1264172.35	0.0045	0.0394	16.5	683	3.5104E-34
CN2/12947	elongation factor EF-G [EC:3.6.5.3]	33825.02	9960815.84	0.0031	0.3107	40.2	204	2.6602E-89
CN2/13081	FOG: TPR repeat	85533.30	943662.23	0.0078	0.0294	25.9	586	1.7471E-78
CN2/13230	F0F1-type ATP synthase, beta subunit	433110.09	3924207.02	0.0396	0.1224	59.5	126	5.2963E-30
CN2/13268	Uncharacterized conserved protein	0	3271681.80	0.0000	0.1020	14.6	144	4.6796E-27
CN2/13353	serine 3-dehydrogenase [EC:1.1.1.276]	0	5060486.38	0.0000	0.1578	41.3	254	1.6732E-112
CN2/13383	Unknown function	0	2723745.03	0.0000	0.0850	27.4	223	6.2166E-66
CN2/13435	iron complex outermembrane recepter protein	115055.10	2764109.05	0.0105	0.0862	38.9	620	0
CN2/13463	polysaccharide export outer membrane protein	123993.41	5636789.31	0.0113	0.1758	64	358	0
CN2/13558	Mlr7403 protein	0	787992.19	0.0000	0.0246	31.9	213	3.3739E-44
CN2/13562	type IV pilus assembly protein PilF	0	716688.96	0.0000	0.0224	18.7	251	1.4606E-81
CN2/1357	electron-transferring-flavoprotein dehydrogenase [EC:1.5.5.1]	67584.73	2384738.94	0.0062	0.0744	29.2	551	2.1173E-51
CN2/1359	electron transfer flavoprotein beta subunit	0	5420774.53	0.0000	0.1691	32.1	249	1.618E-98
CN2/1360	electron transfer flavoprotein alpha subunit	4631609.93	9016180.12	0.4235	0.2812	58.3	309	7.3995E-204
CN2/1369	Unknown function	0	3567363.75	0.0000	0.1113	36	353	5.7053E-77
CN2/13795	ATPases with chaperone activity, ATP-binding subunit	0	4138363.04	0.0000	0.1291	33.1	118	3.0135E-27
CN2/13911	hypothetical protein	0	1222144.56	0.0000	0.0381	19.7	142	0.00015178
CN2/13966	ribonuclease G [EC:3.1.26.-]	0	772906.01	0.0000	0.0241	16.5	485	8.5347E-10
CN2/13969	TldD protein	0	794024.16	0.0000	0.0248	17.3	480	1.0763E-58
CN2/13971	PmbA protein	0	1252041.10	0.0000	0.0391	31	448	3.6327E-87
CN2/13975	putative sigma-54 modulation protein	0	4986807.56	0.0000	0.1555	44.1	102	1.4657E-20
CN2/14058	Unknown function	47748.90	10571177.97	0.0044	0.3297	63.1	317	0
CN2/14145	Molecular chaperone (small heat shock protein)	194150.89	32774659.48	0.0178	1.0223	56.5	177	3.8639E-57
CN2/14220	ammonium transporter, Amt family	0	4485764.06	0.0000	0.1399	6.3	426	0.000038407
CN2/14221	nitrogen regulatory protein P-II 2	57690.17	14778951.74	0.0053	0.4610	57.1	112	3.3448E-37
CN2/14233	GTP pyrophosphokinase (EC 2.7.6.5), (p)ppGpp synthetase	200180.17	0	0.0183	0.0000	13	123	0.00070318
CN2/14333	bacterioferritin	51633.19	16503074.74	0.0047	0.5147	70.5	156	1.1713E-81
CN2/14338	NADH:flavin oxidoreductases, Old Yellow Enzyme family	28045.61	2881386.50	0.0026	0.0899	45	369	6.426E-145
CN2/14563	OmpA-OmpF porin, OOP family	36700.90	6219913.32	0.0034	0.1940	60	260	1.8188E-286
CN2/14613	F0F1-type ATP synthase, beta subunit	0	1235464.26	0.0000	0.0385	44.5	119	1.5657E-17
CN2/14664	Unknown function	0	1508266.43	0.0000	0.0470	27.5	222	1.5186E-29
CN2/14724	Unknown function	0	1635746.11	0.0000	0.0510	46.2	158	2.4142E-106
CN2/1475	Lipid A biosynthesis lauroyl acyltransferase (EC 2.3.1.-)	0	3680780.15	0.0000	0.1148	49.5	192	8.0859E-42
CN2/14940	outer membrane lipoproteins carrier protein	0	1655572.56	0.0000	0.0516	28.4	215	1.043E-54
CN2/15053	Predicted periplasmic/secreted protein	0	1446205.28	0.0000	0.0451	18.8	234	3.0867E-10
CN2/15061	outer membrane protein	21917.10	5084840.04	0.0020	0.1586	40.9	171	2.3406E-52
CN2/15062	outer membrane protein	0	2151054.58	0.0000	0.0671	19	795	2.1055E-40
CN2/15071	methionyl aminopeptidase [EC:3.4.11.18]	0	2659417.43	0.0000	0.0829	18.5	260	9.6316E-20
CN2/15079	outer membrane lipoprotein SlyB	59610.56	16837870.15	0.0055	0.5252	64.1	153	4.3291E-118
CN2/15156	solute:Na+ symporter, SSS family	0	403570.58	0.0000	0.0126	7.9	393	1.7167E-06

CN2/15164	Glycine cleavage system protein P (pyridoxal-binding), C-terminal domain	1154580.41	17193765.40	0.1056	0.5363	55.4	157	6.6122E-113
CN2/15309	Unknown function	0	1755987.83	0.0000	0.0548	13.5	222	0.00029571
CN2/15310	Unknown function	81538.01	2921473.52	0.0075	0.0911	50.5	216	1.2617E-39
CN2/15318	glycine cleavage system transcriptional repressor	0	2564665.64	0.0000	0.0800	52.3	172	1.6899E-28
CN2/15512	Uncharacterized protein with LysM domain, COG1652	0	1501096.45	0.0000	0.0468	37.2	341	1.6129E-78
CN2/15523	iron complex outermembrane recepter protein	0	402484.01	0.0000	0.0126	5.3	681	2.2529E-46
CN2/15751	NADPH2:quinone reductase [EC:1.6.5.5]	31992.18	1308617.80	0.0029	0.0408	17.2	320	8.2474E-11
CN2/15752	carboxyl-terminal processing protease [EC:3.4.21.102]	0	1564642.98	0.0000	0.0488	11.7	693	1.8461E-22
CN2/15844	biotin biosynthesis protein BioH	0	2724089.18	0.0000	0.0850	32	241	6.6231E-207
CN2/1587	aminotransferase [EC:2.6.1.-]	17135.58	1763103.82	0.0016	0.0550	16.6	403	1.0245E-67
CN2/15914	pyrroloquinoline-quinone synthase [EC:1.3.3.11]	0	978982.47	0.0000	0.0305	54	63	3.9488E-07
CN2/15987	ATP-dependent protease La (EC 3.4.21.53) Type	0	333423.74	0.0000	0.0104	7.5	560	3.4074E-14
CN2/15990	FKBP-type peptidyl-prolyl cis-trans isomerase FklB [EC:5.2.1.8]	0	1348022.97	0.0000	0.0420	34.1	205	2.8932E-139
CN2/1610	elongation factor EF-P	487165.19	14154952.94	0.0445	0.4415	48.9	188	9.248E-140
CN2/16119	Na+-transporting NADH:ubiquinone oxidoreductase subunit F	0	1762110.86	0.0000	0.0550	11.9	387	0.000079395
CN2/16129	lipid A 3-O-deacylase [EC:3.1.-.-]	11719.37	9642140.73	0.0011	0.3007	31.4	172	2.6663E-16
CN2/16135	Argininosuccinate lyase	27251.95	3079187.27	0.0025	0.0960	34.8	626	3.7884E-153
CN2/16136	Unknown function	112580.83	7195672.58	0.0103	0.2244	76.2	433	0
CN2/16145	peroxiredoxin (alkyl hydroperoxide reductase subunit C)	55482.43	4717227.71	0.0051	0.1471	74.3	187	7.3615E-165
CN2/16154	hypothetical protein	0	2076897.17	0.0000	0.0648	21.4	159	4.3449E-07
CN2/16183	Methyl-accepting chemotaxis protein	183821.65	937481.74	0.0168	0.0292	29.3	297	4.5525E-55
CN2/16185	Unknown function	146455.05	14603270.94	0.0134	0.4555	46.4	321	2.0129E-265
CN2/16206	thiol:disulfide interchange protein DsbA	0	2586822.77	0.0000	0.0807	48.3	209	1.9503E-58
CN2/16207	Cytochrome c4	85409.56	2023905.01	0.0078	0.0631	18	205	4.6888E-11
CN2/16209	glycine cleavage system H protein	0	12018776.55	0.0000	0.3749	15.7	127	8.4792E-06
CN2/16229	carboxyl-terminal processing protease [EC:3.4.21.102]	0	4543983.08	0.0000	0.1417	17.8	445	4.1714E-59
CN2/16246	Outer membrane protein and related peptidoglycan-associated (lipo)proteins	19764.40	585492.62	0.0018	0.0183	18.9	977	1.6519E-213
CN2/16498	large conductance mechanosensitive channel, MscL family	0	3884423.82	0.0000	0.1212	41.3	138	6.3829E-127
CN2/16506	TPR repeat containing exported protein	11358390.87	1504446.11	1.0385	0.0469	36	272	2.0417E-191
CN2/16507	peptidoglycan-associated lipoprotein	0	37010548.34	0.0000	1.1544	45.5	167	1.6526E-106
CN2/16518	starvation-inducible DNA-binding protein	0	13178643.91	0.0000	0.4111	54.5	156	4.0843E-105
CN2/16522	Outer membrane porin, OprD family	123143.39	24060158.35	0.0113	0.7505	63.4	421	0
CN2/16571	putative acyl-CoA dehydrogenase	18730.15	570782.15	0.0017	0.0178	27.7	548	1.7546E-43
CN2/16737	Unknown function	0	1044783.24	0.0000	0.0326	12.6	183	1.5153E-06
CN2/16772	DNA gyrase subunit A [EC:5.99.1.3]	0	642887.72	0.0000	0.0201	8.7	935	1.1146E-68
CN2/16833	Unknown function	0	1340292.37	0.0000	0.0418	31.9	248	1.1232E-81
CN2/17000	peptidyl-prolyl cis-trans isomerase B (cyclophilin B) [EC:5.2.1.8]	0	8717506.40	0.0000	0.2719	68.3	180	3.9648E-133
CN2/17003	Universal stress protein family 7	0	2804243.59	0.0000	0.0875	36.5	285	9.0154E-110
CN2/17056	peptidyl-prolyl cis-trans isomerase SurA [EC:5.2.1.8]	140952.23	1660800.32	0.0129	0.0518	62.5	435	4.385E-273
CN2/17080	Unknown function	15566.65	1712825.63	0.0014	0.0534	27.6	605	2.7642E-174
CN2/17083	21 kDa hemolysin precursor	0	1109402.97	0.0000	0.0346	28.8	191	2.7343E-73
CN2/17107	Unknown function	17945.09	2844780.32	0.0016	0.0887	18	339	5.5602E-22
CN2/17108	LPS-assembly protein	0	1393107.46	0.0000	0.0435	31	384	2.0207E-108
CN2/17120	2-alkenal reductase [EC:1.3.1.74]	182818.40	3390284.71	0.0167	0.1057	34.8	469	4.7444E-80
CN2/17122	sigma-E factor negative regulatory protein RseB	31754.19	1883869.37	0.0029	0.0588	26.8	310	3.8689E-27
CN2/17170	mitochondrial processing peptidase [EC:3.4.24.64]	0	347547.60	0.0000	0.0108	16.9	503	3.4158E-44
CN2/17408	ATP-dependent Clp protease, protease subunit [EC:3.4.21.92]	0	10951794.12	0.0000	0.3416	70	90	5.1931E-36
CN2/17440	peptidyl-prolyl cis-trans isomerase A (cyclophilin A) [EC:5.2.1.8]	0	7337805.77	0.0000	0.2289	67.4	184	4.4589E-165
CN2/17462	S-adenosyl-methyltransferase [EC:2.1.1.-]	0	1030041.03	0.0000	0.0321	14	315	5.8631E-16

CN2/17493	Unknown function	107509.90	839173.83	0.0098	0.0262	26.9	520	9.0687E-57
CN2/17514	OmpA-OmpF porin, OOP family	128694.26	17981993.51	0.0118	0.5609	66.6	329	0
CN2/17537	Unknown function	654383.81	16013290.05	0.0598	0.4995	44.1	93	1.2143E-99
CN2/17567	iron complex outermembrane recepter protein	112456.60	1020434.74	0.0103	0.0318	43.2	621	0
CN2/17613	Uncharacterized protein involved in an early stage of isoprenoid biosynthesis	0	2747885.88	0.0000	0.0857	32.1	221	3.9224E-72
CN2/17658	lipoprotein-34	21035.53	2138759.82	0.0019	0.0667	51.2	365	4.7276E-56
CN2/17677	Unknown function	0	2091655.56	0.0000	0.0652	8.3	218	0.00093854
CN2/17686	Outer membrane protein A precursor	0	3065529.09	0.0000	0.0956	46.9	245	2.3159E-14
CN2/17687	Signal transduction histidine kinase	0	384388.84	0.0000	0.0120	2.7	1058	0.011227
CN2/17717	3-hydroxyacyl-CoA dehydrogenase (EC 1.1.1.35)	0	974071.00	0.0000	0.0304	22.6	412	1.7345E-37
CN2/17879	putative glutamine amidotransferase	0	2444294.19	0.0000	0.0762	34.7	248	9.7782E-80
CN2/17900	HlyD family secretion protein	0	834199.35	0.0000	0.0260	10.3	358	3.5496E-06
CN2/17902	putative ABC transport system substrate-binding protein	104421.43	1086979.51	0.0095	0.0339	26.3	312	5.3589E-55
CN2/17912	polar amino acid transport system substrate-binding protein	149493.67	9825621.49	0.0137	0.3065	81.7	251	4.1354E-252
CN2/17961	thiol peroxidase, atypical 2-Cys peroxiredoxin [EC:1.11.1.15]	0	13081306.68	0.0000	0.4080	60.2	166	8.8471E-224
CN2/17991	LPS-assembly lipoprotein	22131.58	1066004.71	0.0020	0.0332	40.3	201	9.7684E-93
CN2/18021	thiosulfate sulfurtransferase [EC:2.8.1.1]	0	636381.47	0.0000	0.0198	31	271	2.2184E-13
CN2/18182	Predicted oxidoreductases (related to aryl-alcohol dehydrogenases)	0	799101.33	0.0000	0.0249	38.3	269	2.1516E-56
CN2/18183	Glycerol-3-phosphate dehydrogenase	0	2308220.17	0.0000	0.0720	18.7	390	1.2802E-15
CN2/18212	bacterioferritin	0	3125646.07	0.0000	0.0975	72.9	155	2.1391E-64
CN2/18227	AsmA protein	169955.61	1732677.50	0.0155	0.0540	42.7	740	6.4875E-275
CN2/1838	thioredoxin 2 [EC:1.8.1.8]	0	1585110.57	0.0000	0.0494	43.8	144	9.229E-41
CN2/2174	hypothetical protein	39081.58	2627784.65	0.0036	0.0820	41.8	225	3.2895E-39
CN2/2288	Unknown function	0	1925449.96	0.0000	0.0601	21.6	389	2.336E-26
CN2/2291	alkanal monooxygenase (FMN-linked) [EC:1.14.14.3]	0	1299689.33	0.0000	0.0405	48.3	331	1.0088E-51
CN2/2292	hypothetical protein	0	7217821.84	0.0000	0.2251	38.1	105	9.3946E-21
CN2/2298	GST-like protein	60695.72	11589993.17	0.0055	0.3615	52.9	280	5.368E-292
CN2/2911	Valyl-tRNA synthetase	0	1031213.50	0.0000	0.0322	32.7	223	2.3766E-29
CN2/3584	Protein Yic	218455.10	1308712.15	0.0200	0.0408	30.7	287	3.1816E-47
CN2/3585	ribonuclease PH [EC:2.7.7.56]	0	5423855.93	0.0000	0.1692	21.7	240	5.2974E-34
CN2/4442	multiple sugar transport system substrate-binding protein	106170.45	1069825.65	0.0097	0.0334	20.2	415	1.3377E-90
CN2/5114	hypothetical protein	105308.22	744580.82	0.0096	0.0232	42.8	346	1.4244E-67
CN2/5117	Unknown function	41026.64	724670.63	0.0038	0.0226	36	236	8.5089E-94
CN2/5132	cytochrome c peroxidase [EC:1.11.1.5]	59616.24	3423809.37	0.0055	0.1068	36	464	3.2561E-179
CN2/5239	acyl carrier protein	811525.51	35460358.62	0.0742	1.1060	32.1	78	3.0172E-51
CN2/5738	Enoyl-[acyl-carrier-protein] reductase [FMN] (EC 1.3.1.9)	0	18220224.23	0.0000	0.5683	44.7	347	5.1944E-146
CN2/7223	protease I [EC:3.2.-.-]	0	12124659.45	0.0000	0.3782	48.7	193	1.7484E-39
CN2/7241	Unknown function	0	1348437.60	0.0000	0.0421	8	698	2.8877E-26
CN2/0264	OmpA-OmpF porin, OOP family	472805.95	40000641.97	0.0432	1.2477	61.3	142	4.4408E-242
CN2/0835	glycogen operon protein GlgX [EC:3.2.1.-]	1231414.88	1144748.84	0.1126	0.0357	18.7	717	1.4655E-32
CN2/12279	molecular chaperone DnaJ	0	879851.78	0.0000	0.0274	14.4	375	1.7976E-40
CN2/12850	TRAP-type mannitol/chloroaromatic compound transport system, periplasmic component	168032.53	16294287.92	0.0154	0.5082	77.6	98	1.706E-179
CN2/13249	cold shock protein (beta-ribbon, CspA family)	0	33270824.16	0.0000	1.0377	40	70	4.206E-07
CN2/13319	translation initiation factor IF-2	0	2672424.01	0.0000	0.0834	8.5	831	1.1858E-154
CN2/14339	Uncharacterized protein conserved in bacteria	0	2878836.19	0.0000	0.0898	33.3	240	4.2545E-58
CN2/15069	elongation factor EF-Ts	289130.96	8914916.95	0.0264	0.2781	70.7	287	0
CN2/15328	iron complex outermembrane recepter protein	0	3487200.84	0.0000	0.1088	20.4	457	1.299E-142
CN2/16716	Anaerobic dehydrogenases, typically selenocysteine-containing	28858.25	1685035.09	0.0026	0.0526	38.1	544	1.1393E-109

CN2/17057	LPS-assembly protein	749180.46	3594574.52	0.0685	0.1121	36	514	2.8594E-108
CN2/17618	thioredoxin 1	0	11398517.96	0.0000	0.3555	27	137	4.3644E-29
CN2/17638	Unknown function	0	906360.61	0.0000	0.0283	16.9	496	1.7355E-61
CN2/18040	periplasmic glucans biosynthesis protein	0	1310522.65	0.0000	0.0409	33.3	366	1.2892E-102
CN2/2240	osmotically inducible protein OsmC	0	4073243.27	0.0000	0.1270	37.1	143	9.0132E-103
CN2/0281	ferredoxin--NADP+ reductase [EC:1.18.1.2]	169179.02	1370599.60	0.0155	0.0428	46.5	258	5.7885E-136
CN2/12083	twitching motility protein PilT	61250.58	408621.77	0.0056	0.0127	23.5	344	6.7428E-26
CN2/17050	serine protein kinase	0	2054950.05	0.0000	0.0641	28.3	640	1.5496E-108
CN2/17407	ATP-dependent Clp protease ATP-binding subunit ClpX	415803.99	931210.59	0.0380	0.0290	43.9	426	5.1554E-77
CN2/18124	ATP-dependent Clp protease ATP-binding subunit ClpB	53402.38	2831498.45	0.0049	0.0883	53	728	0
CN1/8266	Unknown function	6392606.61	0	0.5845	0.0000	39.1	197	3.2133E-122
CN1/15566	iron complex outermembrane recepter protein	840301.09	107300.63	0.0768	0.0033	16.3	619	9.491E-18
CN1/24122	multiple sugar transport system substrate-binding protein	3974650.30	401575.44	0.3634	0.0125	54.9	364	3.7509E-133
CN1/0231	Outer membrane protein/protective antigen OMA87	2051377.87	0	0.1876	0.0000	25	276	9.9649E-76
CN1/0554	2-hydroxychromene-2-carboxylate isomerase	453901.86	0	0.0415	0.0000	51.9	206	2.017E-52
CN1/0936	LPS-assembly protein	250123.06	0	0.0229	0.0000	7.3	648	3.5883E-06
CN1/10146	outer membrane protein	521346.60	51246.90	0.0477	0.0016	30.6	209	1.163E-26
CN1/1049	ABC-type transport system involved in resistance to organic solvents, auxiliary component	730936.68	0	0.0668	0.0000	36.4	132	4.5867E-24
CN1/10649	starvation-inducible DNA-binding protein	7320473.35	0	0.6693	0.0000	42.2	166	7.2014E-83
CN1/11214	putative thioredoxin	1349192.37	243319.19	0.1234	0.0076	27.5	316	5.919E-20
CN1/11840	Unknown function	2889432.20	0	0.2642	0.0000	20.6	170	2.4923E-07
CN1/13415	polar amino acid transport system substrate-binding protein	2761870.80	163799.11	0.2525	0.0051	33.1	284	1.9952E-153
CN1/1478	solute:Na+ symporter, SSS family	213099.72	0	0.0195	0.0000	8.3	363	2.6205E-17
CN1/1887	Periplasmic thiol:disulfide interchange protein Dsb	1037331.31	105759.74	0.0948	0.0033	15.9	220	3.074E-28
CN1/1965	DNA-binding protein HU-beta	930643.80	0	0.0851	0.0000	32.2	90	0.00017931
CN1/20855	transcription elongation factor GreA	4000823.59	0	0.3658	0.0000	19.7	137	0.000001193
CN1/2103	peptidoglycan-associated lipoprotein	7316913.83	0	0.6690	0.0000	34.1	164	3.6465E-36
CN1/2196	Unknown function	442603.64	0	0.0405	0.0000	7.5	575	1.1968E-07
CN1/2218	peroxiredoxin (alkyl hydroperoxide reductase subunit C)	2514793.37	0	0.2299	0.0000	31.7	104	1.0879E-08
CN1/2378	outer membrane protein	1418055.65	0	0.1297	0.0000	20.2	535	8.6348E-38
CN1/25231	chaperonin GroES	9275254.25	0	0.8480	0.0000	39.6	96	2.9108E-70
CN1/3522	simple sugar transport system substrate-binding protein	588176.54	0	0.0538	0.0000	17.5	348	1.2496E-132
CN1/5314	basic membrane protein A and related proteins	3959616.08	0	0.3620	0.0000	10.1	335	2.2892E-12
CN1/5616	Unknown function	286381.06	0	0.0262	0.0000	14.6	521	1.4184E-13
CN1/6969	Uncharacterized protein conserved in bacteria	481061.18	0	0.0440	0.0000	20.9	211	0.000011592
CN1/9997	Unknown function	4212298.07	1010176.44	0.3851	0.0315	17.4	357	7.8037E-13
CN1/22264	ATP-dependent Clp protease, protease subunit [EC:3.4.21.92]	812588.62	132190.13	0.0743	0.0041	16.8	208	6.2133E-09
CN1/20701	electron transfer flavoprotein alpha subunit	1391430.45	31091.71	0.1272	0.0010	29.5	129	1.2023E-09
CN1/0115	putative ABC transport system substrate-binding protein	0.00	0	0.0000	0.0000	40.1	272	1.1521E-47
CN1/16960	polar amino acid transport system substrate-binding protein	415293.75	0	0.0380	0.0000	13	216	1.9437E-07
CN1/23563	Unknown function	380270.44	0	0.0348	0.0000	23.7	295	9.4439E-24
CN1/6723	Unknown function	1257365.42	0	0.1150	0.0000	31.3	83	0.000083739

[1] The metaproteomes of CN1 and CN2 were investigated via 1-dimensional (1-DE) gel-based pre-separation of proteins and subsequent tryptic digestion and fractionation and identification of the resulting peptides by a nano-UPLC system coupled to an LTQ-Orbitrap mass spectrometer using the DNA metagenome sequences as templates. Following criteria established in **Supplementary Materials and Methods** proteins were unambiguously identified using the annotation pipeline. Theoretical amino acid protein length (based on data provided in the annotation pipeline) and the coverage of identified peptides within the entire protein sequence are specifically shown. Full details are given in **Supplementary Materials and Methods**.

[2] Normalized *per protein* intensities were calculated as average of *per peptide* intensities in replicas. These *per peptide* intensities were summed and relative concentration of individual proteins calculated by dividing *per protein* intensities and the averages of the *summed* intensities.

[3] PEP is defined as the Posterior Error Probability of the identification. This value essentially operates as a *p*-value.

Table S7 Proteins of CN1 and CN2 communities identified and quantified by metaproteomic approaches putatively involved in the naphthalene degradation pathway. Protein annotation and tentative phylogenetic affiliations are specifically shown. For statistical significance and content description (including footnotes) see **Tables S5** and **S6**.

List of unique proteins	**Protein (ORFs)**	**Best hit**	**Tentative phylogenetic affiliation and presumptive plasmid/genome location**	**Annotation**	**Rel. conc. CN1 (%)**	**Rel. conc. CN2 (%)**
1	CN1/10034	99% *Ralstonia* sp. U2 (AAD12610)	*Betaproteobacteria* (*Achromobacter*)	Naphthalene dioxygenase (alpha)	1.7436	0.3546
	CN1/21052	89% *Ralstonia* sp. U2 (AP12611)		Naphthalene dioxygenase (beta)	1.8884	0.4876
	CN2/8666	95% NagAc *Ralstonia* sp. U2 (AAD12610)		Naphthalene dioxygenase (alpha)	0.0063	0.1100
	CN2/15794	97% NagAc *Ralstonia* sp. U2 (AAD12610)		Naphthalene dioxygenase (alpha)	0.5228	0.0721
2	CN1/14978	100% *Ralstonia* sp. U2 (AP12614)	*Betaproteobacteria* (*Achromobacter*)	Dihydroxynaphthalene dioxygenase	2.1339	0.4816
	CN2/8417	93% *Ralstonia* sp. U2 (AAD12614)		Dihydroxynaphthalene dioxygenase	0.5554	0.1127
3	CN2/17233	94% *Ralstonia* sp. U2 (AAD12612)	*Betaproteobacteria* (*Achromobacter*)	Naphthalene dihydrodiol dehydrogenase	0.0247	0.0984
4	CN1/24001	100% *Ralstonia* sp. U2 (AAD12617)	*Betaproteobacteria* (*Achromobacter*)	2-Hydroxychromene-2-carboxylate isomerase	0.5107	0.1396
5	CN1/23998	99% *Ralstonia* sp. U2 (AAD12166)	*Betaproteobacteria* (*Achromobacter*)	Hydroxybenzylidenepyruvate hydratase-aldolase	0.0577	0.0151
6	CN1/14979	98% *Ralstonia* sp. U2 (AAD12613)	*Betaproteobacteria* (*Achromobacter*)	Salicylaldehyde dehydrogenase	0.7457	1.7291
7	CN1/24000	100% *Ralstonia* sp. U2 (AAD12618)	*Betaproteobacteria* (*Achromobacter*)	Glutathione S transferase (NagJ-like)	0.2015	0.1110
8	CN1/6354	100% *Ralstonia* sp. U2 (AAD12619)	*Betaproteobacteria* (*Achromobacter*)	Gentisate dioxygenase	1.9205	0.3978
	CN1/23999	100% *Ralstonia* sp. U2 (AAD12619)		Gentisate dioxygenase	2.0062	0.4507
9	CN2/3389	82% *Leptothrix cholodnii* SP-6 Lcho3671 (ACB35925)	*Betaproteobacteria* (genus unclear)	Fumarylpyruvate hydrolase	0.5060	0.0000
10	CN1/20819	83% *Polaromonas* sp. JS666 (Bpro_0983)	*Betaproteobacteria* (*Achromobacter*)	Salicylate 5-hydroxylase	0.0000	0.0248
11	CN2/11945	99% NahAd *Pseudomonas* sp. ND6 (NP_943189) and 78%	*Gammaproteobacteria* (*Pseudomonas*)	Naphthalene dioxygenase (beta)	0.0224	1.0212

		Ralstonia sp. U2 (AAD12611)				
	CN2/11946	99% NahAc *Pseudomonas* sp. ND6 (NP_943188) and 89% *Ralstonia* sp. U2 (AAD12610)		Naphthalene dioxygenase (alpha)	0.0981	0.5599
	CN2/11948	NahAa 99% *Pseudomonas* sp. ND6 (NP_943186) and 67% *Ralstonia* sp. U2 (AAD12606)		Naphthalene dioxygenase (reductase)	0.0550	0.0978
12	CN2/11942	100% NahC *Pseudomonas* sp. ND6 (NP_943192) and 88% *Ralstonia* sp. U2 (AAD12614)	*Gammaproteobacteria* (*Pseudomonas*)	Dihydroxynaphthalene dioxygenase	0.1129	0.7332
13	CN2/11944	99% NahB *Pseudomonas* sp. ND6 (NP_943190) and 85% *Ralstonia* sp. U2 (AAD12612)	*Gammaproteobacteria* (*Pseudomonas*)	Naphthalene dihydrodiol dehydrogenase	0.0130	0.7907
14	CN2/11940	100% NahE *Pseudomonas* sp. ND6 (NP_943093) and 89% *Ralstonia* sp. U2 (AAD12616)	*Gammaproteobacteria* (*Pseudomonas*)	Hydroxybenzylidenepyruvate hydratase-aldolase	0.0221	0.4242
15	CN2/11943	100% NahF *Pseudomonas* sp. ND6 (NP_943191) and 89% *Ralstonia* sp. U2 (AAD12613)	*Gammaproteobacteria* (*Pseudomonas*)	Salicylaldehyde dehydrogenase	0.2739	0.9920
16	CN2/16710	100% *Pseudomonas putida* AK5 (ACO92378)	*Gammaproteobacteria* (*Pseudomonas*)	Salicylate 5-hydroxylase	0.0023	0.4404
	CN2/16711	100% *Pseudomonas putida* AK5 (ACO92377)		Salicylate 5-hydroxylase	0.0074	0.2445
17	CN2/16713	100% *Pseudomonas putida* AK5 (ACO92375)	*Gammaproteobacteria* (*Pseudomonas*)	Gentisate dioxygenase	0.0120	0.6514
18	CN2/16708	100% *Pseudomonas putida* AK5 (ACO92383)	*Gammaproteobacteria* (*Pseudomonas*)	Fumarylpyruvate hydrolase	0.0037	0.2162
19	CN2/16712	100% *Pseudomonas putida* AK5 (ACO92376)	*Gammaproteobacteria* (*Pseudomonas*)	Fumarylpyruvate hydrolase	0.1889	0.5318
20	CN2/17371	100% NahH *Pseudomonas* sp. ND6 (AAP44220)	*Gammaproteobacteria* (*Pseudomonas*)	Catechol 2,3-dioxygenase	0.0000	0.0698
21	CN2/8952	100% NahL *Pseudomonas putida* GJ31 (AAX38568)	*Gammaproteobacteria* (*Pseudomonas*)	2-Oxopent-4-enoate hydratase	0.0000	0.0966
22	CN2/17367	100% NahM *Pseudomonas* sp. ND6 (AAP44215)	*Gammaproteobacteria* (*Pseudomonas*)	4-Hydroxy-2-oxovalerate aldolase	0.0027	0.5603
23	CN2/17368	100% NahO *Pseudomonas* sp. ND6 (AAP44216)	*Gammaproteobacteria* (*Pseudomonas*)	Acetaldehyde dehydrogenase	0.0082	0.2935
24	CN2/11941	98% NahQ *Pseudomonas* sp. ND6 (AAP44191) and 76% *Ralstonia* sp. U2 (AAD12615)	*Gammaproteobacteria* (*Pseudomonas*)	Membrane protein	0.0000	1.5654

25	CN1/0554	89% NahD *Azospirillum* B510 Azl004590 (YP_003447641)	*Alphaproteobacteria* (*Azospirillum*)	2-Hydroxychromene-2-carboxylate isomerase	0.0415	0.0000
26	CN1/3872	53% *Ralstonia* sp. U2 (AAD12619)	*Alphaproteobacteria* (*Azospirillum*)	Gentisate dioxygenase	0.3299	0.0000

Table S8 In-house database containing sequences of enzymes known to be involved in the aerobic metabolism of aromatics via di- and trihydroxylated intermediates. The Table is provided as Excel document as separate file.

Table S9 Complete information regarding gene prediction and annotation of metagenome sequences herein obtained. Full description of methods used for gene annotation is provided in **Supplementary Materials and Methods**. The Table is provided as Excel document as separate file.

SUPPORTING FIGURE LEGENDS

Figure S1 Features and location of the PAH-contaminated soil. This soil was affected by different PAH spills for decades by a chemical plant (see text for details). The area of the initial site remediation study covers about 25.000 m^2, and the samples N and Nbs were taken as a composite and representative sample of the central area. This parcel was included in the “National inventory of contaminated soils” in the region of Asturias (Spain) in May 2001.

Figure S2 GC-MS total ion chromatogram of the sample N used in the present study. Major pollutants are indicated, namely 1 (indene), N (naphthalene), MN (methylnaphthalenes), 2 (alkanes), 3 (phenanthrene), 4 (phtalates and fatty acids) and 5 (non identified nitrogen-containing heterocycles). Note: it was not possible to obtain actually (and neither in the past) unpolluted same type of soil given the edaphic and hydric conditions in the area and the extension of contamination; therefore, we have developed a detailed forensic study of the site (not included in this work) that includes the following conclusions. First, PAHs (predominant contaminants at this site) are linked to former activities of coal tar processing. Second, the presence of alkanes is related to former fuel and gasoil spills from the heating systems of the factory. Third, fatty acids were identified in coherence with the former storage of ‘Tinol’, a residue from steel plants, composed of a mixture of animal fats, mineral lubricants and chips. Fourth, phtalates are usually associated with the manipulation of samples, although in our case the important signals found are clearly linked to one of the main activities of the factory (resin manufacture). Fifth, the N heterocycles found in the samples are derived of the use of these products in other minoritary processes carried out in the factory (chemical products manufacturing). As a conclusion it is possible that traces of some of these contaminants had a natural origin or were attributable to the manipulation of samples; however, in our case chemical, physical and even "historical" evidences justify the classification of the indicated peaks in the chromatogram as pollutants.

Figure S3 DGGE fingerprints of PCR-amplified 16S rDNA gene fragments from contaminated soil (N), contaminated soil after 231 days of biostimulation (Nbs) and enrichment cultures (CN1, CN2) after four different transfers. 1: CN1 Transfer 20; 2: CN1 Transfer 35; 3: CN1 Transfer 45; 4: CN1 Transfer 62.5: CN2 Tranfer 20; 6:CN2 Transfer 35; 7: CN2 Transfer 45;

8:CN2 Transfer 60; 9: Contaminated soil (N); 10: Biostimulated soil after 231 days of treatment. DGGE fingerprints after transfer 62 for CN1 and transfer 60 for CN2 were equally the same and are not shown. Gels were stained with SYBR Gold, and visualized by UV-fluorescence. Major bands excised and sequenced are indicated: A, *Flavobacterium ferrugineum*; B, Endosymbiont of *Achantomaeba*; C, *Acidovorax* sp.; D, *Achromobacter piechaudii*; E, *Microbacterium*; F, *Pseudomonas stutzeri*.

Figure S4 Rarefaction curves for the bacterial 16S rDNA gene sequences. The curves were plotted as Operational Phylogenetic Units (OPUs) at 97% sequence similarity *versus* the number of 16S rDNA sequences. For Nbs, both almost complete cloned 16S rDNA gene sequences and 16S rDNA gene partial sequences obtained in the metagenome survey, were considered.

Figures S5 and **S6** Neighbor-joining tree of proteobacterial (**S5**), and other (**S6**) 16S rDNA gene from N, Nbs, CN1 and CN2 communities this figure corresponds to Figures 1, 2 and 3, but in this case showing the internal topology of each OPU. As shown in the tree a large fraction of sequences affiliated with known families and clustered with branches represented by cultured microorganisms, typical inhabiting (non) contaminated environments, and thus they do not appear to be specific for the soil and enrichment samples herein investigated; however, a number of sequences were closely related to uncultured microorganisms mostly recovered from non contaminated soils (DQ256350 related to *Alkanindigenes*) as well as hydrocarbon contaminated soils (DQ297946 related to *Novosphingobium*), soils containing natural asphalts (EF157271 related to *Luteolibacter*), soil wastes from oil-shale chemical industry (EF540436 related to *Rhodobacteraceae*), wastewater digesters (EU104101 and FJ535514) and waters of petroleum reservoirs (AY570578 related to *Arcobacter* and AY570579) and mines (DQ469227 related to *Alcaligenaceae*) and chloroaromatic consortia (DQ286556 related to *Azospirillum*), to cite some.

Figure S7 Contigs length distribution of N, Nbs, CN1 and CN2 metagenomes. Graph represents the number of contigs with length ≤ 5000 bp (A), 5000-10000 bp (B), 10200-15000 (C) and 15200-25000 (D).

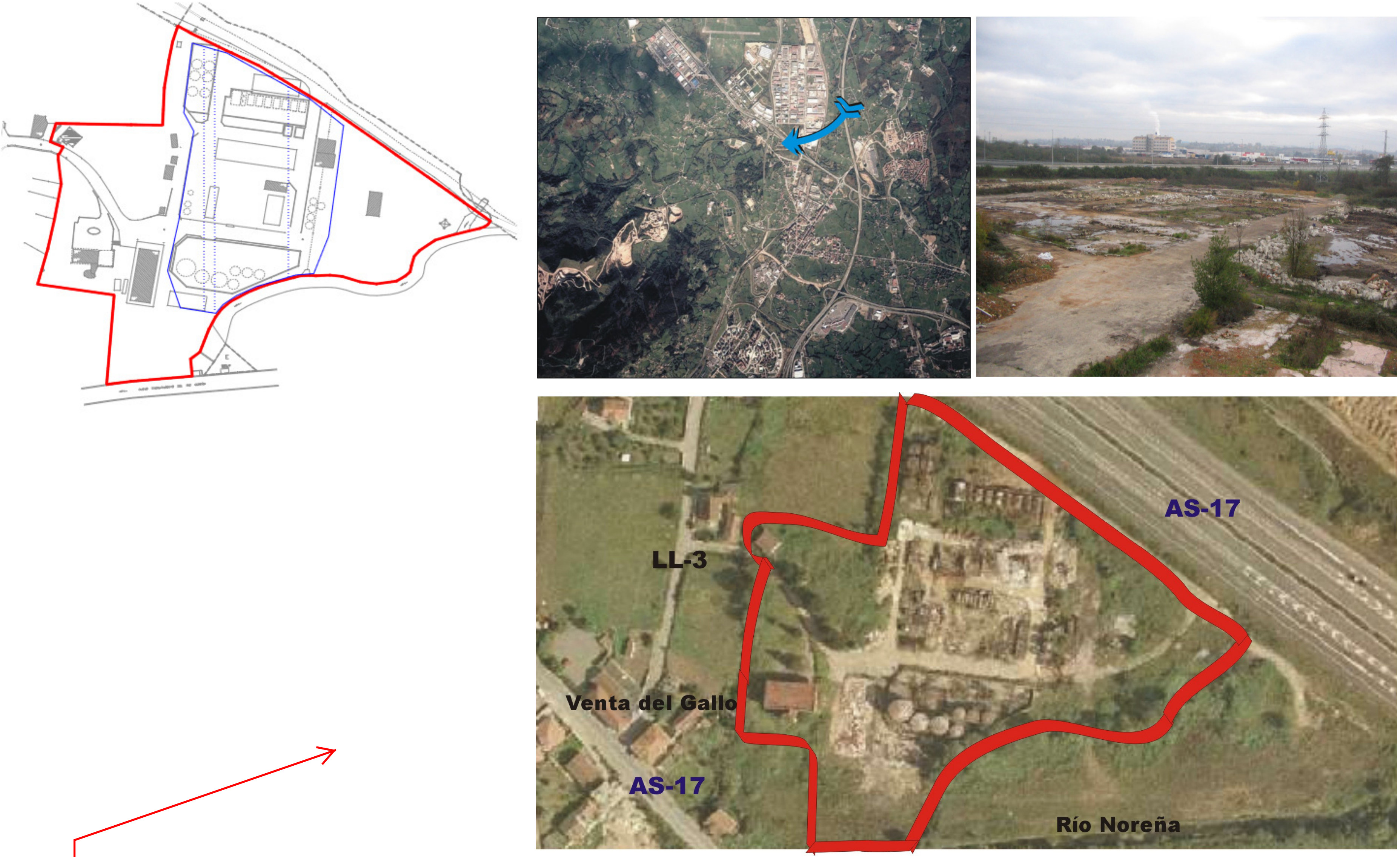


Figure S1

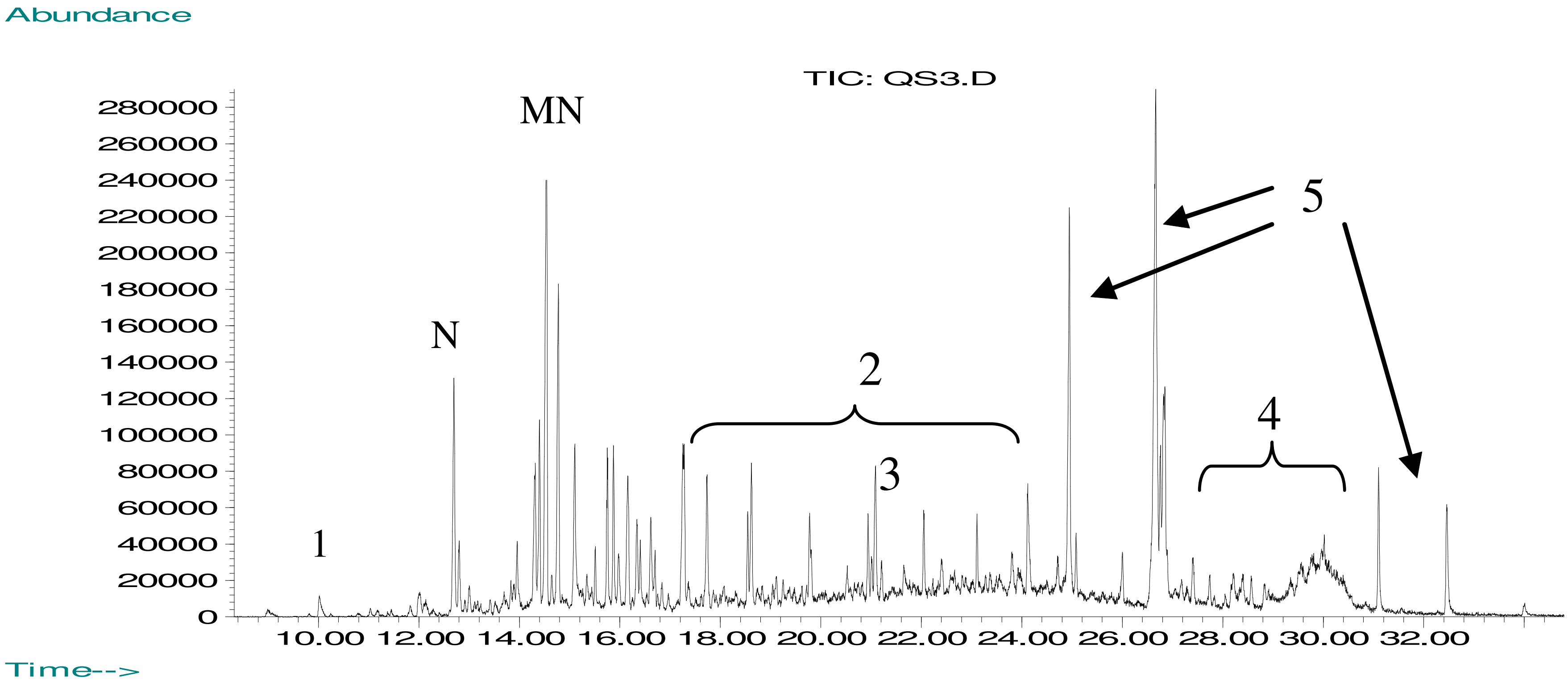

Abundance
TIC: QS3.D
280000
260000
240000
220000
200000
180000
160000
140000
120000
100000
80000
60000
40000
20000
0
MN
N
1
2
3
4
5
10.00
12.00
14.00
16.00
18.00
20.00
22.00
24.00
26.00
28.00
30.00
32.00
Time-->

Figure S2

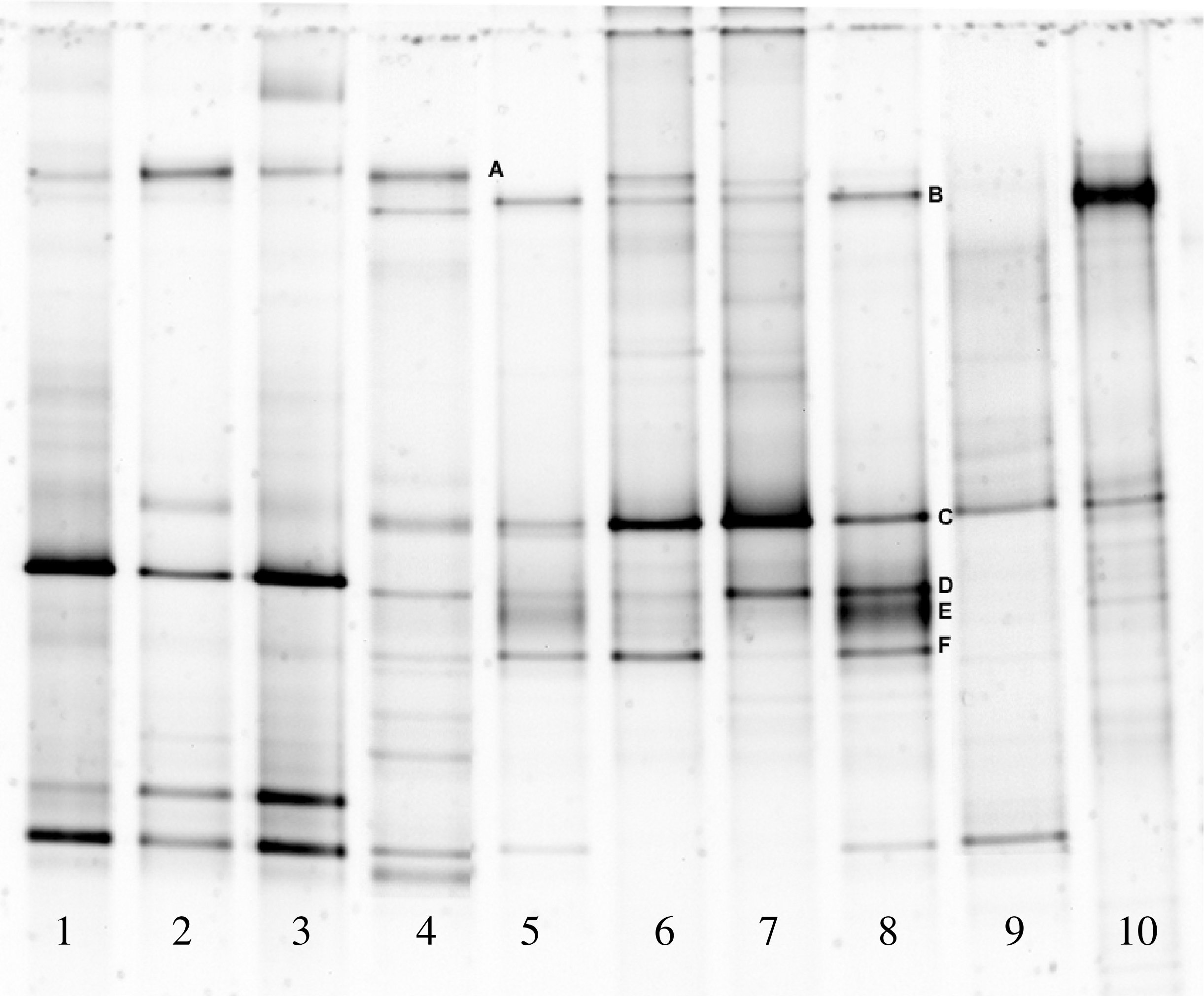


Figure S3

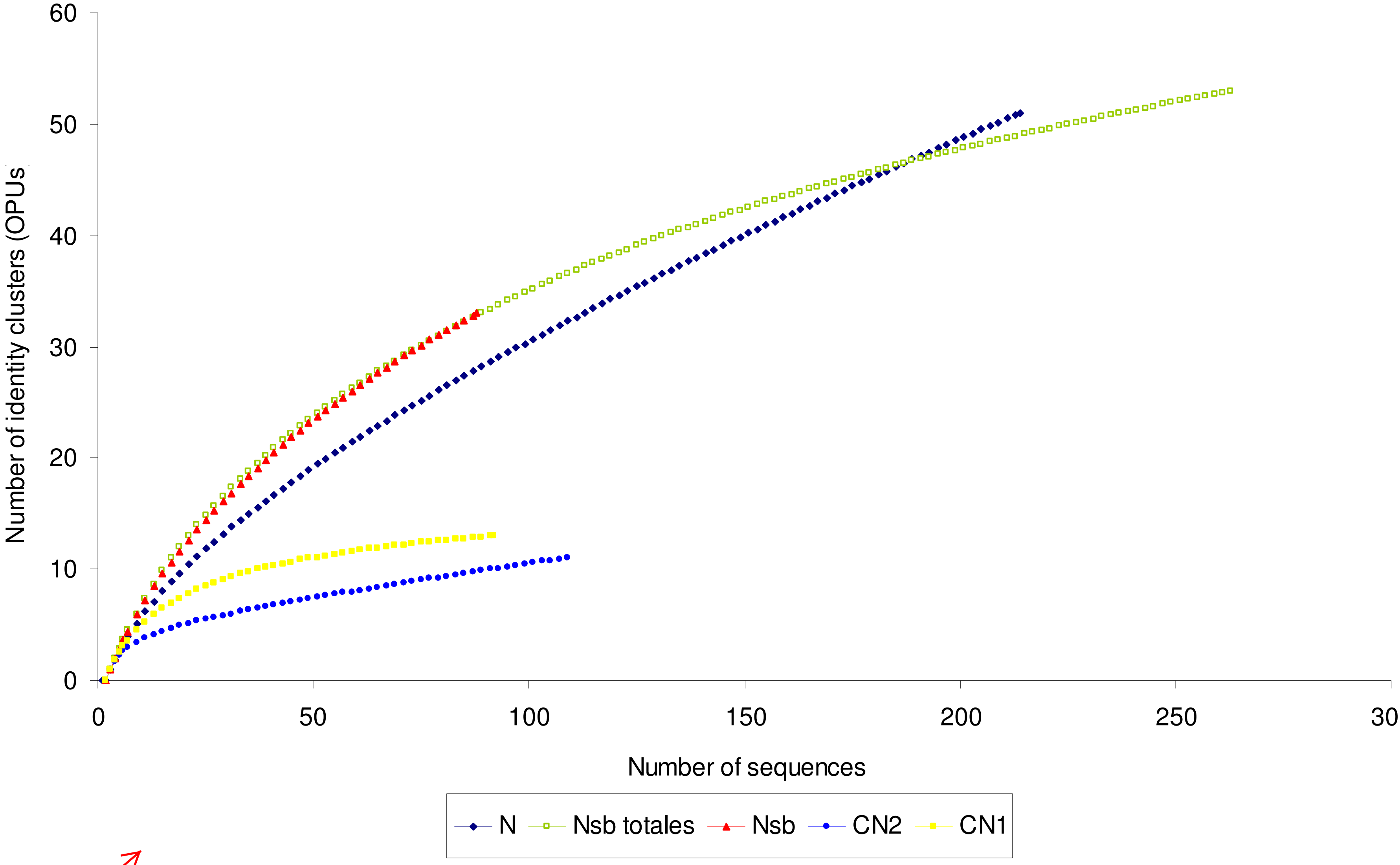

60
50
40
30
20
10
0
Number of identity clusters (OPUs
0
50
100
150
200
250
300
Number of sequences
N
Nsb totales
Nsb
CN2
CN1

Figure S4

Figure S5

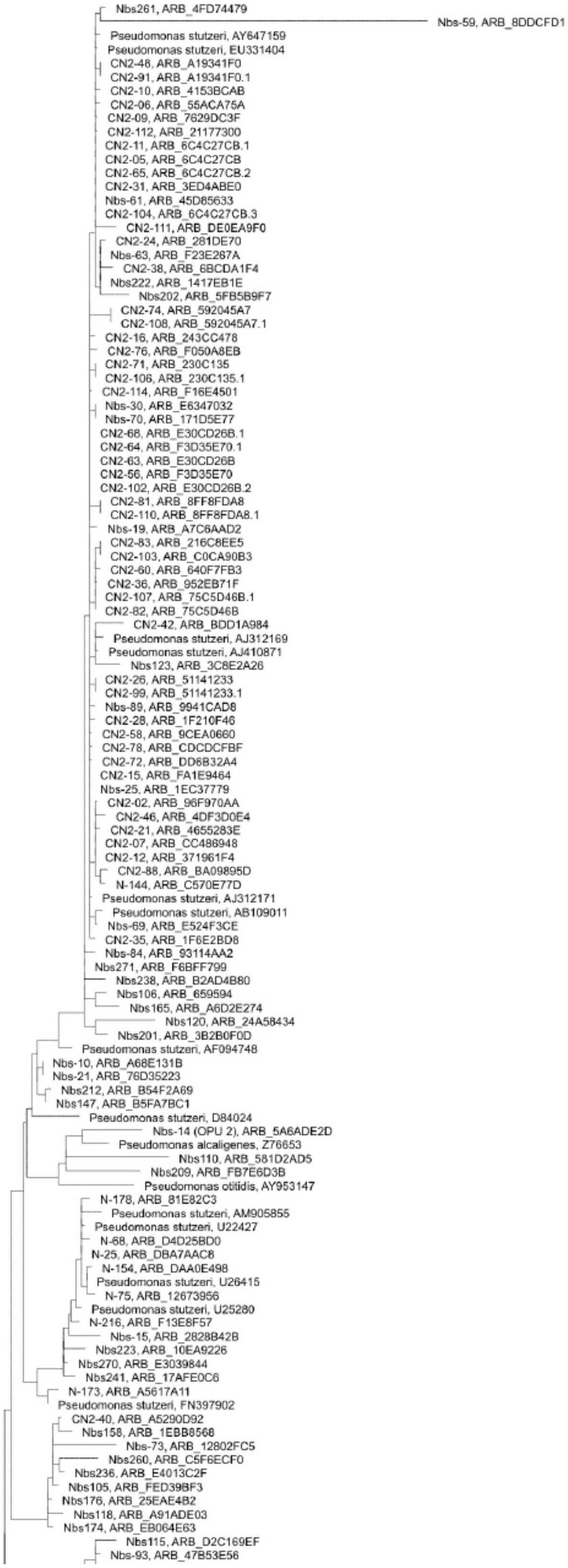

Nbs261, ARB_4FD74479
Nbs-59, ARB_8DDCFD1
Pseudomonas stutzeri, AY647159
Pseudomonas stutzeri, EU331404
CN2-48, ARB_A19341F0
CN2-91, ARB_A19341F0.1
CN2-10, ARB_4153BCAB
CN2-06, ARB_55ACA75A
CN2-09, ARB_7629DC3F
CN2-112, ARB_21177300
CN2-11, ARB_6C4C27CB.1
CN2-05, ARB_6C4C27CB
CN2-65, ARB_6C4C27CB.2
CN2-31, ARB_3ED4ABE0
Nbs-61, ARB_45D85633
CN2-104, ARB_6C4C27CB.3
CN2-111, ARB_DE0EA9F0
CN2-24, ARB_281DE70
Nbs-63, ARB_F23E267A
CN2-38, ARB_6BCDA1F4
Nbs222, ARB_1417EB1E
Nbs202, ARB_5FB5B9F7
CN2-74, ARB_592045A7
CN2-108, ARB_592045A7.1
CN2-16, ARB_243CC478
CN2-76, ARB_F050A8EB
CN2-71, ARB_230C135
CN2-106, ARB_230C135.1
CN2-114, ARB_F16E4501
Nbs-30, ARB_E6347032
Nbs-70, ARB_171D5E77
CN2-68, ARB_E30CD26B.1
CN2-64, ARB_F3D35E70.1
CN2-63, ARB_E30CD26B
CN2-56, ARB_F3D35E70
CN2-102, ARB_E30CD26B.2
CN2-81, ARB_8FF8FDA8
CN2-110, ARB_8FF8FDA8.1
Nbs-19, ARB_A7C6AAD2
CN2-83, ARB_216C8EE5
CN2-103, ARB_C0CA90B3
CN2-60, ARB_640F7FB3
CN2-36, ARB_952EB71F
CN2-107, ARB_75C5D46B.1
CN2-82, ARB_75C5D46B
CN2-42, ARB_BDD1A984
Pseudomonas stutzeri, AJ312169
Pseudomonas stutzeri, AJ410871
Nbs123, ARB_3C8E2A26
CN2-26, ARB_51141233
CN2-99, ARB_51141233.1
Nbs-89, ARB_9941CAD8
CN2-28, ARB_1F210F46
CN2-58, ARB_9CEA0660
CN2-78, ARB_CDCDCFBF
CN2-72, ARB_DD6B32A4
CN2-15, ARB_FA1E9464
Nbs-25, ARB_1EC37779
CN2-02, ARB_96F970AA
CN2-46, ARB_4DF3D0E4
CN2-21, ARB_4655283E
CN2-07, ARB_CC486948
CN2-12, ARB_371961F4
CN2-88, ARB_BA09895D
N-144, ARB_C570E77D
Pseudomonas stutzeri, AJ312171
Pseudomonas stutzeri, AB109011
Nbs-69, ARB_E524F3CE
CN2-35, ARB_1F6E2BD8
Nbs-84, ARB_93114AA2
Nbs271, ARB_F6BFF799
Nbs238, ARB_B2AD4B80
Nbs106, ARB_659594
Nbs165, ARB_A6D2E274
Nbs120, ARB_24A58434
Nbs201, ARB_3B2B0F0D
Pseudomonas stutzeri, AF094748
Nbs-10, ARB_A68E131B
Nbs-21, ARB_76D35223
Nbs212, ARB_B54F2A69
Nbs147, ARB_B5FA7BC1
Pseudomonas stutzeri, D84024
Nbs-14 (OPU 2), ARB_5A6ADE2D
Pseudomonas alcaligenes, Z76653
Nbs110, ARB_581D2AD5
Nbs209, ARB_FB7E6D3B
Pseudomonas otitidis, AY953147
N-178, ARB_81E82C3
Pseudomonas stutzeri, AM905855
Pseudomonas stutzeri, U22427
N-68, ARB_D4D25BD0
N-25, ARB_DBA7AAC8
N-154, ARB_DAA0E498
Pseudomonas stutzeri, U26415
N-75, ARB_12673956
Pseudomonas stutzeri, U25280
N-216, ARB_F13E8F57
Nbs-15, ARB_2828B42B
Nbs223, ARB_10EA9226
Nbs270, ARB_E3039844
Nbs241, ARB_17AFE0C6
N-173, ARB_A5617A11
Pseudomonas stutzeri, FN397902
CN2-40, ARB_A5290D92
Nbs158, ARB_1EBB8568
Nbs-73, ARB_12802FC5
Nbs260, ARB_C5F6ECF0
Nbs236, ARB_E4013C2F
Nbs105, ARB_FED39BF3
Nbs176, ARB_25EAE4B2
Nbs118, ARB_A91ADE03
Nbs174, ARB_EB064E63
Nbs115, ARB_D2C169EF
Nbs-93, ARB_47B53E56

Figure S5 cont. (1)

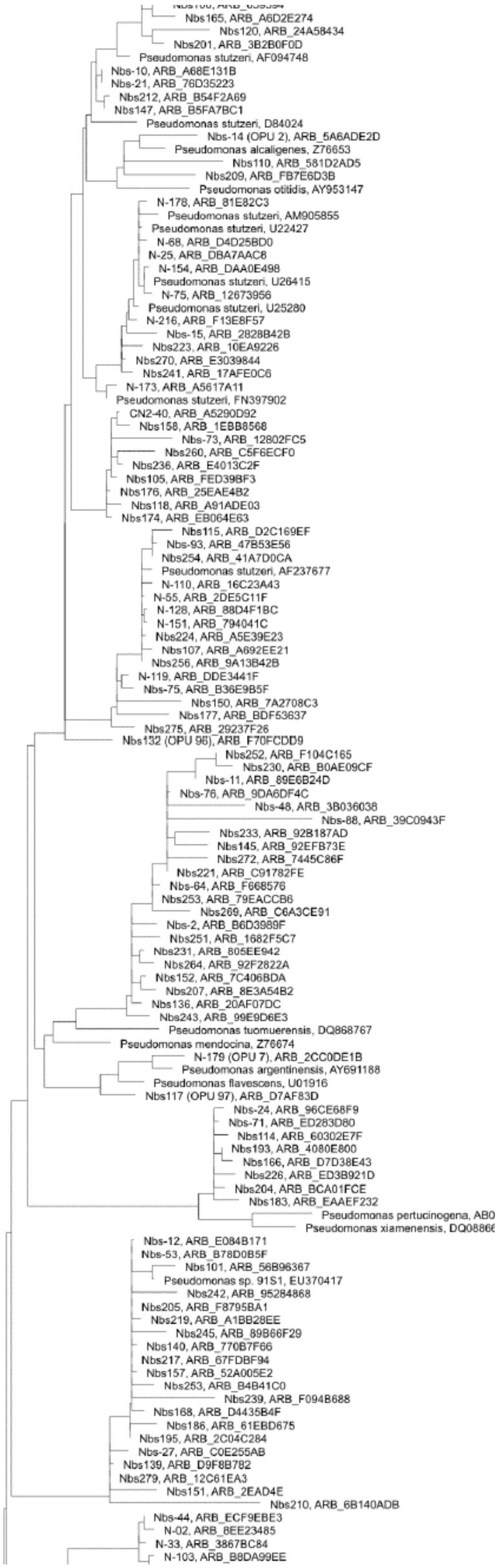


Figure S5 cont. (2)

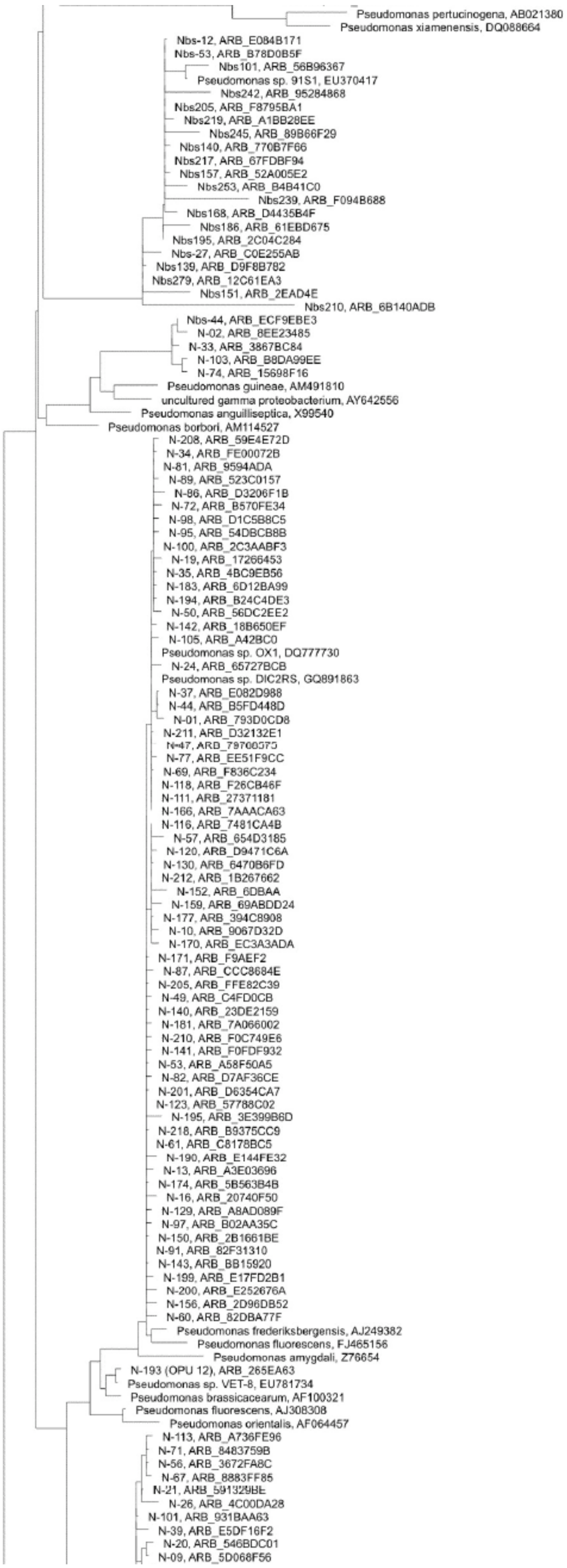


Figure S5 cont. (3)

Figure S5 cont. (4)

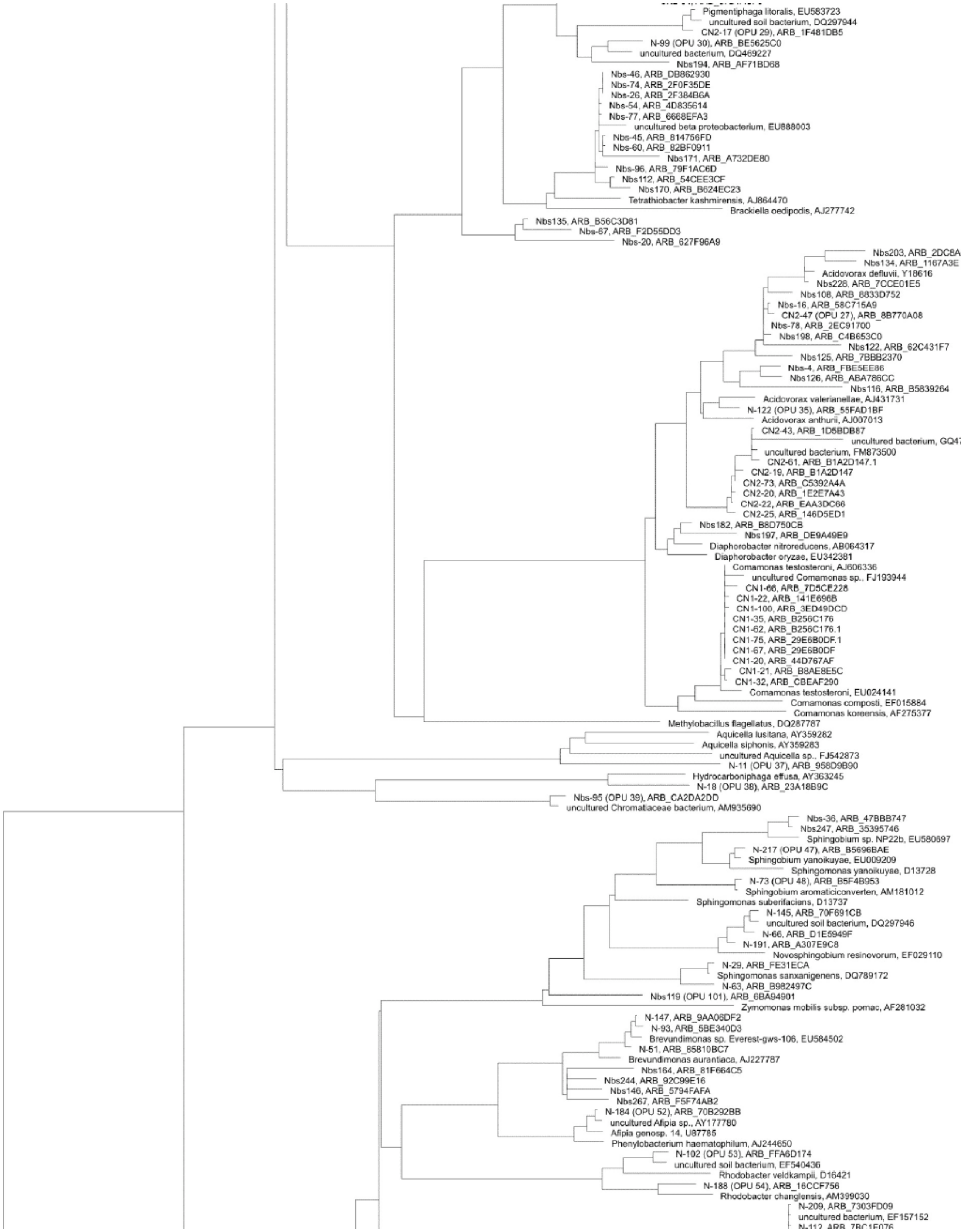
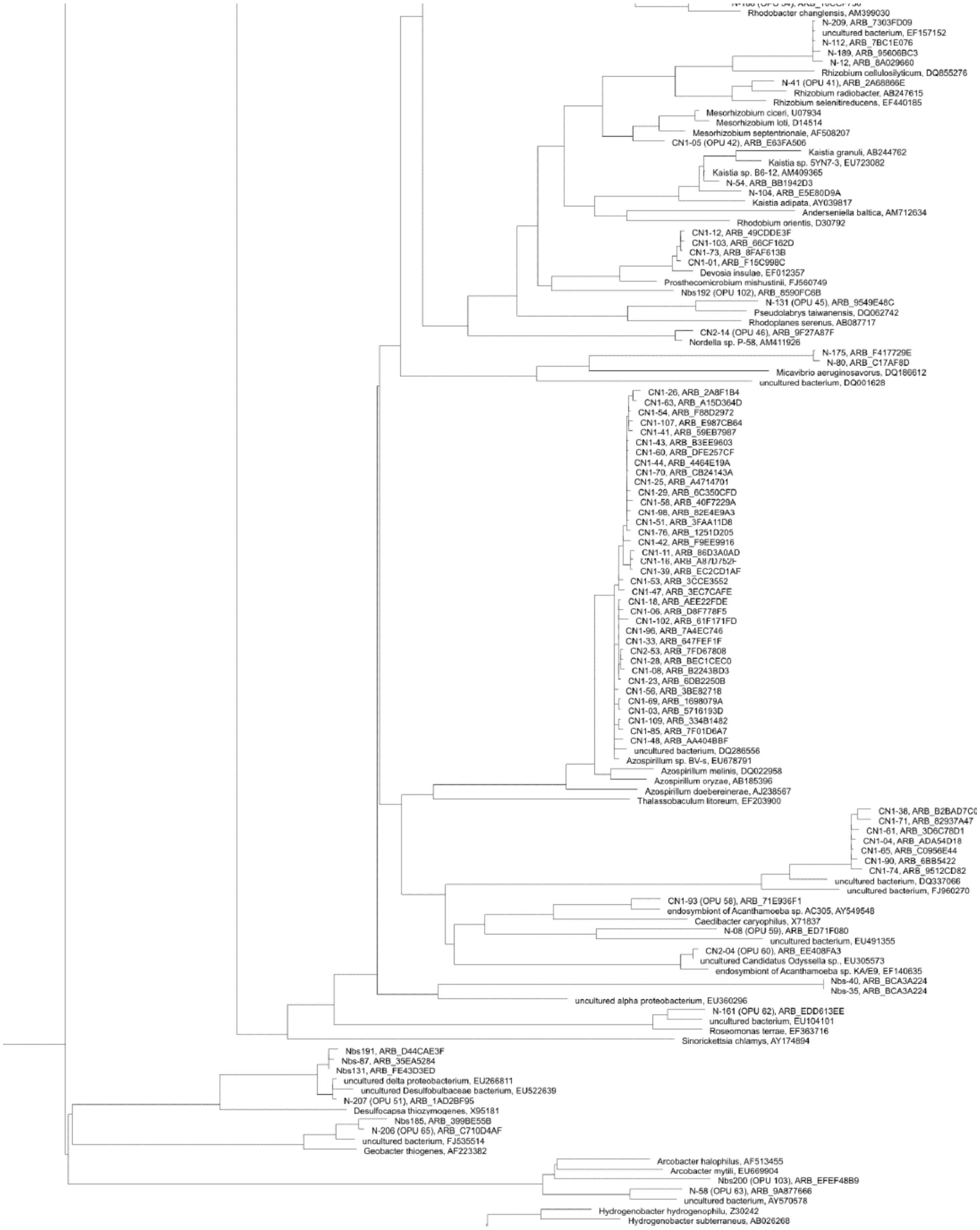


Figure S5 cont. (5)

Figure S5 cont. (6)



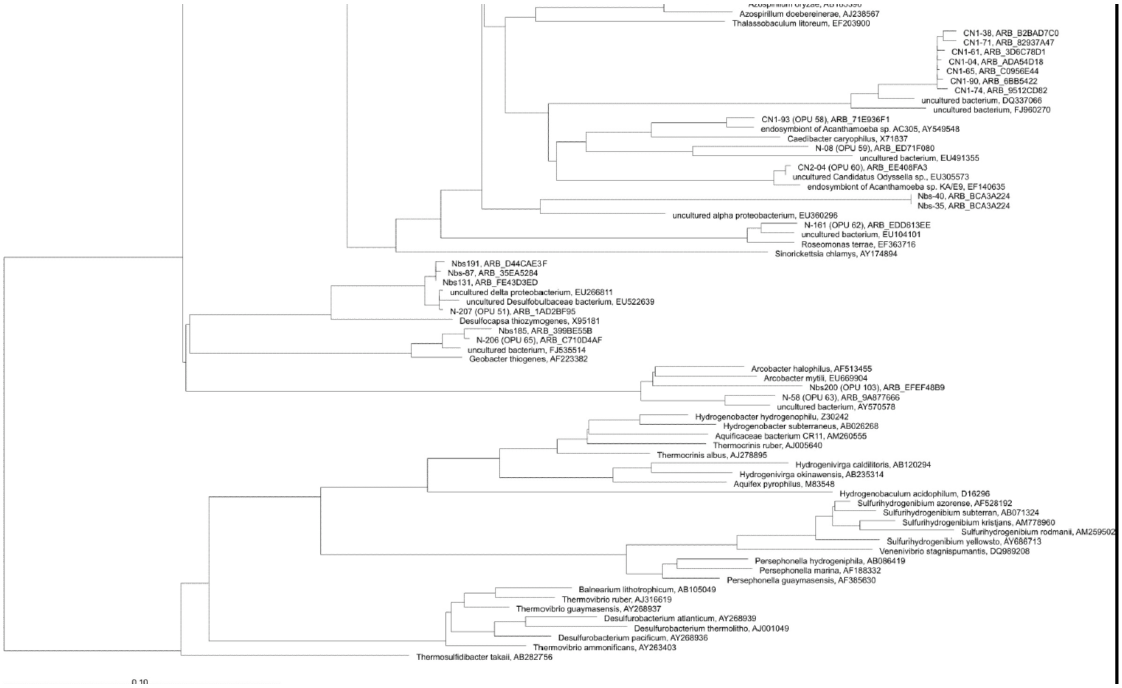
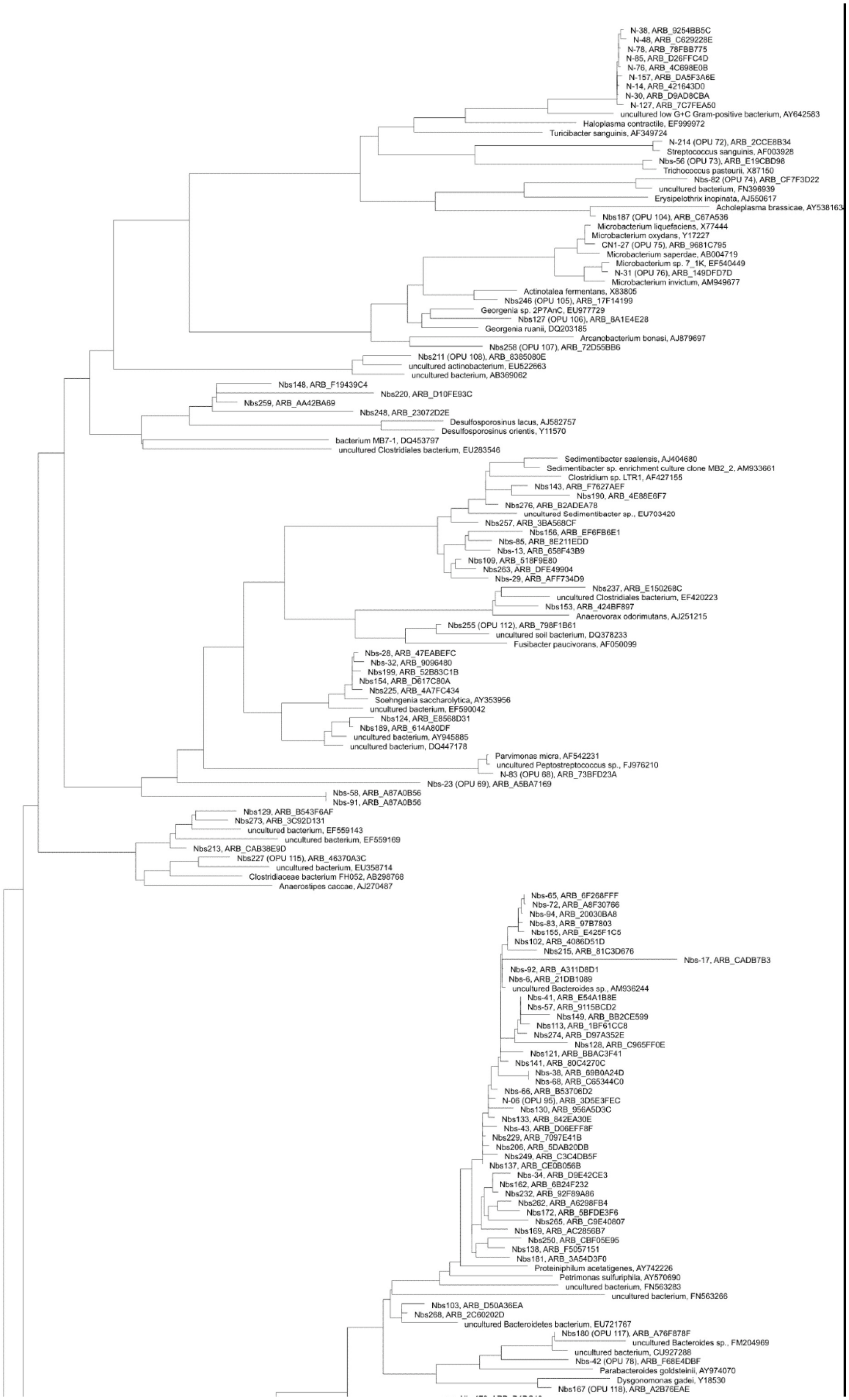


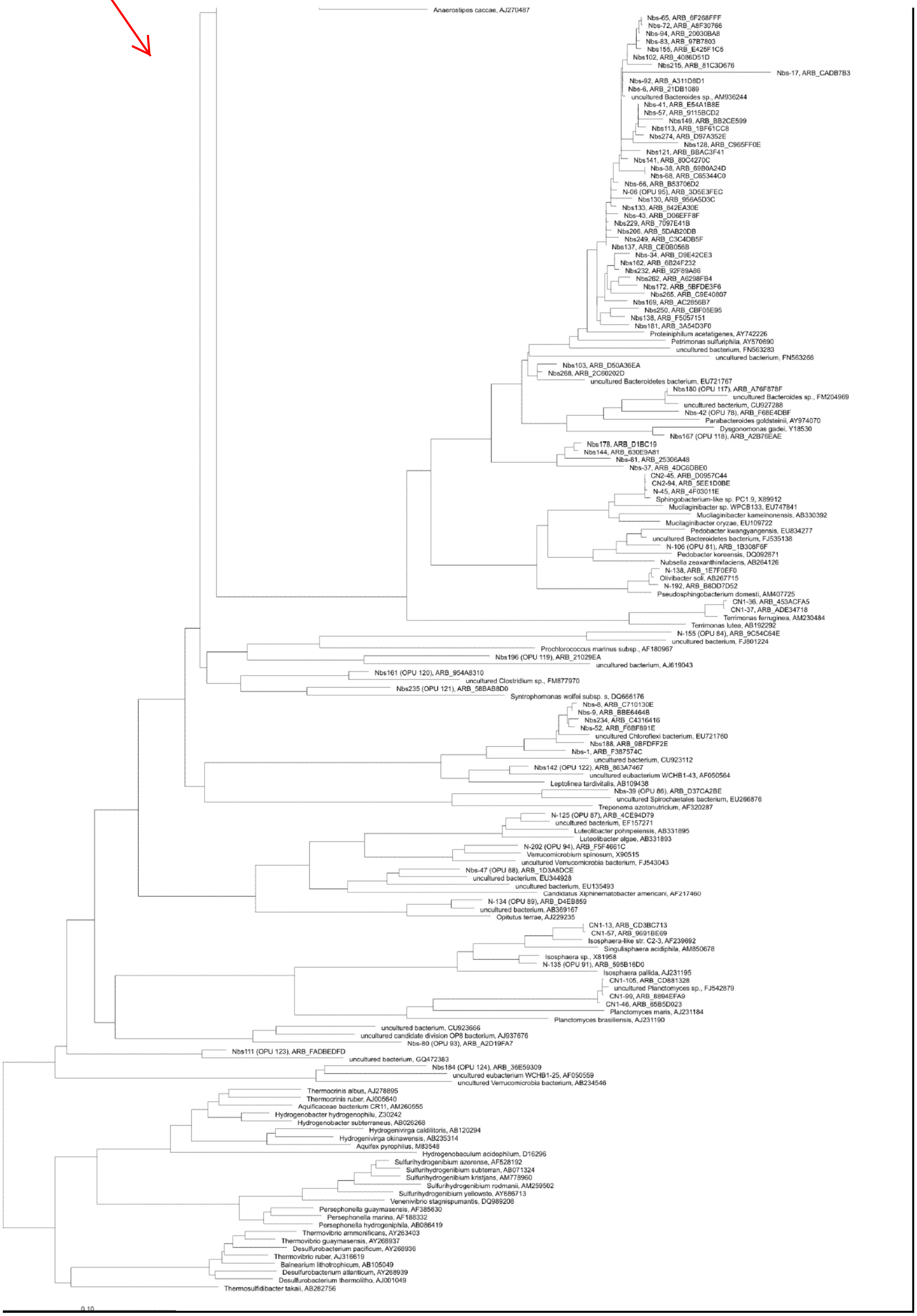
Figure S5 cont. (7)

Figure S6



N-38, ARB_9254BB5C
N-48, ARB_C629228E
N-78, ARB_78FBB775
N-85, ARB_D26FFC4D
N-76, ARB_4C698E0B
N-157, ARB_DA5F3A6E
N-14, ARB_421643D0
N-30, ARB_D9AD8CBA
N-127, ARB_7C7FEA50
uncultured low G+C Gram-positive bacterium, AY642583
Haloplasma contractile, EF999972
Turicibacter sanguinis, AF349724
N-214 (OPU 72), ARB_2CCE8B34
Streptococcus sanguinis, AF003928
Nbs-56 (OPU 73), ARB_E19CBD98
Trichococcus pasteurii, X87150
Nbs-82 (OPU 74), ARB_CF7F3D22
uncultured bacterium, FN396939
Erysipelothrix inopinata, AJ550617
Acholeplasma brassicae, AY538163
Nbs187 (OPU 104), ARB_C67A536
Microbacterium liquefaciens, X77444
Microbacterium oxydans, Y17227
CN1-27 (OPU 75), ARB_9681C795
Microbacterium saperdae, AB004719
Microbacterium sp. 7_1K, EF540449
N-31 (OPU 76), ARB_149DFD7D
Microbacterium invictum, AM949677
Actinotalea fermentans, X83805
Nbs246 (OPU 105), ARB_17F14199
Georgenia sp. 2P7AnC, EU977729
Nbs127 (OPU 106), ARB_8A1E4E28
Georgenia ruanii, DQ203185
Arcanobacterium bonasi, AJ879697
Nbs258 (OPU 107), ARB_72D55BB6
Nbs211 (OPU 108), ARB_8385080E
uncultured actinobacterium, EU522663
uncultured bacterium, AB369062
Nbs148, ARB_F19439C4
Nbs220, ARB_D10FE93C
Nbs259, ARB_AA42BA69
Nbs248, ARB_23072D2E
Desulfosporosinus lacus, AJ582757
Desulfosporosinus orientis, Y11570
bacterium MB7-1, DQ453797
uncultured Clostridiales bacterium, EU283546
Sedimentibacter saalensis, AJ404680
Sedimentibacter sp. enrichment culture clone MB2_2, AM933661
Clostridium sp. LTR1, AF427155
Nbs143, ARB_F7627AEF
Nbs190, ARB_4E88E6F7
Nbs276, ARB_B2ADEA78
uncultured Sedimentibacter sp., EU703420
Nbs257, ARB_3BA568CF
Nbs156, ARB_EF6FB6E1
Nbs-85, ARB_8E211EDD
Nbs-13, ARB_658F43B9
Nbs109, ARB_518F9E80
Nbs263, ARB_DFE49904
Nbs-29, ARB_AFF734D9
Nbs237, ARB_E150268C
uncultured Clostridiales bacterium, EF420223
Nbs153, ARB_424BF897
Anaerovorax odorimutans, AJ251215
Nbs255 (OPU 112), ARB_798F1B61
uncultured soil bacterium, DQ378233
Fusibacter paucivorans, AF050099
Nbs-28, ARB_47EABEFC
Nbs-32, ARB_9096480
Nbs199, ARB_52B83C1B
Nbs154, ARB_D617C80A
Nbs225, ARB_4A7FC434
Soehngenia saccharolytica, AY353956
uncultured bacterium, EF590042
Nbs124, ARB_E8568D31
Nbs189, ARB_614A80DF
uncultured bacterium, AY945885
uncultured bacterium, DQ447178
Parvimonas micra, AF542231
uncultured Peptostreptococcus sp., FJ976210
N-83 (OPU 68), ARB_73BFD23A
Nbs-23 (OPU 69), ARB_A5BA7169
Nbs-58, ARB_A87A0B56
Nbs-91, ARB_A87A0B56
Nbs129, ARB_B543F6AF
Nbs273, ARB_3C92D131
uncultured bacterium, EF559143
uncultured bacterium, EF559169
Nbs213, ARB_CAB38E9D
Nbs227 (OPU 115), ARB_46370A3C
uncultured bacterium, EU358714
Clostridiaceae bacterium FH052, AB298768
Anaerostipes caccae, AJ270487
Nbs-65, ARB_6F268FFF
Nbs-72, ARB_A8F30766
Nbs-94, ARB_20030BA8
Nbs-83, ARB_97B7803
Nbs155, ARB_E425F1C5
Nbs102, ARB_4086D51D
Nbs215, ARB_81C3D676
Nbs-17, ARB_CADB7B3
Nbs-92, ARB_A311D8D1
Nbs-6, ARB_21DB1089
uncultured Bacteroides sp., AM936244
Nbs-41, ARB_E54A1B8E
Nbs-57, ARB_9115BCD2
Nbs149, ARB_BB2CE599
Nbs113, ARB_1BF61CC8
Nbs274, ARB_D97A352E
Nbs128, ARB_C965FF0E
Nbs121, ARB_BBAC3F41
Nbs141, ARB_80C4270C
Nbs-38, ARB_69B0A24D
Nbs-68, ARB_C65344C0
Nbs-66, ARB_B53706D2
N-06 (OPU 95), ARB_3D5E3FEC
Nbs130, ARB_956A5D3C
Nbs133, ARB_842EA30E
Nbs-43, ARB_D06EFF8F
Nbs229, ARB_7097E41B
Nbs206, ARB_5DAB20DB
Nbs249, ARB_C3C4DB5F
Nbs137, ARB_CE0B056B
Nbs-34, ARB_D9E42CE3
Nbs162, ARB_6B24F232
Nbs232, ARB_92F89A86
Nbs262, ARB_A6298FB4
Nbs172, ARB_5BFDE3F6
Nbs265, ARB_C9E40807
Nbs169, ARB_AC2856B7
Nbs250, ARB_CBF05E95
Nbs138, ARB_F5057151
Nbs181, ARB_3A54D3F0
Proteiniphilum acetatigenes, AY742226
Petrimonas sulfuriphila, AY570690
uncultured bacterium, FN563283
uncultured bacterium, FN563266
Nbs103, ARB_D50A36EA
Nbs268, ARB_2C60202D
uncultured Bacteroidetes bacterium, EU721767
Nbs180 (OPU 117), ARB_A76F878F
uncultured Bacteroides sp., FM204969
uncultured bacterium, CU927288
Nbs-42 (OPU 78), ARB_F68E4DBF
Parabacteroides goldsteinii, AY974070
Dysgonomonas gadei, Y18530
Nbs167 (OPU 118), ARB_A2B76EAE

Figure S6 cont. (1)



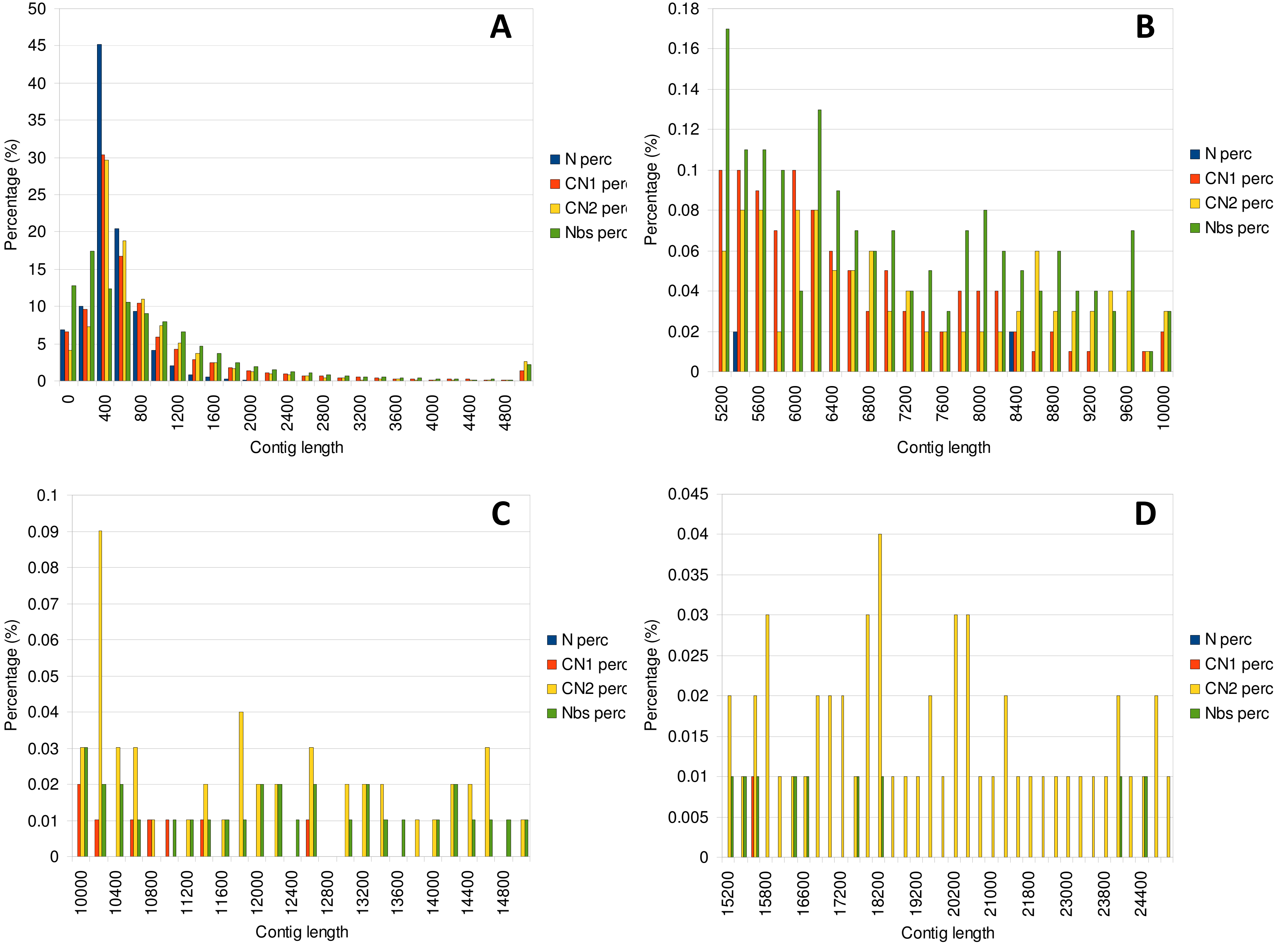

A
50
45
40
35
30
25
20
15
10
5
0
Percentage (%)
0
400
800
1200
1600
2000
2400
2800
3200
3600
4000
4400
4800
Contig length
N perc
CN1 perc
CN2 perc
Nbs perc
B
0.18
0.16
0.14
0.12
0.1
0.08
0.06
0.04
0.02
0
Percentage (%)
5200
5600
6000
6400
6800
7200
7600
8000
8400
8800
9200
9600
10000
Contig length
N perc
CN1 perc
CN2 perc
Nbs perc
C
0.1
0.09
0.08
0.07
0.06
0.05
0.04
0.03
0.02
0.01
0
Percentage (%)
10000
10400
10800
11200
11600
12000
12400
12800
13200
13600
14000
14400
14800
Contig length
N perc
CN1 perc
CN2 perc
Nbs perc
D
0.045
0.04
0.035
0.03
0.025
0.02
0.015
0.01
0.005
0
Percentage (%)
15200
15800
16600
17200
18200
19200
20200
21000
21800
23000
23800
24400
Contig length
N perc
CN1 perc
CN2 perc
Nbs perc

Figure S7