

UC Merced

UC Merced Previously Published Works

Title

Horizontal gene transfer of the Mer operon is associated with large effects on the transcriptome and increased tolerance to mercury in nitrogen-fixing bacteria

Permalink

<https://escholarship.org/uc/item/4371r81q>

Journal

BMC Microbiology, 24(1)

ISSN

1471-2180

Authors

Bhat, Aditi

Sharma, Reena

Desigan, Kumaran

et al.

Publication Date

2024-07-01

DOI

10.1186/s12866-024-03391-5

Copyright Information

This work is made available under the terms of a Creative Commons Attribution License, available at <https://creativecommons.org/licenses/by/4.0/>

Peer reviewed

RESEARCH

Open Access



Horizontal gene transfer of the Mer operon is associated with large effects on the transcriptome and increased tolerance to mercury in nitrogen-fixing bacteria

Aditi Bhat¹, Reena Sharma¹, Kumaran Desigan¹, M. Mercedes Lucas², Ankita Mishra³, Robert M. Bowers⁴, Tanja Woyke⁴, Brendan Epstein⁵, Peter Tiffin⁵, José J. Pueyo² and Tim Paape^{3,6*}

Abstract

Background Mercury (Hg) is highly toxic and has the potential to cause severe health problems for humans and foraging animals when transported into edible plant parts. Soil rhizobia that form symbiosis with legumes may possess mechanisms to prevent heavy metal translocation from roots to shoots in plants by exporting metals from nodules or compartmentalizing metal ions inside nodules. Horizontal gene transfer has potential to confer immediate de novo adaptations to stress. We used comparative genomics of high quality de novo assemblies to identify structural differences in the genomes of nitrogen-fixing rhizobia that were isolated from a mercury (Hg) mine site that show high variation in their tolerance to Hg.

Results Our analyses identified multiple structurally conserved merA homologs in the genomes of *Sinorhizobium medicae* and *Rhizobium leguminosarum* but only the strains that possessed a Mer operon exhibited 10-fold increased tolerance to Hg. RNAseq analysis revealed nearly all genes in the Mer operon were significantly up-regulated in response to Hg stress in free-living conditions and in nodules. In both free-living and nodule environments, we found the Hg-tolerant strains with a Mer operon exhibited the fewest number of differentially expressed genes (DEGs) in the genome, indicating a rapid and efficient detoxification of Hg from the cells that reduced general stress responses to the Hg-treatment. Expression changes in *S. medicae* while in bacteroids showed that both rhizobia strain and host-plant tolerance affected the number of DEGs. Aside from Mer operon genes, *nif* genes which are involved in nitrogenase activity in *S. medicae* showed significant up-regulation in the most Hg-tolerant strain while inside the most Hg-accumulating host-plant. Transfer of a plasmid containing the Mer operon from the most tolerant strain to low-tolerant strains resulted in an immediate increase in Hg tolerance, indicating that the Mer operon is able to confer hyper tolerance to Hg.

Conclusions Mer operons have not been previously reported in nitrogen-fixing rhizobia. This study demonstrates a pivotal role of the Mer operon in effective mercury detoxification and hypertolerance in nitrogen-fixing rhizobia. This

*Correspondence:

Tim Paape

timothy.paape@usda.gov; tim.paape@ag.tamu.edu

Full list of author information is available at the end of the article



This is a U.S. Government work and not under copyright protection in the US; foreign copyright protection may apply 2024. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

finding has major implications not only for soil bioremediation, but also host plants growing in mercury contaminated soils.

Keywords Mercury reductase, Rhizobia, Mer operon, Horizontal gene transfer

Introduction

Heavy metal stress can have severe consequences for both plant and bacterial cells which exerts strong selective pressure on plant and microbial populations. Mercury (Hg) is a pervasive global pollutant [1] where the predominant sources of environmental heavy metal contamination, besides natural accumulation or volcanic eruptions, are mines, foundries, and smelters [2, 3]. Additional sources of heavy metals can be from pesticide and herbicides that are applied to crops [4, 5]. Bioaccumulation in the food chain is commonly in the form of methylmercury (CH_3Hg^+), which is highly toxic to humans and other organisms [6, 7], but inorganic mercury in the form of Hg^{2+} is also highly toxic to most organisms. The main difference between both lies in their sources as methylmercury, which is most commonly derived from anaerobic microorganisms [8], and inorganic mercury which typically has an atmospheric or soil deposition source (e.g. mines), or enters through biological systems as a byproduct of demethylation of CH_3Hg^+ [9].

Molecular adaptations to mercury stress have been found predominantly in bacteria [10, 11]. The most well characterized genetic mechanism responsible for detoxifying mercury in bacteria is the Mer operon that evolved in geothermal bacteria, but is often found in aerobic, anaerobic, aquatic, and terrestrial bacteria as well [12]. The Mer operon has the capacity to convert, transport, and detoxify both methylmercury (using MerB) and inorganic mercury (using MerA) along with a suite of membrane transporter genes that are typically tightly linked in the operon [11]. Central to the Mer operon is MerA, a flavin-dependent NAD(P)-disulfide (FAD) oxidoreductase that converts Hg^{2+} to Hg^0 , which can be released from the cell as a gas. This differs from most other heavy metal detoxification systems in microbes which primarily use ATPase based transport systems to move metal ions across cell membranes, or to compartmentalize toxic ions in extracellular structures [13–15]. The Mer operon also contains several proteins that span the inner membrane transport proteins such as MerC, MerE, MerF, MerG, MerP, and MerT [11]. In addition to the known Hg^{2+} ion transporters, glutathione reductase genes (sometimes called MerK, [16]) can be found in Mer operons as glutathione can potentially transfer Hg^{2+} to MerA [17]. The overall expression of Mer genes is regulated by MerR, acting as a transcriptional repressor or activator in the absence and presence of Hg^{2+} . In some strains, when the MerA gene is accompanied by MerB, which encodes for an organomercury lyase that catalyzes methylmercury

into Hg^{2+} and methane, Hg^{2+} is passed along to MerA where it is transformed to Hg^0 . The MerA and MerB genes were found to be at low frequencies (8% and 2% respectively) among broad taxonomic surveys of bacteria genomes [12], and the difference in frequencies indicate that methyl-mercury detoxification via MerB gene is less prevalent in bacterial genomes than detoxification of ionic mercury.

While the Mer operons originated from micro-organisms that survive in geothermal environments [11], MerA and MerB genes are found in broad taxonomic groups. Christakis et al. [12] have done the most exhaustive study of the diversity of the MerAB complex in bacteria using phylogenetic approaches of publicly available genomic data and found the presence of these genes in a wide variety of terrestrial bacteria, including soil-borne, nitrogen-fixing rhizobia in the class α -proteobacteria which form symbiosis with plant roots. Some of the most studied nitrogen-fixing bacteria are *Sinorhizobium medicae* (*syn. Ensifer*) and *S. meliloti* due to their symbiotic interactions with the model legume species *Medicago truncatula* and alfalfa (*M. sativa*) [18, 19], and *Rhizobium leguminosarum* bv. *trifolii* which mainly interacts with clovers, but also peas [20, 21]. While the majority of molecular research in legume host-plants and symbiotic rhizobia has focused on nitrogen fixation mechanisms, which provide obvious benefits to both partners through nitrogen and carbon exchange, more recognition of overlapping stress responses in rhizobia and host plant signaling has been realized to be essential for establishing symbiosis and improved resilience in legume-rhizobia systems [22]. Symbiotic interactions between legume host plants and nitrogen-fixing rhizobia have profound impacts on soil quality and plant health, and soil microbes in symbiosis with legumes have potential for bioremediation of toxic soils [15, 23], particularly if one or both symbiotic partners becomes adapted to a novel stress condition such as heavy metals in the environment.

Tolerance to soils contaminated with toxic heavy metals depends on molecular adaptations that allowed rhizobia to colonize contaminated environments, or adaptations that arose de novo through mutation in rhizobia living in the contaminated soils. Because free-living bacteria have such large effective population sizes, adaptive evolution through natural selection on mutations that confer higher tolerance is expected to be strong [24]. Moreover, bacterial genome evolution can also occur through horizontal gene transfer (HGT) between strains or between species, providing a source of rapid

adaptation to heavy metal stress [25, 26]. Indeed, nitrogen-fixing rhizobia strains isolated from toxic mine sites have revealed they adapted to increases in tolerance to heavy metals, typically through enhanced gene expression or presence-absence variation among strains [13, 14, 27, 28]. Transcriptomics studies have identified candidate genes that respond to the predominant heavy metals in mining sites by enhanced ion transport mechanisms (e.g., ATPase efflux activity using *cadA*), but genome wide studies in rhizobia have been mostly focused on adaptations to Cd, Cu, Pb and Zn [15, 29–31]. Studies of adaptive evolution and tolerance to Hg in rhizobia have been limited to a small number of genes [23, 32], rather than whole genome evolution or genome-wide responses in the transcriptome.

The Almadén mine region in Spain contains some of the greatest concentrations of cinnabar and mercury anywhere on Earth [33]. A large collection of *Sinorhizobium medicae* and *Rhizobium leguminosarum* isolated from root nodules of *Medicago* and *Trifolium* host plants growing at Almadén was measured for tolerance to various heavy metals [28]. Minimum inhibitory concentrations (MIC) showed that the most Hg-tolerant strains had 10-fold higher tolerance than non-Hg-tolerant strains from Almadén, and the increase in tolerance was attributed to *cis*-regulation of *merA1* and *merA2* [32], the two mercury reductase A homologs not associated with a Mer operon in the genome. In this paper, we will capitalize genes on the Mer operon (e.g., MerA) and homologous mercury reductase A genes will be indicated using lowercase gene names (e.g., *merA*) to distinguish the difference between the operon and non-operon homologs. The Arregui et al., [32] study used qPCR to quantify the regulatory changes in two rhizobia strains with different tolerance to Hg, focusing on *merA* genes in response to mercury in free-living conditions and in nodules. However, the genomes of these strains were not sequenced and other potential candidate genes or loci that contributed to hypertolerance to Hg could not be identified.

In α -proteobacteria, the class containing the most studied nitrogen-fixing bacteria including the genera *Bradyrhizobium*, *Mesorhizobium*, *Rhizobium*, and *Sinorhizobium*, all possess *merA* genes. However, most of the sequenced strains lack a Mer operon and these *merA* genes exist in the genome independent of any Mer operon. Very little is known about Mer operons in nitrogen fixing rhizobia or whether the *merA* genes contribute to Hg-tolerance through gene regulation. Furthermore, comparisons of secondary and tertiary protein structures of MerA and *merA* homologs in rhizobia genomes could provide insight into the conservation of function and whether the homologs provide redundant or elevated tolerance due to additive or multiplied gene expression in genomes that possess MerA, *merA1* and *merA2*.

In addition to heavy metal detoxification responses in rhizobia, it is also important to identify regulatory changes in nitrogen fixation genes (e.g. *nif*) that would potentially limit nitrogen export to host-plants during symbiosis if these genes are impacted by heavy metals [34]. Expression of the *nif* gene cluster by rhizobia controls the production of nitrogenase, the necessary enzyme for converting atmospheric nitrogen N₂ to ammonia (NH₃), that can be used as a macronutrient by the host plant. In nodules, the environment for rhizobia changes from purely aerobic in their free-living conditions to a much more anaerobic, oxygen limited environment, inside of nodules. Because heavy metal stress is known to trigger oxidative stress responses in soil microbes including rhizobia [30, 35–38], and in *Medicago truncatula* nodules which contain symbiotic *Sinorhizobium* [39], heavy metal stress on rhizobia may be even more severe during symbiosis due to the combined metal and oxygen stresses [40].

The Almadén rhizobia strains provide a unique opportunity for comparative genomics and transcriptomics due to known quantitative differences in their tolerance to Hg and other heavy metals. We sequenced the genomes of 9 *S. medicae* and 5 *R. leguminosarum* strains from the Almadén mine, using PacBio long-read sequencing to identify structural genomic differences between low-tolerant (LT) and hypertolerant (HT) strains. Using RNAseq in *S. medicae*, we identified gene expression differences between strains varying in their Hg tolerance in both free-living conditions and inside of nodules. The comparative genome analyses and differential expression results indicated the presence of complete Mer operons in Hg-tolerant strains, which we claim is the main determinant of HT rhizobia strains, despite the presence of generic *merA* genes in all *Sinorhizobium* and *Rhizobium* genomes. To test this hypothesis, we transferred a plasmid containing the Mer operon from the most tolerant *S. medicae* strain into two LT *S. medicae* strains and our findings demonstrated that HGT can instantly confer tolerance to Hg.

Methods

Isolation and DNA extraction from rhizobia

The rhizobia used in this study included 33 *Sinorhizobium medicae* and 26 *Rhizobium leguminosarum* bv. *trifolii* strains, which were isolated from nodules of *Medicago* and *Trifolium* species growing at different locations near the abandoned Almadén mercury mine in Spain (See [28, 41] for specific information regarding each strain including their tolerance to various metals). Their tolerance was assessed using minimum inhibitory concentrations (MIC) over ranges of 25–250 μ M [28]. Among these, we selected 9 *S. medicae* and 5 *R. leguminosarum* strains (Supplementary Data 1, Table S1) that

showed a broad range in Hg tolerance. Rhizobia strains that were grown for DNA isolation for whole genome sequencing, were cultivated in liquid TY medium at 28°C in a shaking incubator until they reached an OD₆₀₀ of 0.7. Total DNA was extracted using the Machery-Nagel NucleoSpin Microbial DNA Mini kit for DNA from microorganisms (Item number: 740235.50). During the procedure, bacteria cells were lysed using a Qiagen TissueLyser bead mill for 5–8 min.

Genome sequencing, assembly, and annotation

Genomic sequence data were generated using a Pacific Biosciences (PacBio) RS II sequencer at the University of Minnesota Genomics Center with one PacBio single molecule real time (SMRT) cell per strain. Genomes were assembled using hgap version 4.0 [42] in the SMRT-link Server software (version 7) with default assembly parameter, using a similar pipeline described in [43]. Subsequently, the genomes of almost all strains were circularized and contigs were verified using Circulator version 1.5.5 [44] <https://github.com/sanger-pathogens/circulator>. To annotate genomes we used RASTtk <https://rast.nmpdr.org> [45] using default settings. For the strains used for RNAseq, we also annotated the genome using the NCBI Prokaryotic Genome Annotation Pipeline (PGAP https://www.ncbi.nlm.nih.gov/genome/annotation_prok/) [46] to use these annotations to link GO-terms for GO-enrichment analysis. The annotations used throughout this manuscript are based on RASTtk, as this method was able to identify mercuric ion reductases better than PGAP. Protein (amino acid) sequences were used in subsequent phylogenetic analyses and for running OrthoFinder [47] to label homologous genes used the RASTtk annotations. We included PGAP annotations of strain *S. medicae* AMp08 in the OrthoFinder analysis to link all annotations to a common strain. The AMp08 strain was used because it has a Mer operon on an accessory plasmid and the genome was completely assembled and circularized into four complete chromosomes and plasmids.

Synteny analysis of whole genomes and mer operon and phylogenetic analysis of merA homologs

To check for synteny between focal strains with previously published reference strains of *S. medicae*, we used Sibelia 2.2.1 [48] using the command `./Sibelia -s loose -q`. We aligned the two Hg tolerant Almadén strains AMp08 and SMp01 strains to *S. medicae* WSM419 to assign the chromosomes and pSym plasmids according to [19]. For the *R. leguminosarum* strain Stf07, we aligned the *R. leguminosarum* WSM1325 genome to assign chromosomes and accessory plasmids [20]. We used OrthoFinder [47] to cluster genes in order to obtain orthogroups that can connect the genes from each independent annotation in

this study and other publicly available genomes. For the synteny analyses, we created orthologue tables using OrthoFinder [47]. IMG annotations (cog, ko and pfam) were mapped onto the orthologue tables, and each gene annotation contains the gene start, end, and strand direction information for each gene. Additionally, we used ortho groups to assign a consensus annotation when a given protein was not annotated as 100% the same annotation or to fill genes of unknown function when a consensus annotation was available. Note, these situations are rare, but the orthologue group clustering helps to make sense of either missing or misannotated genes. To produce the synteny plot displayed in Fig. S4, we used the R package gggenomes (<https://github.com/thackl/gggenomes>). We could find MerA homologs that possessed flanking MerT, MerC, MerP and MerB genes for the synteny analysis of the Mer operons in our dataset. We also built two separate phylogenies within this pipeline, individual gene trees, and species trees based on a set of 56 concatenated markers from 72 bacterial genomes. To produce the gene trees, we aligned the markers using mafft [49] then produced a tree with FastTree2 [50], that was used to orient the Almadén genomes with others in IMG to make a circular phylogeny that represented relationships of α -proteobacteria (Fig. S1). Mercuric acid reductase A (merA) homologs were identified using Orthofinder, and we aligned amino acid sequences of the four homologs merA1 (OG0004493), merA2 (OG0001010) merA2 and dihydrolipoamide 2-oxoglutarate dehydrogenase (D2OD) (OG0000957), and MerA (OG0009124) using Geneious and Clustal. We used Blast to obtain outgroup sequences, many of which were annotated as MerBA in NCBI/GenBank. We used Mr. Bayes v3.2.7 on the CIPRES webserver (<https://www.phylo.org/index.php>) and sampled from 500,000 trees to obtain a consensus tree with posterior probability scores.

Modeling the protein structure of mercury reductase A homologs

Protein structures were modeled through Google ColabFold with MSA mode MMseqs2, pair mode unpaired+paired, and other default settings (<https://colab.research.google.com/github/sokrypton/ColabFold/blob/main/AlphaFold2.ipynb>). ColabFold was developed from AlphaFold2 for accelerated structure prediction of proteins and their complexes, and it matches AlphaFold2 and AlphaFold Multimer in prediction quality of tested datasets [51]. Modelled structures were superposed with COOT [52] for the structural comparison. PyMOL [53] was used to visualize the structures and prepare figures.

Isolation of RNA for rhizobia grown under free-living conditions

We selected four *S. medicae* strains (AMp08, AMp07, SMp01 and VMo01) which show the highest and lowest tolerance to Hg in our collection and had high quality reference genome assemblies for transcriptomics analysis (Table S1). In all experiments in this work, HgCl₂ has been used as the source of Hg. From here on, Hg tolerance, sensitivity, and stress, refer to HgCl₂. *S. medicae* strains AMp08 and SMp01 (Hg-tolerant) which we refer to as “high tolerant” abbreviated as “HT”, AMp07 and VMo01 (Hg-sensitive) which we refer to as “low tolerant” abbreviated as “LT”, were grown in liquid TY medium in the absence or presence of 4 μM HgCl₂ at 28°C in a shaking incubator until they reached an OD₆₀₀ of 0.8. Subsequently, we directly added the Qiagen RNeasy Protect Bacteria Reagent to the bacterial culture in a 1:2 ratio (culture: RNA stabilizer). The mixtures were vortexed for 5 s and incubated for 5 min at room temperature (23°C), then centrifuged at 5,000 × g for 10 min at 20°C. Pellets were frozen in liquid nitrogen and bacterial lysis was performed by resuspension and incubation of the cell pellet in 1 mg/ml lysozyme from chicken egg whites (Sigma-Aldrich) in Tris-EDTA buffer, pH 8.0. Total RNA was extracted using the Qiagen RNeasy Mini Kit using the manufacturer’s conditions specified.

Isolation of RNA from nodules

To evaluate the effects of Hg on *S. medicae* gene expression inside of nodules of *M. truncatula* host-plants, we grew two *M. truncatula* genotypes (HM302, HM304) in sterilized surface as a soil substrate, then inoculated the plants with two *Sinorhizobium medicae* strains (AMp07, AMp08) that have different Hg-tolerance. Plants were well-replicated (average of 20 plants per assay) and completely randomized. Prior to planting, seeds were treated in concentrated sulfuric acid for 3 min and washed 5 times with MiliQ water. Then seeds were surface sterilized using 50% bleach with 0.1% Tween20 for 3 min and washed 8–10 times with sterilized MiliQ and left in sterile water for 2 h at room temperature in dark. The seeds were placed on sterilized filter paper on petri dishes and kept overnight at 4 °C in dark. Later the petri dishes were kept in growth chamber at 21 °C, 40% humidity, on a 16-h-light/8-h-dark photoperiod for a week, then seedlings were transferred to the sterilized Turface: vermiculite (2:1) soil (LESCO-turface all sport pro soil) in plant growth chamber. Plants were cultivated in a walk-in growth chamber in 16-h-light/8-h-dark photoperiod and 22 °C conditions and watered every 5–7 days with sterilized ½ strength of B&D medium with low concentration of KNO₃ (0.5mM) (Supplementary Data 1, Table S7). Prior to inoculation, KNO₃ was not added while watering the seedlings.

To prepare the inoculum, rhizobia strains were cultured at 28 °C in sterilized liquid TY medium until an OD₆₀₀ of 0.8 was reached. Subsequently, the culture was centrifuged at 4000 rpm for 5 min and the supernatant removed and washed 2 times with autoclaved 0.9% NaCl. The pellet was resuspended in 0.9% NaCl for inoculation. Seedlings were inoculated with *S. medicae* AMp07 and AMp08 rhizobia strains 5–7 days after being transplanted to the sterilized surface soil. 21 days post inoculation (21dpi), the control plants were watered with ½ strength of B&D medium supplemented with 0.5 mM concentration of KNO₃, while the Hg-treated plants received the same media with an addition of 100μM HgCl₂. After 7 days of treatment, the nodules were harvested and quickly frozen in liquid nitrogen and stored at –80 °C for RNA-seq. For RNA extraction, the nodules were first crushed in Qiagen TissueLyser II for 2 min in 30 s intervals. Subsequently, RNA was extracted using Spectrum Plant Total RNA kit (Sigma) by following the manufacturer’s manual. The concentration and quality of RNA was quantified by using nanodrop (Thermo Scientific) and Bioanalyzer (Agilent) respectively.

Library preparation and RNAseq of rhizobia grown under free-living conditions and in nodules

Library preparation and transcriptome sequencing of the free-living rhizobia samples was done at the DOE Joint Genome Institute under Community Science Project (Functional Genomics) 506402. RNAseq data for each library was generated using the Illumina NovaSeq S4 platform which generated 2×151 bp reads. BBDuk (version 38.90) was used to remove contaminants (BBDuk, BBDuk and BBMerge commands used for filtering are placed in file: 52510.2.364064.AAGAAGGC-AAGAAG GC.filter_cmd-MICROTRANS.sh), trim reads that contained adapter sequence and homopolymers of G’s of size 5 or more at the ends of the reads, right quality trim reads where quality drops below 6, remove reads containing 1 or more ‘N’ bases, remove reads with average quality score across the read less than 10, having minimum length ≤ 49 bp or 33% of the full read length.

For nodule samples, construction of the RNAseq libraries and sequencing on the Illumina NovaSeq 6000 were performed at the Roy J. Carver Biotechnology Center at the University of Illinois at Urbana-Champaign. Purified DNase treated total RNAs were run on a Fragment Analyzer (Agilent) to evaluate RNA integrity. The total RNAs were converted into individually bar-coded RNAseq libraries with the Universal Plus mRNA-Seq Library Preparation kit from Tecan, using custom probes against *Medicago* and *S. melliloti* rRNA designed by Tecan. Libraries were barcoded with Unique Dual Indexes (UDI’s) which have been developed to prevent index switching. The adaptor-ligated double-stranded

cDNAs were amplified by PCR for 10 cycles. The final libraries were quantitated with Qubit (ThermoFisher) and the average cDNA fragment sizes were determined on a Fragment Analyzer. The libraries were diluted to 10nM and further quantitated by qPCR on a CFX Connect Real-Time qPCR system (Biorad) for accurate pooling of barcoded libraries and maximization of number of clusters in the flowcell. The barcoded RNAseq libraries were loaded on one SP lane on an Illumina NovaSeq 6000 for cluster formation and sequencing. The libraries were sequenced from both ends of the fragments for 150 bp from each end. The fastq read files were generated and demultiplexed with the bcl2fastq v2.20 Conversion Software (Illumina).

Mapping of RNAseq reads and differential expression analysis

Free-living rhizobia RNAseq reads from each *S. medicae* strain were mapped to the newly assembled genomes specific to each strain. For nodule samples, concatenated reference genomes of *Medicago truncatula* v5.0 and either *S. medicae* AMp07 or AMp08 were used for mapping the dual-transcriptomics samples. Indexing of each genome was done using STAR v2.7.10a [54]. FastQC v0.11.9 [55] and Trimmomatic v0.39 [56] were used for quality control and adapter removal with default parameters. The resulting high-quality reads were mapped to the appropriate reference genome assembly using STAR v2.7.10a and raw counts were generated using the GeneCounts feature. Differentially expressed genes were identified using DESeq2 v1.34.0 and p-values were adjusted using Benjamini and Hochberg's correction [57]. PhyloFlash v3.4 [58] was used to check for contamination of samples using bacteria databases to align raw data. To aggregate read counts for each gene, we used a custom python script `aggregate_counts.py`. (https://github.com/mclear73/RNAseq_Miniworkshop/blob/main/Processing%20Raw%20RNAseq%20Data%20into%20Counts.md) which generates a count table and metadata file to be used for input for DESeq2, which we used to calculate differential expression in Hg-treated samples compared with control samples. Genes with an adjusted p-value < 0.05 and \log_2 fold-change ≥ 1 were considered differentially expressed. For direct comparison of gene counts between samples, raw gene counts were normalized using the trimmed mean of m-value (TMM) method [59] from edgeR v3.36.0 [60]. Transcripts per kilobase million (TPM) were calculated using a custom python function.

Ortholog comparison and GO-enrichment analysis

Orthogroups were identified for 20 related bacterial strains including AMp08 annotated with the Prokaryotic Genome Annotation Pipeline (PGAP) [46] with

OrthoFinder v2.5.4 [47] with default parameters. The complete genome annotation for AMp08 was generated using NCBI PGAP (v2022-02-10.build5872) because the RAST annotations cannot be directly linked to GO-terms. To allow for a standardized comparison of enriched GO-terms between strains, all GO enrichments were performed with the PGAP annotation of AMp08 via orthologs determined by OrthoFinder. Using custom python scripts, DEGs were mapped to the Orthogroup provided by Orthofinder and all AMp08 genes within the Orthogroup were used as input for GO Enrichment. Enriched GO-terms for DEGs were identified using a custom python script for g:Profiler2 using the built in "g_SCS" function that corrects for multiple tests [61]. Tests for overrepresentation were performed for the biological processes (BP), Cellular components (CC), and molecular function (MF) databases based on the PGAP annotation for AMp08.

Transfer of plasmid containing mer operon into low-tolerant strains

The *Sinorhizobium medicae* strain AMp08 from the Almadén mine is highly tolerant to Hg and harbors a Mer operon on plasmid. To isolate the plasmid, the AMp08 strain was grown in liquid TY medium (28 °C, 220 rpm) until $OD_{600}=0.8$ was reached. Afterward the plasmid was isolated using GenElute plasmid miniprep kit (Sigma Aldrich). The recipient strains (AMp07 and WSM419) were made electrocompetent using the protocol described in [62]. Specifically, the strains were grown in liquid TY medium at 28 °C for 2 days at 220 rpm until they reached mid-logarithmic stage ($OD_{600}=0.8-1.0$). Subsequently the cells were chilled on ice for 15 to 30 min, and then collected from the pellet after centrifugation at 3700 g for 10 min at 4 °C. The cell pellet was washed four times with chilled deionized water, then with 10% glycerol. The cells were resuspended in 10% glycerol and stored at -80 °C in 100 μ l aliquots. For electroporation, 4 μ l (~250–300 ng) of AMp08 plasmid DNA was mixed with the recipient strain by gentle vortexing and kept on ice for 30 min. The cell-DNA mixture was then loaded into a pre-chilled electroporation cuvette with a 0.1-cm gap and was exposed to a single pulse of high voltage of field strength 14 kV/cm for 5.8 ms. After the electroporation, the cuvettes were kept on ice for an additional 10 min. Once the cells were allowed to grow in liquid TY media for an additional 1 h, the cell culture was plated on TY plates containing 200 and 250 μ M HgCl₂ to select for transformed strains. Finally, to confirm the positive transformation, genomic DNA was isolated from the positive colonies and the presence of AMp08 MerA gene was confirmed using PCR. To confirm if the Mer operon was functional in the transformed strains, qPCR was conducted to check the expression levels of key Mer

operon genes (MerA and MerT) as well as merA1, which is expressed in all the *S. medicae* strains. Briefly, RNA was isolated using the methodology described above. 1 µg of total RNA was used to synthesize cDNA using SuperScript III cDNA synthesis kit (Invitrogen, U.S.A) with random hexamer primers, and the cDNA samples were diluted 1:10 in dH₂O. qRT-PCR assays were conducted using gene specific primers (Supplementary Data 1, Table S6) using SYBR green master mix (Bio-Rad). The reaction was run for a total of 40 cycles, and the relative expression was calculated using the comparative C_t method (2^{-DDCt}) method. The threshold cycle (C_t) value of the reference gene, namely 16 S rRNA was used to normalize the transcript levels of all genes. All the results shown here include data averaged from three biological replicates, which in turn included three technical replicates each.

Quantifying minimum inhibitory concentration (MIC)

The MIC represents the lowest concentration of Hg at which the growth of a strain is inhibited. To quantify MIC, a stock solution of TY medium and 1% agar was prepared by adding 100 mM of filter sterilized HgCl₂ stock solution to autoclaved TY to get the required concentrations containing 0 µM, 25 µM, 50 µM, 75 µM, 100 µM, 150 µM, 200 µM, 250 µM HgCl₂. All the rhizobia strains used in this study (AMp07, AMp08, WSM419 and the engineered AMp07/pAMp08 and WSM419/pAMp08) were grown in autoclaved liquid TY medium until an OD₆₀₀ of 0.8 was reached. Subsequently, the cultures were serially diluted 1:10, 1:100, 1:1000 using TY solution and 10 µl each of the non-diluted and diluted cultures were spotted to all the Hg-containing TY plates. The plates were incubated at 28 °C for 2 days and images were taken. MIC was confirmed for all the original strains (AMp07, AMp08, WSM419) and determined for the engineered AMp07/pAMp08 and WSM419/pAMp08 strains.

Results

Complete chromosomes and plasmids of high mercury tolerant and low-tolerant strains

The genomes of the *S. medicae* strains were each assembled into 3–6 contigs with one exception (SupplementaryData1, Table S1), and the *R. leguminosarum* genomes into 6–11 contigs. In most cases these contigs constituted complete circularized chromosomes and plasmids. Three of the most Hg-tolerant Almadén strains (*S. medicae* strains AMp08, SMp01 and *R. leguminosarum* strain STf07) each had a mercury reductase (Mer) operon which was previously unknown in this collection of rhizobia. Furthermore, Mer operons have not been reported in *S. meliloti*, *S. medicae* or *R. leguminosarum* despite two decades of genomic studies in these species. The

genome sizes and contig assemblies are consistent with previously published reference genomes of both species, which show *S. medicae* possessing one main chromosome and two symbiotic plasmids (pSymA and pSymB) which contain the symbiotic signaling genes [63] which is true of the genomes of both HT and LT strains, and often an additional accessory plasmid that is unique to individual strains [19, 64]. *R. leguminosarum* tends to have a single large chromosome like *S. medicae*, but typically more plasmids [19, 65]. We performed a phylogenomic analysis using 7 of the Almadén genomes with 65 additional genomes from the Integrated Microbial Genomes and Microbiomes (IMG) database (<https://img.jgi.doe.gov>) and found that four of the Almadén genomes were located in clades containing *S. medicae* and *S. meliloti*, and the other three strains in an adjacent clade containing *R. leguminosarum* (Fig. S1 - S9). The closest related strain to the Almadén strains is *Mesorhizobium* sp. LSJC277A00 which contains a Mer operon, but data showing its high Hg-tolerance could not be found. Therefore, the genome assemblies would allow us to make valuable insights into the evolution of local adaptation to Hg tolerance in the Almadén rhizobia strains due to presence-absence variation of the Mer operon.

The genomes of the two most Hg-tolerant *S. medicae* strains, AMp08 and SMp01 were assembled into 4 and 6 contigs respectively. We aligned these two focal genomes to the reference *S. medicae* strain WSM419 [19] using Sibelia to examine synteny and to assign each contig to chromosomes and symbiotic plasmids. In the case of the Hg-tolerant strain AMp08, the main chromosome and two symbiotic plasmids (pSym) are syntenic with the main chromosome and symbiotic plasmids of *S. medicae* WSM419: one chromosome (~3.8 Mb), pSymA (~1.2 Mb), pSymB (~1.6 Mb). It also included an accessory plasmid (~71 kb) that was not syntenic or shared any homology with the accessory plasmid of WSM419 (Fig. 1a). The accessory plasmid in AMp08 which contained a complete mercury reductase (Mer) operon, (Fig. S2) is likely derived from a HGT from another rhizobia strain or bacteria species and contains many genes of unknown function (hypothetical proteins).

We also checked the homology and synteny of the main chromosome and two pSym plasmids of Hg-tolerant *S. medicae* strain SMp01 against WSM419, which showed that 4 of the 6 contigs could be merged (contig_0 with contig_2 and contig_3 with contig_4) resulting in 4 contigs total (Fig. 1b). Similar to AMp08, the SMp01 strain also has a Mer operon, but at a different genomic location. In SMp01, the Mer operon is located on the main chromosome and not on the pSym or accessory plasmids, and this region is not syntenic with WSM419 as it lacks the Mer operon. Finally, we aligned the Hg-tolerant *R. leguminosarum* strain STf07 against the *R.*

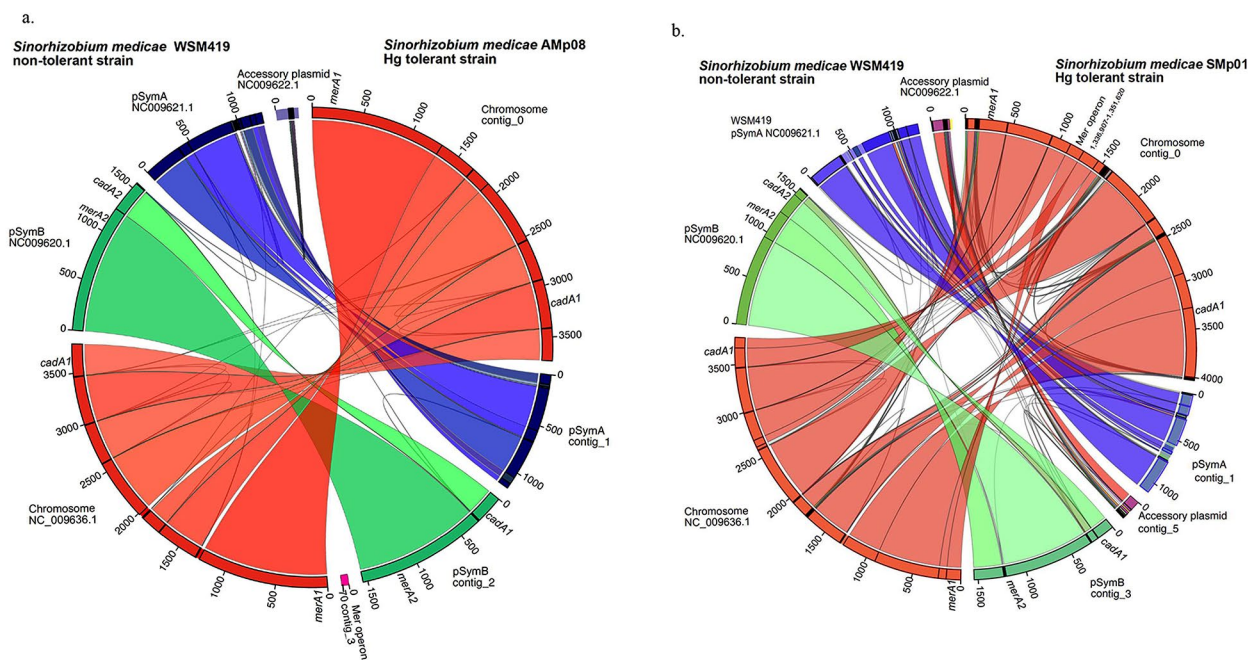


Fig. 1 (a) Comparison of the Almadén mine *Sinorhizobium medicae* strain AMP08 (the most Hg tolerant strain which contains a mercury reductase (Mer) operon on an accessory plasmid) with the model *S. medicae* strain WSM419. (b) Comparison of the Almadén mine *S. medicae* strain SMP01 (the most second most Hg-tolerant strain which contains a Mer operon on the main chromosome, at position 1,336,907- 1,351,620) with the model *S. medicae* strain WSM419. On the outer ring, common colors were used for homologous chromosomes between the two strains (e.g., red = chromosome, blue = pSymA, green = pSymB), and syntenic regions within chromosomes or plasmids spanning the center of the Fig. share common colors between the two strains. Numbers on outer ring are $\times 1000 = \text{Megabase (Mb)}$

leguminosarum reference strain WSM1325 [20] and found that this genome was also assembled into complete chromosomes and plasmids that aligned with the reference genome, where both strains possessed one large chromosome and five plasmids. Strain STf07 also possesses a Mer operon, which is located on a smaller plasmid and not on the main chromosome (Fig. S3). We therefore have high quality reference genomes for the three most Hg-tolerant rhizobia strains in our collection, each with Mer operons.

Mercury reductase A phylogeny and synteny of the mer operon in the Almadén strains and across α -proteobacteria

The three most Hg-tolerant strains among our assembled genomes from Almadén (*S. medicae* AMP08, SMP01 and *R. leguminosarum* STf07) had minimum inhibitory concentrations (MIC) ranging from 200–250 μM , while low-tolerant (LT) strains had MIC of 25 μM [28], which is a 10-fold difference in Hg-tolerance. Those with $\text{MIC} \geq 200 \mu\text{M}$ were defined as high-tolerant (HT) based on the range of variation shown by Nonnoi et al., (2012) where most strains had $\text{MIC} = 25 \mu\text{M}$ (LT) and only a few between 50 and 150 μM . Because mercury reductase A genes (*merA1* and *merA2*) are pervasive in *Sinorhizobium* and *Rhizobium*, we conducted a phylogenetic analysis to identify whether the *merA* homologs

from the highly tolerant strains formed separate clades from the LT strains, including non-Almadén strains *S. medicae* (WSM419), *S. meliloti* (1021) and *R. leguminosarum* (WSM1325), and other outgroup bacteria species. The phylogenetic analysis of *merA* homologs in the *S. medicae* and *R. leguminosarum* Almadén strains showed four clades (Fig. 2). One clade (green clade) was comprised of mercury reductase A genes that had the greatest sequence diversity, and only included our three HT strains (*S. medicae* AMP08, SMP01 and *R. leguminosarum* STf07) along with other outgroup *MerA* sequences from non-Almadén rhizobia. This clade was comprised of *MerA* genes that were part of a Mer operon, while the *merA* genes from less tolerant rhizobia strains that lack a Mer operon were not present in this clade and formed separate clades (red and blue).

Fusions of the *MerA* and *MerB* proteins have been found in bacteria where the N-terminal region of *MerA* was replaced by the alkylmercury lyase domain of *MerB* [7, 16]. Blast searches of the *MerA* gene from the Mer operon in the Hg-tolerant *S. medicae* strains against α -proteobacteria and outgroup bacteria including *Pseudomonas*, *Xanthobacter*, and *E. coli* resulted in hits that were often annotated as “MerBA”, despite many lacking an actual fusion of *MerA* and *MerB* genes. In many cases, this appears to be an annotation artifact where

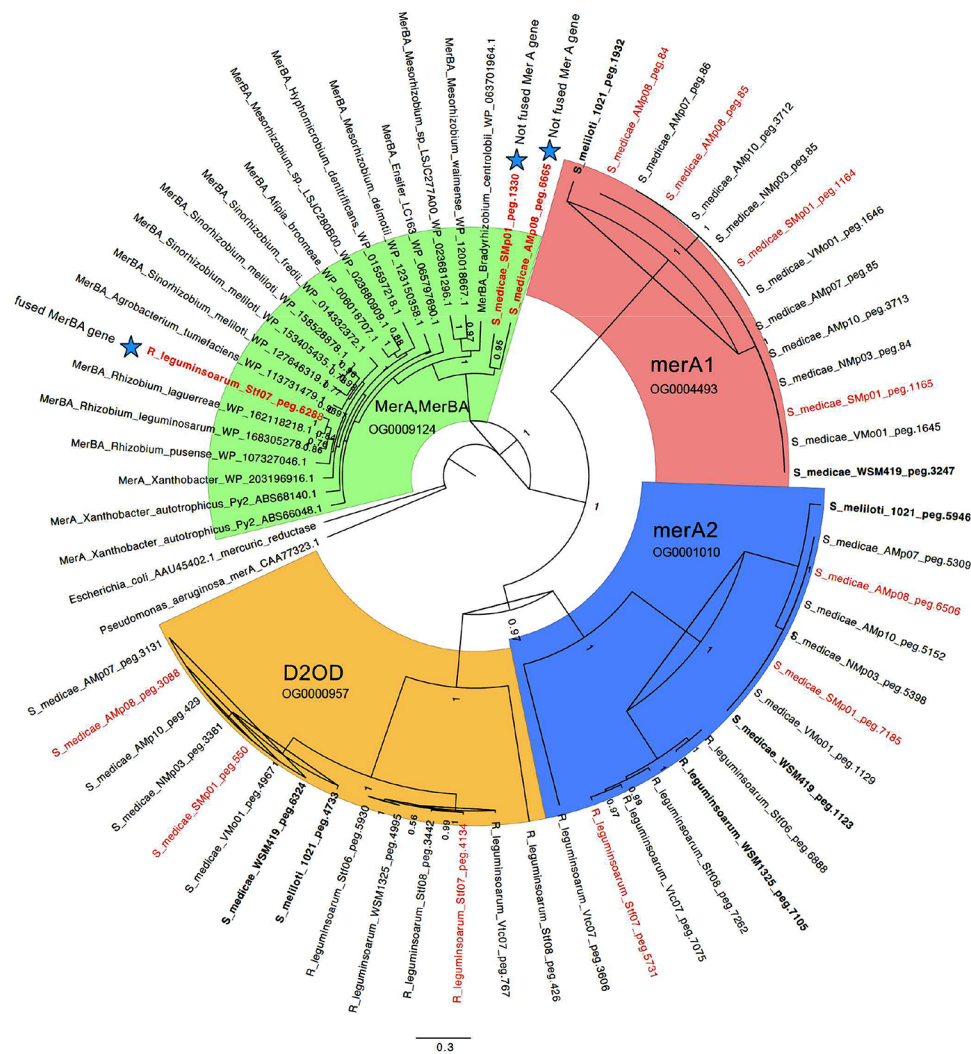


Fig. 2 Phylogeny of mercury reductase A (merA) homologs with a focus on *Sinorhizobium medicae* and *Rhizobium leguminosarum*. The merA homologs from the Almadén mine strains that have high tolerance to Hg are indicated by red text. The phylogeny is colored according to four main clades which includes their homolog name and Orthofinder ID. The green clade includes MerA and fused MerBA genes from multiple species/strains of *Mesorhizobium*, *Sinorhizobium* and *Rhizobium* and various bacteria including *Pseudomonas aeruginosa* and *Escherichia coli* (used to root the phylogeny), *Xanthobacter autotrophicus* which has well studied MerA and MerB genes [16], *Agrobacterium tumefaciens*, *Hyphomicrobion dentrificans*, *Alfipia broomeae*, and multiple species/strains of *Mesorhizobium*, *Sinorhizobium* and *Rhizobium*. In the green clade are the MerA genes from the Almadén *S. medicae* strains AMp08, SMp01 and *R. leguminosarum* strain STf07 are marked with a blue star. The strain STf07 has a fused MerA and MerB (“MerBA”) (OG0009124) while the strains AMp08 and SMp01 do not have fused MerA and MerB. The merA1 (OG0004493), merA2 (OG0001010), and dihydroliipoamide 2-oxoglutarate dehydrogenase (D2OD) (OG0000957) are colored red, blue and orange, respectively. The numbers on the nodes are posterior probability scores from 500,000 sampled trees using Mr. Bayes v3.2.7 on the CIPRES webserver (<https://www.phylo.org/index.php>)

true fusions of MerA and MerB were classified as MerBA during annotation of deposited genomes due to their close homology to the MerA region of the fused MerBA sequences in GenBank. Among the three most Hg-tolerant strains from Almadén that possessed a Mer operon, one had a MerA gene but no MerB gene (AMp08), one had a MerA and three MerB genes (SMp01), and one had a fused MerBA gene (STf07). Therefore, we conducted a synteny analysis of the Mer operon from the three most Hg-tolerant Almadén strains along with other

a-proteobacteria strains available in IMG. The analysis revealed three independent origins of a Mer operon, with each of the Hg-tolerant strains having a Mer operon being more closely related to another species than to each other (Fig. S4), despite all three strains being collected from Almadén. These results indicate that the Mer operons were derived through different HGT sources. Several a-proteobacteria genomes such as *Mesorhizobium* sp. LSJC277A00 had a complete Mer operon, as well as a fused MerBA, but no publications or data reported

a Mer operon in the genomes deposited in GenBank or IMG. However, the non-Almadén strains with a complete operon presumably have elevated tolerance to Hg, but tolerance to Hg was not tested.

In addition to the clade containing MerA homologs that were in Mer operons, our phylogenetic analysis revealed three additional clades (red, blue, and orange clades in Fig. 2). A second clade contained merA1 from all *S. medicae* strains from Almadén, including a short tandem duplication only found in the Almadén strains (merA1'), along with merA1 from *S. medicae* WSM419 and *S. meliloti* 1021 (Fig. 2). The merA1 copies from the strains from Almadén and the non-Almadén reference strains showed almost no amino acid sequence diversity among all strains. This suggests merA1 is under strong selective constraint in *Sinorhizobium*, presumably because it provides basic tolerance to Hg²⁺ ions in the environment, regardless of environmental contamination level. A third clade, which we designated as merA2 (annotated as FAD-dependent NAD(P)-disulphide oxidoreductase) based on [32], contains all *S. medicae* and *R. leguminosarum* strains derived from Almadén, as well as the non-Almadén strains including *S. medicae* WSM419, *S. meliloti* 1021 and *R. leguminosarum* WSM1325 (Fig. 2). *R. leguminosarum* does not possess any homologs in the merA1 clade but has homologs in the merA2 clade. The merA2 gene in this species likely provides the same basic tolerance to Hg²⁺ as does merA1 and merA2 in both *S. medicae* and *S. meliloti*.

A fourth clade which contains dihydroliipoamide 2-oxoglutarate dehydrogenase. (abbreviated as D2OD), is a close homolog of merA1 and merA2. An examination

of the expression of D2OD in response to Hg-stress may help to determine whether this homologous gene plays a role in Hg detoxification in rhizobia. Finally, because of the independent assemblies and annotations of each of the genomes studied here, we used OrthoFinder to identify homologous genes between each of the genomes. Each of the four genes (MerA/MerBA, merA1, merA2, D2OD) were assigned their own orthogroup which was consistent with their separate phylogenetic clades using Mr. Bayes (Fig. 2). We will examine the expression of genes in *S. medicae* represented in the four clades in strains with different tolerance to Hg in later sections of this paper.

Models of protein structure of mercury reductase A homologs in *S. medicae* show high conservation and the presence of a MerBA fusion protein in *R. Leguminosarum*

To estimate whether the tertiary structure showed differences were obvious among the homologous merA proteins (merA1, merA2, D2OD and MerA), we used AlphaFold to create models of each of the homologs and then compared them by superimposing all four structures on each other (Fig. 3). To this end, we used amino acid sequences of the four genes from *S. medicae* strain SMp01, which contains a Mer operon. The average pLDDT (ranging from 93 to 96) and PTM score (ranging from 0.92 to 0.94) in AlphaFold [51] indicated the structural models are highly reliable. The four modelled structures were highly homologous (RMSDs range between 1.0 and 1.4 Å) and each one formed a homodimer (Fig. 3). The structures revealed the N-terminal FAD-binding domain active site CXXXXC motif, the

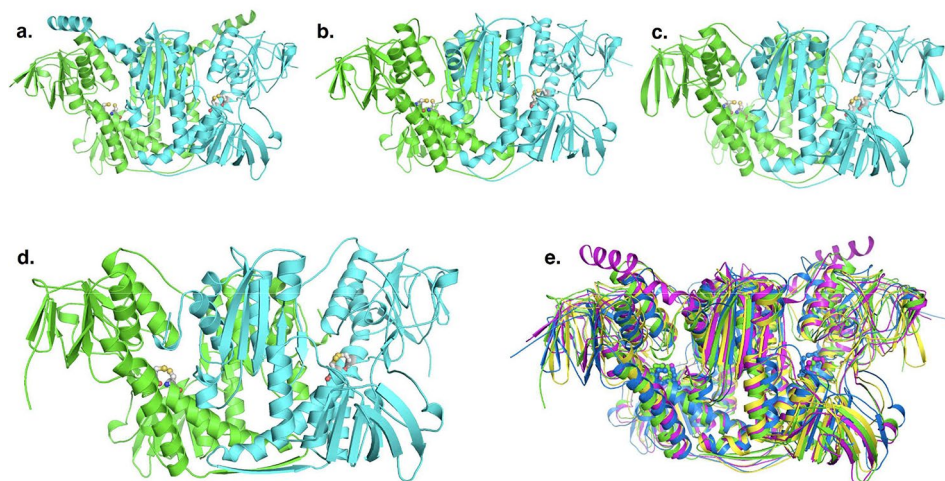


Fig. 3 Ribbon representations of the dimeric protein structures of homologous mercuric reductases from *Sinorhizobium medicae* strain SMp01 modeled using AlphaFold. The generic mercury reductase proteins merA1 (a), merA2 (b) and D2OD (c) are found in all *S. medicae* and *S. meliloti* genomes regardless of their tolerance (Fig. 2). The MerA protein (d) from the Mer operon is only found in the Hg-tolerant strains from Almadén (and other publicly available genomes). The two monomers are shown in green and cyan colors. Superposition of structures of merA1 (magenta), merA2 (blue), D2OD (yellow) and MerA (green) (e) show high levels of conservation and homology. Conserved active site cysteine residues are shown as stick models

NADPH-binding domain, the central domain, and the C-terminal dimer interface domain, were all conserved between the four homologs. Previous studies have demonstrated that the N-terminal domain (NmerA) of MerA is responsible for recruiting Hg^{2+} and transferring it to the cysteine pair located at the flexible C-terminal segments [66, 67]. The C-terminal cysteine pair undergoes a conformational change to deliver Hg^{2+} to the active site of the opposing monomer. However, in the *S. medicae* MerA homologs, while the N-terminal NmerA domain is absent, the overall structural fold and expected reductase activity appear similar to the catalytic core of the *Bacillus* sp. strain RC607 [68], *Pseudomonas aeruginosa* [66], *Lysinibacillus sphaericus* strain G1 [69], and FDR family proteins [67, 70, 71].

To uncover the domain architecture of MerBA fusion proteins in *R. leguminosarum* strain STf07, we modelled the structure of MerBA in this strain and compared it with *Mesorhizobium* sp. LSJC277A00 using AlphaFold (Fig. S6). The structures of MerBA from both *R. leguminosarum* strain STf07 and *Mesorhizobium* sp. LSJC277A00 were homodimeric and the MerA component was highly homologous to the non-fused MerA proteins from *S. medicae*. In both protein structures, the MerB domain is tethered to MerA by a flexible linker and minimally interacts with N-terminal FAD-binding domain of MerA (Fig. S6). The structure of the MerB domain was found to be similar to *E. coli* [72] and the

conserved active site residues Cys 147, Cys 212 and Asp 150 were located at the central core of the domain. To our knowledge, this is the first reported protein prediction for a fused MerBA protein, at least in nitrogen-fixing α -proteobacteria.

Differentially expressed genes (DEGs) in response to hg stress in free-living conditions are mostly mer operon genes in highly tolerant rhizobia

We selected four *S. medicae* strains with variable tolerance to Hg (AMp08, SMp01, AMp07 and VMo01) and treated them with 4 μM $HgCl_2$ in free-living conditions for transcriptomic analysis (Supplementary Data 2). We found that the HT strains had dramatically fewer differentially expressed genes (DEGs) when compared to LT strains (Fig. 4a). The two HT strains AMp08 and SMp01, had only 17 (0.2% of the total genome) and 32 (0.5% of the total genome) DEGs respectively, which were mostly upregulated (60–90%). The LT strains AMp07 and VMo01 had 540 (7.9% of the total genome) and 448 DEGs (10.3% of the total genome) respectively. This amounts to roughly 25 times more DEGs in the LT strains than in HT strains, and there were almost equal proportions of upregulated and downregulated genes (40–55%) in the LT strains. These strain-specific patterns can be easily visualized by comparing the volcano plots (Fig. 4b–d, OrthoFinder IDs were used to label genes due to the independent assembly of each reference genome),

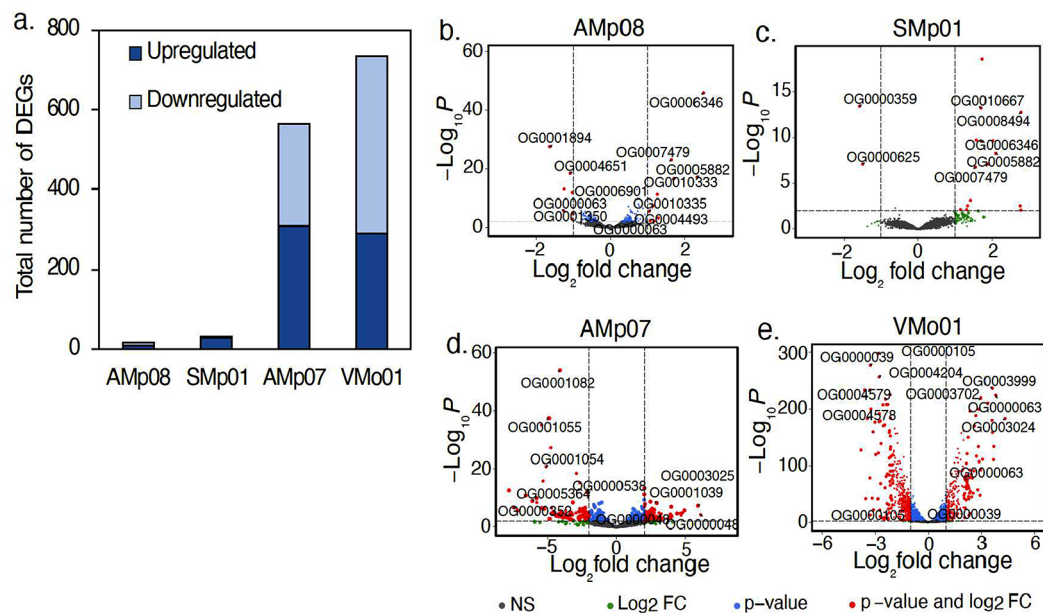


Fig. 4 Differentially expressed genes (DEGs) in free-living conditions following Hg stress (a). Free-living rhizobia were grown in the absence (0 μM , Control) or presence of 4 μM $HgCl_2$ (Hg-treated). DEGs were identified based on the criteria $|\log_2 FC| \geq 1$, adjusted $p < 0.05$, and those that met these criteria are shown in red in the volcano plots for AMp08 (b), SMp01 (c), AMp07 (d) and VMo01 (e). The strains AMp08 (b) and SMp01 (c) each have a complete mercury reductase operon (see Fig. 2). Each point represents a gene, and the top significantly upregulated and downregulated DEGs were labeled using their respective Orthogroup IDs

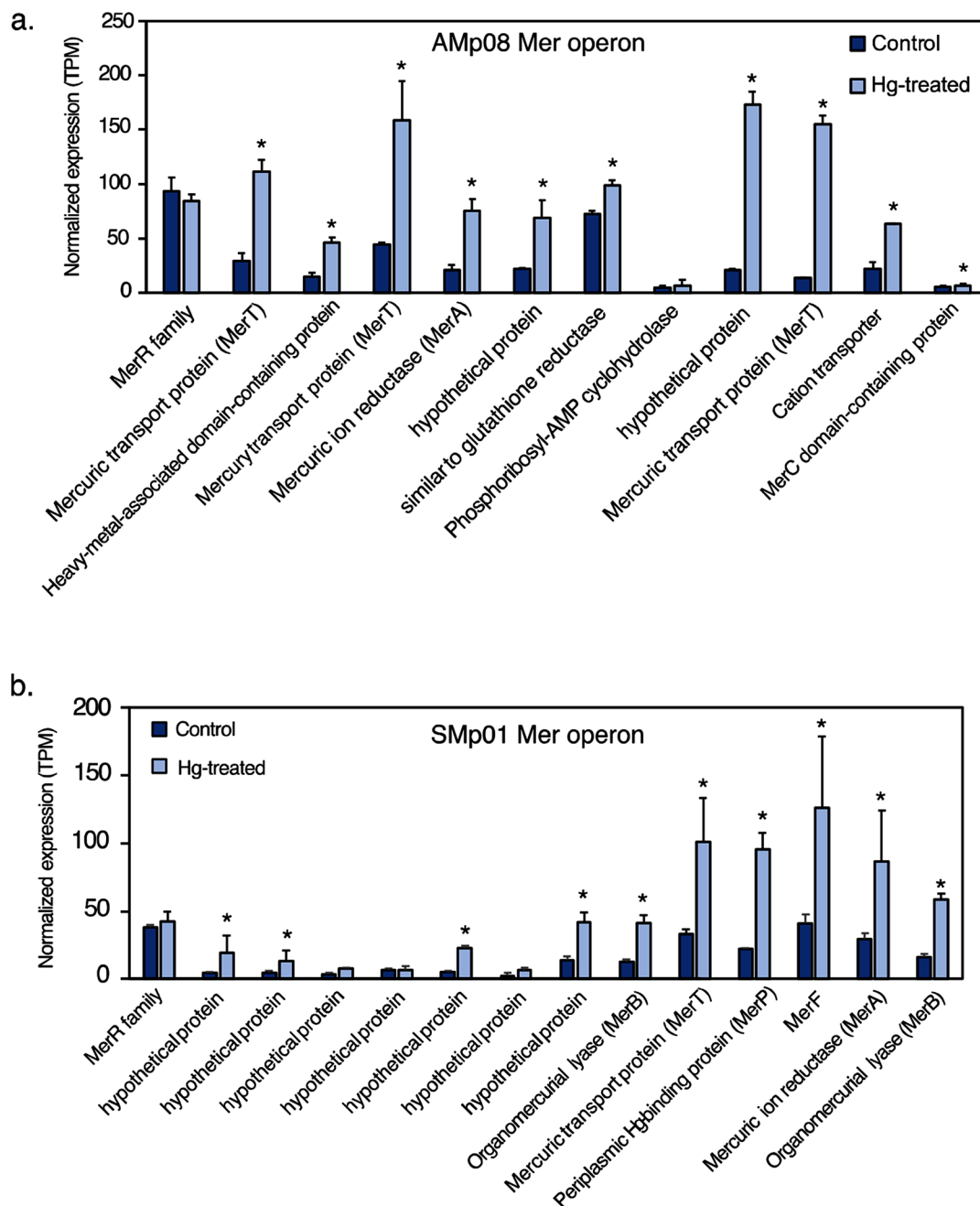


Fig. 5 Expression of Mer operon genes in AMp08 (a) and SMp01 (b) in control conditions compared with Hg treated. The strain AMp08 does not possess a MerB gene. In SMp01 MerB is present in two copies. Free-living rhizobia were grown without Hg (Control) or presence of 4 μM HgCl_2 (Hg-treated). Three biological replicate cell cultures were grown for control and Hg-treated. Normalized expression levels are reported as the mean TPM of three biological replicates

which also show key genes involved in Hg detoxification such as MerA (OG0007479), MerT (OG0006346), MerB (OG0010667), and MerP (OG0005882) among the other most significant DEGs (Supplementary Data 2). The very low number of DEGs, and mostly upregulated genes in HT strains, suggests quick triggering of the Mer operon gene expression and highly efficient removal of Hg from

the cells. On the other hand, the higher number of DEGs in LT strains, and more equal ratios of up to down-regulated genes, suggests greater cellular and genomic impact of Hg stress on the susceptible strains.

For the LT strains AMp07 and VMo01 which lack a Mer operon, several genes from the cytochrome O-ubiquinol oxidase gene family (OG0001054, OG0001055)

and nitric oxide dioxygenase (OG0001082) were significantly downregulated, and genes from ABC-transporter families such as ABC-transporter substrate-binding protein (OG0003025), ferric-ion iron-binding ABC-transporter protein (OG0003702), and beta-galactoside, ABC-transporter substrate-binding protein (OG0003024, OG0003025) were significantly upregulated (Fig. 4d and e; Table S2 in Supplementary Data 1). These genes and gene families are known to be involved in mediating oxidative stress responses and ABC-transport mechanisms of cellular compartmentalization of toxic ions [73], which can be expected to be more responsive in the absence of an effective mercury detoxification mechanism such as a Mer operon.

Significant up-regulation of mercury reductase operon (mer) genes in Hg-tolerant strains but not in the generic merA genes

As reported above, many of the significant DEGs that responded to Hg stress belonged to the Mer operon. While the generic mercury reductase genes merA1 and merA2 are present in all the *S. medicae* strains as stand-alone genes, only the Hg-tolerant strains possess the Mer operon (*S. medicae* strains AMp08 and SMp01) which includes a MerA gene surrounded by other genes that transport and detoxify Hg in the operon locus. In the two HT strains, AMp08 and SMp01, a majority of the DEGs, including MerA, showed significant up-regulation in response to Hg. In the Mer operon, we also observed significant upregulation for MerT (mercuric transport protein), MerB (organomercurial lyase, all three copies in Smp01), MerC (cellular membrane transport), MerP (periplasmic mercury transport), MerF and other non-characterized hypothetical proteins (Fig. S5 a, b). It is noteworthy that upregulation of MerB occurred despite no methyl-mercury (CH_3Hg^+) present in our culture media, indicating that ionic Hg is enough to trigger upregulation of all genes in the Mer operon. Interestingly, we did not find any significant expression change of the transcriptional regulator MerR in the Mer operon in either strain, suggesting constitutive expression of MerR.

The common mercury reductase gene merA1 (OG0004493), which transforms Hg^{2+} into the volatile form Hg^0 , is present in all the strains analyzed in the current study regardless of Hg-tolerance. We observed upregulation of merA1 in response to Hg stress in the Almadén strains, however the change was not significant in neither tolerant nor susceptible strains (Fig. 6a). Similarly, another mercury reductase annotated as FAD-dependent NAD(P)-disulphide oxidoreductase, referred to as merA2 (OG0001010) [32], showed non-significant changes in expression (Fig. 6b). D2OD (OG0000957), which is homologous to the merA genes (Fig. 2), showed a significant upregulation in response to Hg treatment

in the LT strains only (Fig. 6c). It is unclear whether this gene contributes to some level of Hg tolerance but an extensive search of merA homologs by Boyd and Barkay (2012) across a broad range of bacteria also found dihydroliipoamide dehydrogenases to be homologous to merA, but experimental evidence of their role in Hg tolerance or detoxification is lacking.

Comparison of DEGs in response to hg stress in nodules shows the effect of symbiosis and genotype x genotype interactions

To test the effect of Hg stress on rhizobia in their symbiotic state inside of host-plant nodules, we performed an experiment using two *M. truncatula* host-plant genotypes (high Hg-accumulating genotype HM304 and a low Hg-accumulating genotype HM302) inoculated with either a HT (AMp08) or LT (AMp07) *S. medicae* strain. Following 21 days post-inoculation, the plants were treated with 100 μM HgCl_2 . The purpose was to provide a biologically relevant rhizobia response as the nodule is a drastically different environment than free-living conditions. The RNA isolated from the nodule is a mixture of both the host plant and rhizobia, and we here focused on the rhizobia responses in the current study, and the plant responses will be reported separately [74].

We found a stark difference between the number of DEGs in the nodules formed by the tolerant strain AMp08 in the nodules of both host-plant genotypes when compared to the LT strain, AMp07. In the nodules, AMp08 showed 186 genes and 330 genes in the host-plants HM304 and HM302 respectively (Fig. 7a). By comparison, AMp07 had 873 DEGs in HM304 while those in HM302 had 639 DEGs, which was a 1.8 and 4.5-fold increase in the number of DEGs in AMp07 compared with AMp08 in the two host-plants (Fig. 7a), respectively. These results show a similar trend with what we observed under free-living conditions where the HT strain also showed lower numbers of DEGs than the LT strain. There was not a major difference between the percent DEGs in free-living conditions versus nodules for the LT strain as in AMp07, 7% of genes were differentially expressed under free-living vs. 9–12% in nodules. However, for the HT strain (AMp08), we observed over 10 times increase in percent DEGs in nodules compared with free-living conditions, as only 0.25% of genes were differentially expressed (mostly Mer operon genes) under free-living conditions, compared to 3–5% DEGs in nodules, which included other genes in addition to the Mer operon. The higher number of rhizobia genes responding in nodules may be due to a stronger stress treatment in the experiment conducted *in planta*, or the presence of other environmental factors inside the nodules, including signaling between host-plants and rhizobia.

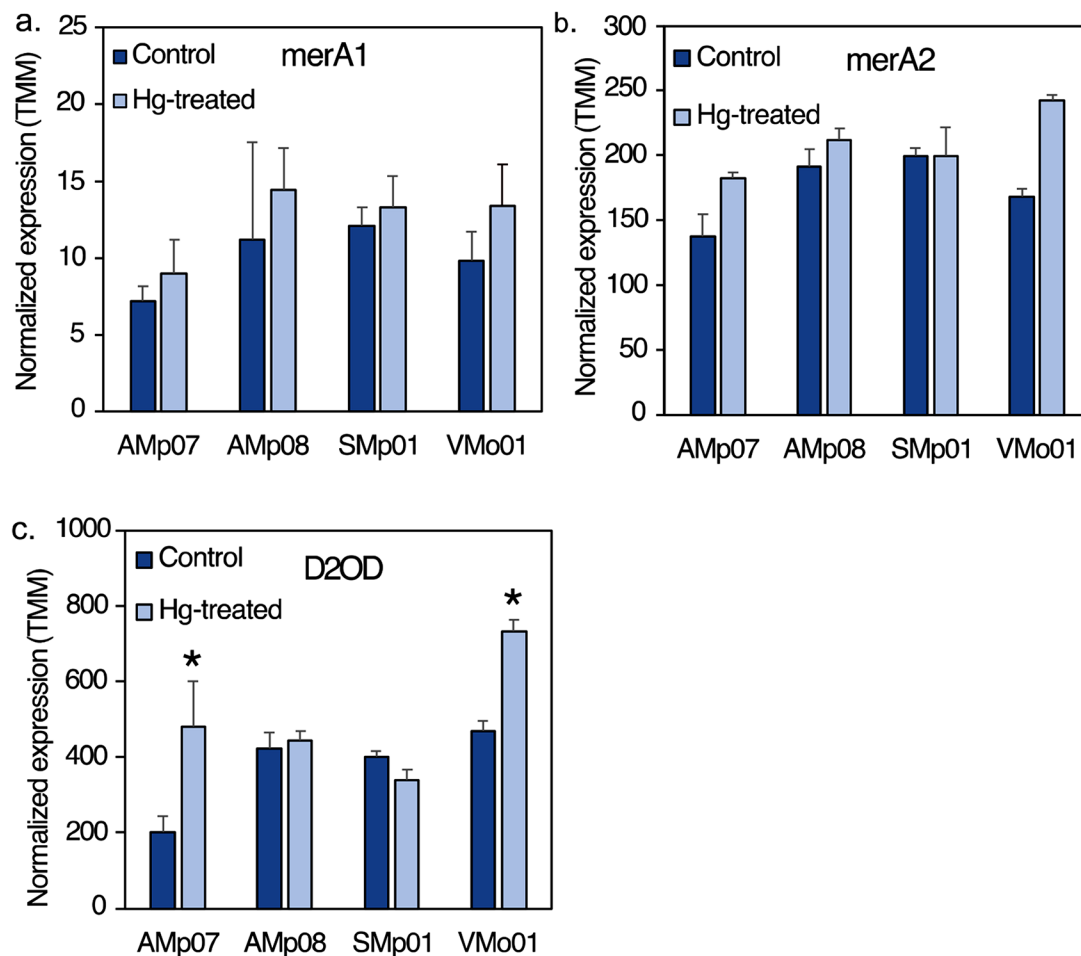


Fig. 6 Changes in expression of mercuric ion reductase gene homologs in different strains following Hg treatment. **(a)** Mercury ion reductase (*merA1*) showed no significant change in response to Hg treatment in any of the strains studied, although the highest expression was found in AMp08. **(b)** Mercury ion reductase (*merA2*) shows no significant change in response to Hg stress in any of the strains tested. All three of these genes, *merA1*, *merA2* and D2OD are homologous to the MerA gene in Fig. S2 and Fig. 5 **(c)** A gene annotated as a Dihydropyridine 2-oxoglutarate dehydrogenase (D2OD), which is homologous to *merA1* and *merA2* and shows a significant upregulation in LT strains (AMp07 and VMo01), and no significant changes in the tolerant strains (AMp08, SMp01). The normalized expression levels are reported as mean TMM values of three biological replicates

For the HT strain AMp08 inside nodules, similar to the free-living conditions, we found the majority of DEGs were located on the Mer operon including MerA, MerT, and glutathione reductase, which had a 3-12-fold increase in expression levels in response to Hg stress in both host-plant genotypes. The total number of shared DEGs between free-living conditions and inside of nodules for the tolerant strain AMp08 amounted to only 0.6% of the DEGs (3 genes total), which included MerA and MerT, indicating the critical role of Mer operon genes in Hg detoxification in free-living conditions and nodules (Fig S7). The *merA1* and *merA2*, which are not present on the operon, showed different patterns than in free-living conditions where *merA1* had greater upregulation but was not significant (Fig. S9). For the LT strain AMp07, 2.8% of the DEGs were common between free-living conditions and the nodules (Fig. S7). The shared

DEGs included multiple genes from the ABC-transporter family which are known to play a role in metal ion transport and cellular detoxification in plants and bacteria. We found that the identity of most of the DEGs under free-living conditions and in nodules is largely different, which could be due to many factors including the cellular replication of rhizobia inside the nodules, signaling between host-plant and rhizobia, and the dilution of the stress by some contribution from the host-plant response. But overall, the response of the genes in the *S. medicae* Mer operon appear to be playing the key role in Hg detoxification based on their up-regulation patterns in the nodules, which is a similar trend in free-living conditions (Fig. 5, Fig. S8).

When globally examining the most significant DEGs in the HT strain AMp08 (with lowest adjusted p-values shown in Fig. 7), we found certain genes such as

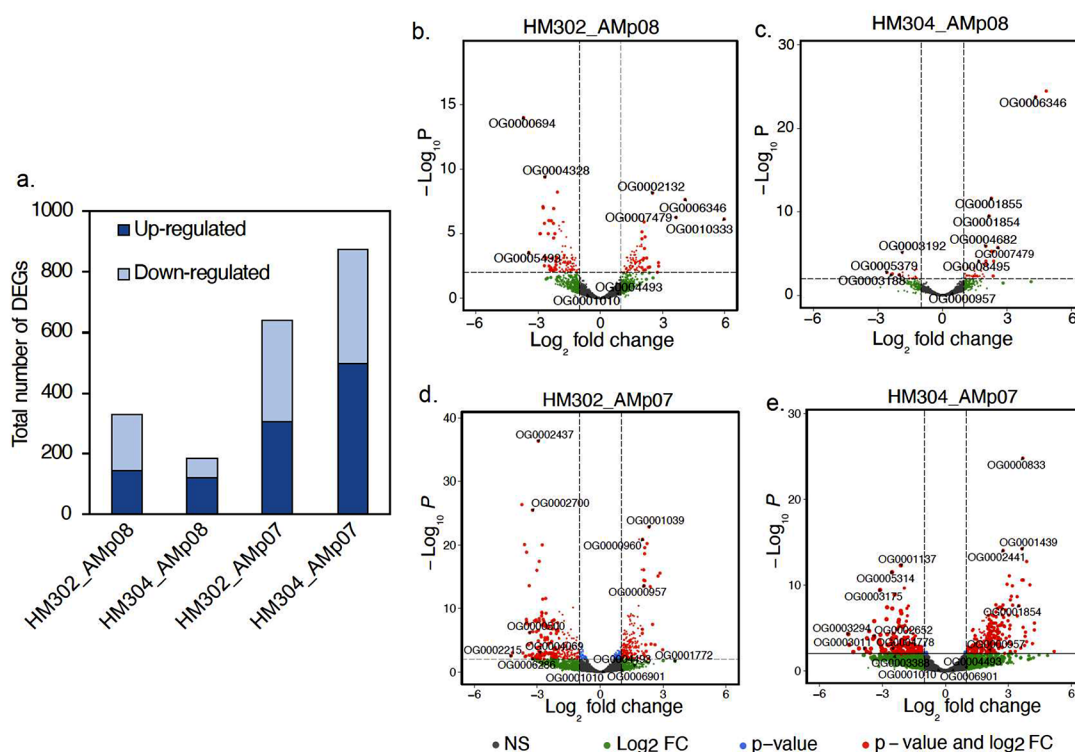


Fig. 7 Total number of differentially expressed genes (DEGs) for rhizobia present in nodules following Hg treatment (a). The host-plants inoculated with rhizobia were treated with either 0 μM HgCl_2 (Control) or 100 μM HgCl_2 (Hg-treated). DEGs were calculated based on the criteria $|\text{log}_2\text{FC}| \geq 1$, adjusted $p < 0.05$. Volcano plots for DEGs for host plants inoculated with AMp08 (HM302_Amp08 (b), HM304_Amp08(c)), and host plants inoculated with AMp07 (HM302_Amp07 (d), and HM304_Amp07 (e)). Each point on the volcano plot represents a DEG, and the top significantly upregulated and downregulated DEGs were labeled using their respective Orthogroup IDs

polyribonucleotide nucleotidyltransferase to be highly upregulated in rhizobia when inside the HM302 host plant, while phosphoserine aminotransferase and D-3-phosphoglycerate dehydrogenase were significantly upregulated in the HM304 host plant (Fig. 7b, c; Supplementary Data 1, Table S3). For the LT strain AMp07 inside HM302 nodules, we found significant downregulation of NAD-dependent protein deacetylase (SIR2 family), manganese catalase, and ABC-transporter substrate-binding protein, while genes such as Na^+/H^+ -dicarboxylate symporter, malate dehydrogenase and dihydrolipoamide dehydrogenase of 2-oxoglutarate dehydrogenase were significantly upregulated (Fig. 7d). Inside of nodules of the HM304 host-plant, AMp07 genes that were differentially expressed included ATP synthase F0 sector subunit a, nitric oxide reductase activation protein, D-3-phosphoglycerate dehydrogenase, all of which were upregulated, while glutathione S-transferase, kup system potassium uptake protein, and ABC-transporter permease protein were downregulated (Fig. 7e). These results indicate there are expression patterns that are unique to the rhizobia strains in different host-plant environments and responses depend on the level of tolerance to Hg by both the rhizobia strain and the host-plant genotype.

To visualize strain specific patterns of differential expression between free-living and nodule environments, we selected the top 10 significantly upregulated and top 10 significantly downregulated genes in each strain, then compared the expression of these 20 genes per strain across all the strains for a total of 160 genes (Fig. 8, Supplementary Data 1, Table S4). This allowed us to visually compare clusters of the most differentially expressed genes in each strain for sets of homologous genes in the other strains. For each strain we found distinct clusters of significantly high- and low-expressed genes which were often not differentially expressed in other strains, and typically coincided with their tolerance levels (measured using MIC) of the strain. For the tolerant strains, there was a clear cluster of genes corresponding to the Mer operon being significantly upregulated in both free-living conditions and in nodules, while no such cluster was observed for the downregulated genes.

For the LT strains growing in free-living conditions, we observed a cluster of downregulated genes belonging to the family of cytochrome c oxidase subunits as well as genes such as pyridoxamine 5'-phosphate oxidase-related and L-proline 3-hydroxylase. By contrast, clusters of up-regulated genes included transporters such as

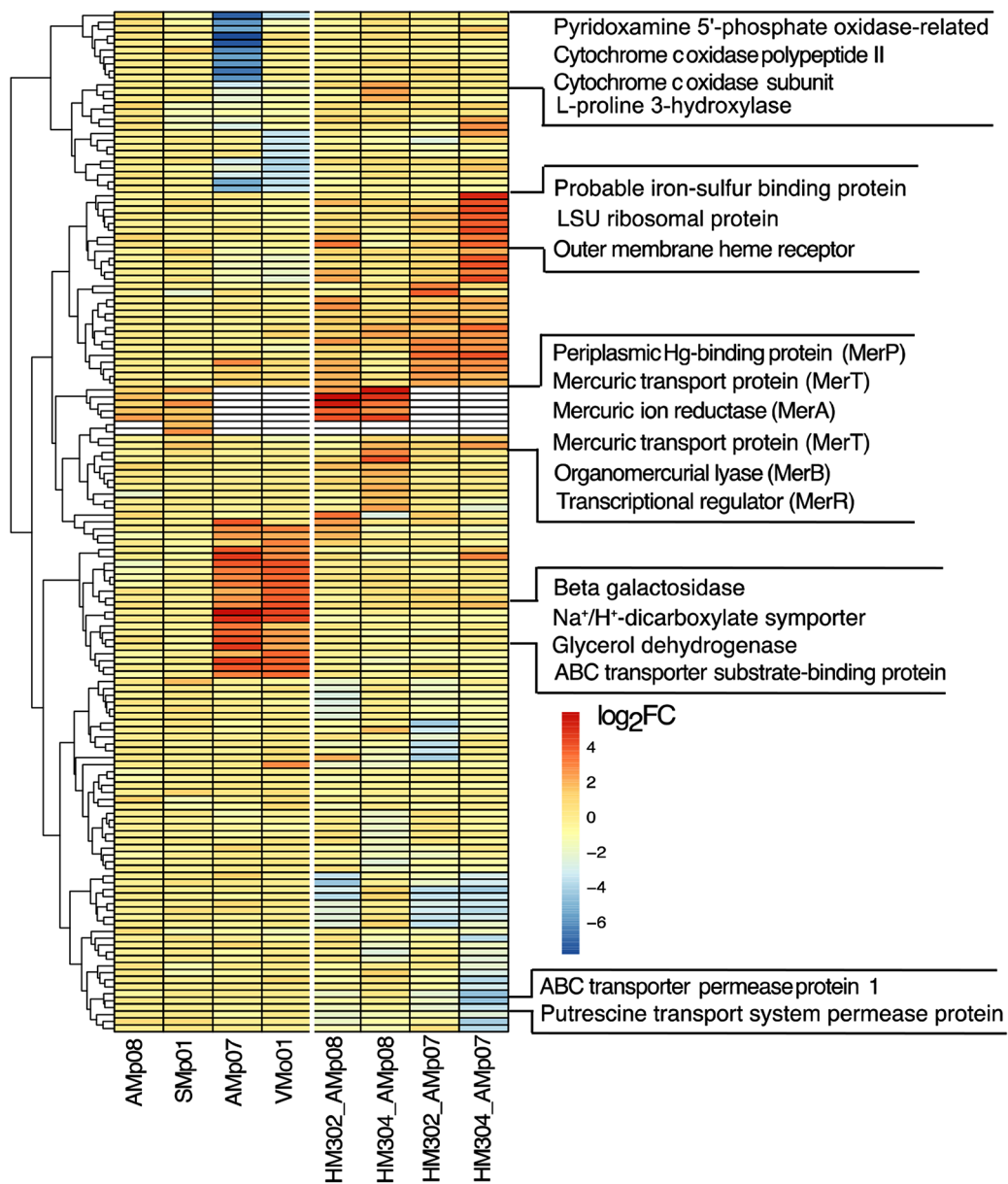


Fig. 8 Heatmap representing the top 20 DEGs (top 10 upregulated and top 10 downregulated DEGs) of each sample set for a total of 160 DEGs across all samples encompassing free-living and nodules conditions. On the y-axis, genes and gene families with biological relevance to Hg transport and detoxification were labeled. For the rhizobia strains with no mercuric ion reductase (Mer) operon genes (MerP, MerT, MerA, MerB, MerR), the cells that are white indicate no data point as these strains do not possess a Mer operon or one or more of the Mer operon genes

ABC-transporter substrate-binding proteins, Na⁺/H⁺ dicarboxylate symporter genes, beta galactosidase and glycerol dehydrogenases. Inside nodules, the LT strain AMp07 had several highly expressed genes including an iron-sulphur binding protein and an outer membrane heme receptor clustering together, while several permeases such as ABC-transporter permease protein and putrescine transport system permease protein formed a distinct cluster of low-expressed genes. Taken together, this analysis allows us to distinguish the responses of Hg-tolerant strains from LT strains based on the distinct

cluster of genes that are differentially expressed in each strain in response to Hg stress, which showed a much broader impact on the transcriptome while in symbiosis than in free-living conditions.

Hg-stress has an impact on expression of nitrogen fixation (*nif*) genes in symbiosis with host-plants

We conducted GO-enrichment of DEGs from the rhizobia transcriptome in free living conditions and in the nodules to better understand the broader response to Hg stress in *S. medicae* outside of the Mer operon or merA1

and *merA2* genes. Overall, GO-enrichment was not very powerful to identify GO-terms in the HT strains as there were too few DEGs to detect enrichment (Supplementary Data 1, Table S5) but it led us to the GO-term for nitrogen fixation (GO:0009399) in nodules, which came from differentially expressed nitrogen fixation (*nif*) genes in the HM304-AMp08 host-plant-rhizobia combination. We therefore compared expression of *nif* genes in the two strains used to inoculate the host-plant to determine whether *nif* regulation was predicted by the tolerance of the strains in nodule conditions. In nodules, the HT strain AMp08 with the higher Hg-accumulating phenotype host plant (HM304), showed an upregulation of the *nif* genes in response to Hg treatment, whereas the expression levels of the *nif* genes in AMp08 were down-regulated inside the nodules of the low accumulating (HM302) plant genotype (Fig. 9a, c). By contrast for the LT strain AMp07, we observed a significant downregulation of *nif* genes (*nifA*, *nifB*, *nifT* and FNA) in AMp07 inside the nodules of both host plant genotype (Fig. 9b,

d), suggesting that *nif* genes are negatively affected by Hg stress in the LT rhizobia strain regardless of the host-plant tolerance. The finding that AMp08 *nif* genes were up-regulated inside the HM304 (high Hg accumulator) host plant, but down-regulated in the low accumulating HM302 host plant, suggests a genotype-by-genotype interaction which may allow for better nitrogen fixation by the AMp08 strain in the presence of Hg, if the host plant also possesses higher Hg-stress detoxification capacity.

The *mer* operon from Almadén strain AMp08 can confer Hg-tolerance to low-tolerant strains through horizontal gene transfer

Our transcriptomics analysis showed the genes present on the Mer operon were the most significant DEGs in the HT AMp08 strain under free-living conditions and in nodules (e.g. MerT, MerA). Because all strains contain *merA1* and *merA2* and show some response to Hg treatments, but probably not enough to confer hypertolerance

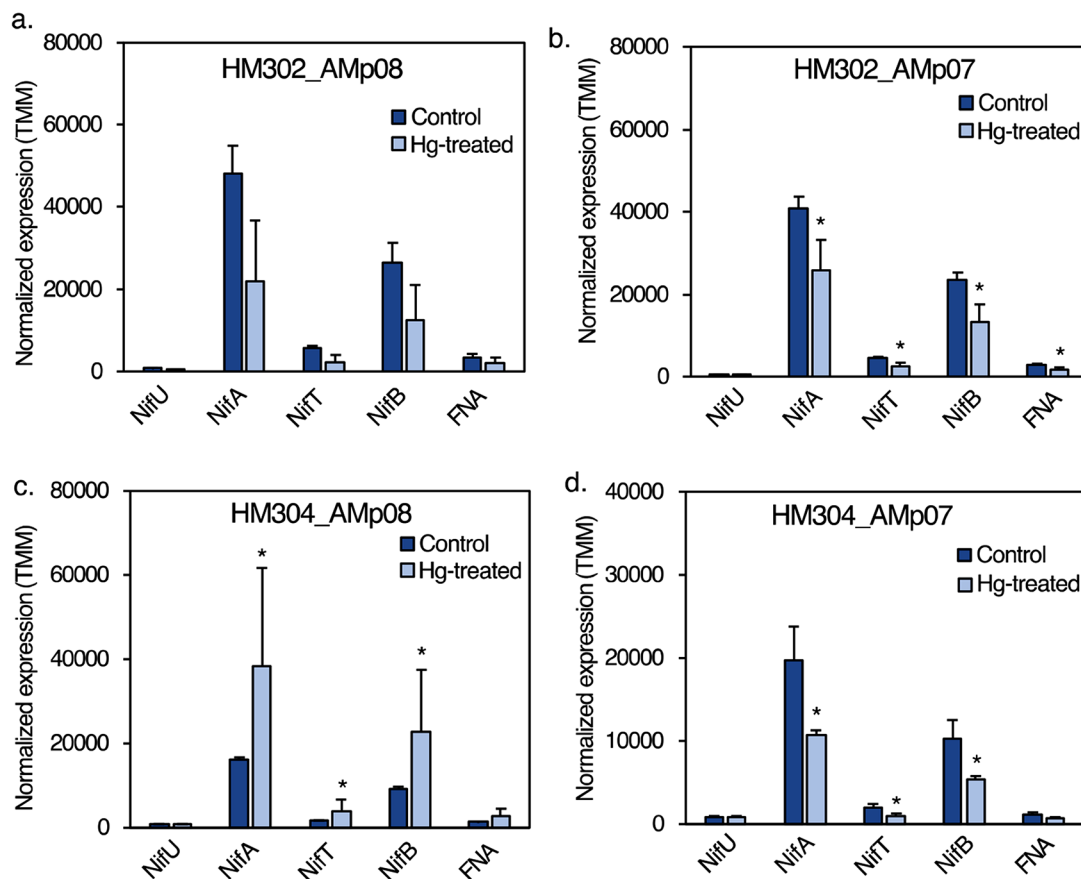


Fig. 9 Expression of *nif* genes in rhizobia present in nodules under Hg treatment. In the nodules formed by tolerant strain AMp08, the expression levels of *nifU*, *nifA*, *nifT* and 4Fe-4 S ferredoxin nitrogenase-associated gene (FNA) were significantly downregulated when the host plant was HM302 (a), and significantly upregulated when the host-plant was HM304 (b). In the nodules formed by LT strain AMp07, the expression levels of *nifU*, *nifA*, *nifT* and 4Fe-4 S ferredoxin nitrogenase-associated gene (FNA) were significantly downregulated when the host plant was either HM302 (c) or HM304 (d). The normalized expression levels are shown as the mean TMM values of three biological replicates

like in AMp08 or SMp01, we set out to demonstrate that the Mer operon was essential for conferring Hg tolerance. We isolated the 71 kb accessory plasmid from AMp08 (Fig. S2) harboring the Mer operon, then transferred the entire plasmid to the two LT strains AMp07 and WSM419 using electroporation. Without the Mer operon, both the wild-type strains, AMp07 and WSM419 could grow at 25 μM HgCl_2 (Fig. 10), but only the wild-type AMp08 strain grew at greater concentrations. We therefore tested the ability of the transformed strains to grow on higher concentrations of HgCl_2 than the wild-type strains without the plasmid (Fig. 10a). After the Mer operon was transferred into AMp07 and WSM419 strains, they gained the ability to grow at 250 μM HgCl_2 , which is equivalent to the tolerant AMp08 strain (Fig. 10a). We confirmed the positive transfer of the Mer operon by amplifying MerA and MerT that are only present on the operon (Fig. 10b-c), while all strains contained merA1 (located on main chromosome) as expected (Fig. 10d). To compare the relative expression levels of the Mer genes in the non-transformed and transformed strains grown in the presence or absence of 4 μM HgCl_2 , we performed qPCR (primers listed in Supplementary Data 1, Table S6). While we did not observe any significant change in the expression levels of merA1 and merA2 in any of the strains tested, we found upregulation in the genes present on the Mer operon (MerA, MerT) which had been transferred into AMp07 and WSM419 (Fig. 10e-g), similar to the native Mer operon of the wild-type AMp08 strain. These results confirm that the Mer operon is key to Hg tolerance, and the tolerance can be transferred to other low-tolerant strains via horizontal gene transfer.

Discussion

The genomes we sequenced have shown the presence of a complete Mer operon in three of the most Hg-tolerant strains known in rhizobia. The Mer operon constitutes a structural adaptation in these genomes which allowed them to tolerate high levels of Hg exposure at Almadén. In the *S. medicae* genomes, we identified four mercury reductase A homologs (merA1, merA2, D2OD, MerA) which were all present in the HT strains, while the LT strains lacked the MerA gene. This new finding advanced our understanding of Hg-tolerance in these strains which was previously attributed to differential expression and regulatory changes in the mercury reductase genes, merA1 and merA2 [32], which are not part of a Mer operon. Protein structure of the merA homologs showed high homology among each other, and three other species based on 3D conservation mapping of active sites using Dali. This suggests conserved function, but the four homologs nevertheless showed distinct clades with strong branch support in our phylogenetic analysis,

indicated that there are differences at individual amino acids. Rigorous testing of conserved function among all the merA homologs will require in vitro tests using isolated proteins of each homolog. The fused MerBA has not yet been shown in α -proteobacteria and provides even more novelty among the three Mer operons that we sequenced. Because the MerB homodimer domains are connected by a loop in a single polypeptide chain to the MerA homodimers, we speculate that conversion of methylmercury and transfer of H^{2+} to the MerA domains could be quicker and more efficient than in non-fused MerA and MerB. Testing this also requires functional analyses using in vitro assays.

The synteny and phylogenetic analyses showed that the Mer operons in the three strains were most likely acquired independently, which is a remarkable example of convergent evolution at a local geographic scale. The diverse synteny is also consistent with previous observations of high levels of diversity in gene organization and structural diversity of Mer operons in bacteria in general [75]. Recently, it has been shown that varying degrees of metal stress in the local population may increase the frequency of HGT events [26], providing a source of adaptive evolution in bacteria [25] in contaminated environments. Moreover, HGT is also crucial in rhizobia to determine their host range [76]. And while we typically expect selection to act on standing genetic variation in a species [77], the mechanism of HGT can cross species boundaries, accelerating adaptation, even if the donor genetic locus is at low frequency in the population [78] or in the microbiome of the rhizosphere. To our knowledge, the presence of a functional Mer operon and its relationship to other mercury reductases in α -proteobacteria has yet to be reported, despite the considerable number of genomic studies of nitrogen-fixing bacteria in this group [43, 79], including strains collected from heavy metal contaminated sites [13, 14, 30, 31].

Regulatory changes in Hg-tolerant rhizobia in free-living conditions compared to low-tolerant strains

Transcriptomic analysis showed several interesting patterns that were dependent on the presence of the Mer operon as well as the environmental conditions associated with growth of the rhizobia (i.e., whether free-living or in symbiosis). The number of DEGs in the two Hg-tolerant strains in free-living conditions was remarkably low (~20–30 genes), while the LT strains had nearly 25 times more DEGs. Among the most significant DEGs, nearly all genes located in the Mer operon in both Hg-tolerant strains were significantly up-regulated. However, there was an exception, the regulatory gene MerR showed constitutive expression in both control and Hg-treated conditions. The very low number of DEGs suggests the rapid and efficient mechanism of Hg^{2+} transport, volatilization,

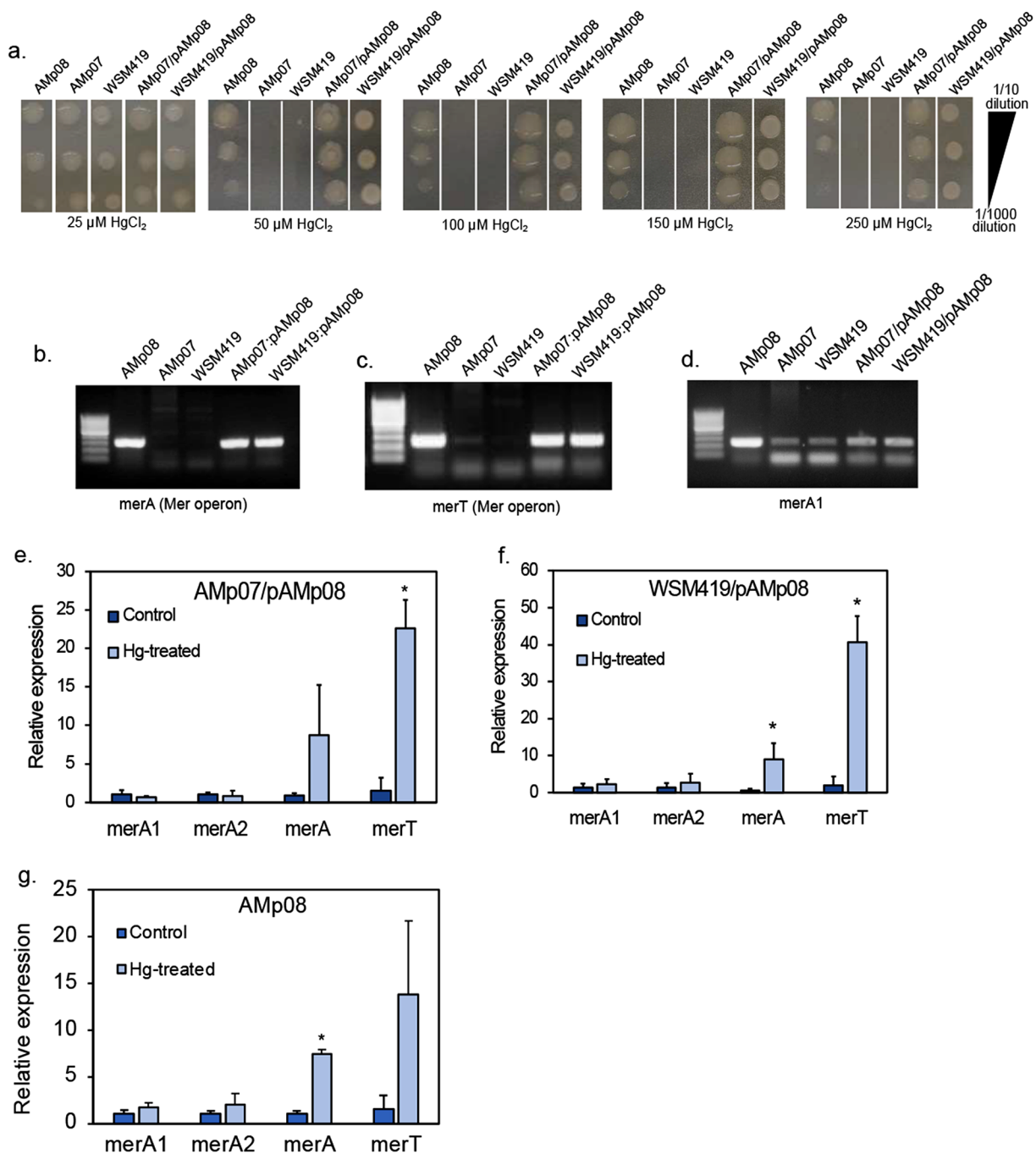


Fig. 10 Minimum inhibition concentration (MIC) assay shows the gain of Hg tolerance in LT strains transformed with plasmid from AMp08 (AMp07/pAMp08; WSM419/pAMp08). Each MIC image represents *S. medicae* strains from left to right: AMp08, AMp07, WSM419, AMp07/pAMp08, WSM419/pAMp08, grown on plates containing 25 μM HgCl₂, 50 μM HgCl₂, 100 μM HgCl₂, 150 μM HgCl₂, 200 μM HgCl₂, 250 μM HgCl₂ respectively. The dilutions are 1:10, 1:100, 1:1000 from top to bottom (a) Gel images show the presence of MerT and MerA in AMp08 and the transformed strains AMp07/pAMp08, WSM419/pAMp08, but not in the untransformed strains (AMp07, WSM419) that lack a Mer operon (b, c). Gel image showing the presence of merA1 (present on the main chromosome) in all the strains (d). The ladder size scale on each gel image is 50 bp. Expression changes in merA1, merA2, MerA, and MerT in AMp07/pAMp08 and WSM419/pAMp08 quantified using qPCR (e, f). Relative expression levels of merA1 (present on the main chromosome), merA2 (present on pSymb) and MerA, MerT (present on the Mer operon accessory plasmid) were calculated using 16 S rRNA as an endogenous control. Three independent experiments were conducted and mean values ± SD are represented. Asterisks indicate p-value < 0.05 calculated using a Student's T-test

and elimination from the cell results in extremely low impact on free-living rhizobia. The low number of DEGs in the Hg-tolerant strains appears to be in line with heavy metal tolerant *Mesorhizobium metallidurans* [30] and *Sinorhizobium meliloti* [31], which also showed few genes differentially expressed (~1.0% or less of the genome showing significant responses), but these studies did not compare tolerant strains with less tolerant strains from the same population or elsewhere. We suspect that adaptation to high levels of Hg (and other toxic heavy metals) in the soil and rhizosphere is most important in free-living conditions due to direct exposure to the metal ions in the environment. However, we must concede that the level of Hg stress in our free-living experiment (4 μM HgCl_2) was a low dose that was necessary for obtaining transcriptomic data from both the LT strains and tolerant strains, but higher doses would potentially result in greater numbers of stress response DEGs in the HT strains.

The non-operon mercury reductases (merA1 and merA2) showed only slight up-regulation between control vs. treatment conditions and appeared more like constitutively expressed genes and showed only slight differences between HT and LT strains. However, the highest expression of merA1 in Hg-treated conditions was in our most tolerant strain, AMp08, which may provide additional tolerance above what the Mer operon provides, but this would need further validation experiments. Overall, our findings suggests that merA1 and merA2 in *S. medicae* (and likely *S. meliloti*) function in basic Hg^{2+} detoxification but they cannot be optimized by *cis*-regulation to achieve HT to the scale that the Mer operon provides, as indicated by the LT strains from the same population. Moreover, we found almost no amino acid sequence diversity for the merA1 gene and this gene appears fixed in *Sinorhizobium*, which would fit the HGT followed by gene specific sweep model shown by [25], which could mean merA1 was also the product of an earlier HGT event and other Mer operon genes (such as MerB, MerC, MerP, etc.) were lost in environments where only basic tolerance to Hg was necessary, since high Hg contamination in soils is uncommon. Furthermore, the very low diversity for the merA2 gene among the several strains that were included in the phylogenetic analysis (including the two non-Almadén strains *S. medicae* WSM419 and *S. meliloti* 1021), is consistent with strong selective constraint on both merA1 and merA2 in *Sinorhizobium* to confer basic tolerance to Hg (i.e., < 25 μM) in free-living conditions. Finally, we speculate that the significant up-regulation of D2OD in the two low Hg-tolerant strains but not in the two most Hg-tolerant strains that possess a Mer operon, may be the result of a local adaptation (i.e., enhanced *cis*-regulation) that may contribute to an inducible mercury reduction process

due to high exposure to Hg^{2+} in their native environment at Almadén, but further study into the function of this gene in Hg detoxification is necessary.

Expression of genes inside bacteroids also shows mer genes dominate the response to hg stress

Because rhizobia also exist in symbiotic life-stages, it was important to examine gene expression patterns while inside of nodules of legume host plants. It is interesting to find a similar trend of fewer DEGs in the Hg-tolerant strain inside bacteroids, but the total number of DEGs for both tolerant and LT strains is higher than in free-living conditions. This could be attributed to the higher concentration of the stress treatment (100 μM Hg compared to 4 μM Hg under free-living conditions), but the differences in free-living and nodule life stages cannot be underestimated due to the tremendous signaling between host-plant and rhizobia that occurs to maintain homeostasis during stress responses in nodules. Like in free-living conditions, some of the most significant genes in nodules were those found on the Mer operon. Globally, genome-wide stress responses based on the most differentially expressed genes in the LT strains revealed many expected types of genes including oxidative stress related genes, cytochrome c oxidases, cytochrome O-ubiquinol oxidase, nitric oxide dioxygenase, ABC-transporters, and other membrane related transporters.

Due to higher number of DEGs compared with free-living conditions, GO-enrichment of the top ranked DEGs revealed some genes that may be involved in the oxidative stress and cellular detoxification response that resulted from the Hg exposure to the rhizobia in the nodules of the host-plant. The GO-term for nitrogen fixation (GO:0009399) was enriched in bacteroids, which included several nitrogen fixation (*nif*) genes which regulate nitrogenase levels. The effect of regulatory changes on *nif* genes could provide a more direct link to stress impacts on nitrogen fixation than the more general responses of oxidative stress. In a previous study using the AMp08 genotype in *M. truncatula* nodules, [32] did not detect any reduction in nitrogenase activity in plants treated with 500 μM Hg when compared with non-treated plants, but the expression levels of neither *nif* nor other nitrogen fixation genes were measured. The *nif* genes in the AMp08 strain showed no significant changes in free-living conditions but in nodules, both the tolerant strain AMp08 and LT strain AMp07 showed down-regulation while residing in the less Hg-accumulating host plant genotype (HM302). In the higher Hg-accumulating genotype (HM304), AMp08 showed up-regulation but AMp07 showed down-regulation of *nif* genes. Therefore, if up-regulation of *nif* genes is positively correlated with nitrogenase production, then the host-plant genotype x rhizobia strain interaction may also be important for

nitrogen fixation under stress conditions. Because higher nitrogen content in the host plants will improve plant fitness under stress conditions, by proxy the nodule number phenotype could be the result of the rhizobia strain tolerance and its ability to fix nitrogen under Hg stress conditions [34].

Horizontal gene transfer of the mer operon accessory plasmid achieved hypertolerance in *S. medicae*

Regardless of the role of other genes involved in dealing with Hg stress, our horizontal transfer experiment clearly indicated that an accessory plasmid containing the Mer operon locus alone would suffice to confer HT to Hg stress in *S. medicae*. This is consistent with functional studies of the Mer operon in other bacteria [16]. The gene expression patterns in free-living conditions and inside of nodules of highly up-regulated Mer operon genes, and the immediate acquisition of HT to Hg when the Mer operon was transferred into LT strains, strongly suggests that the Mer operon is the source of this high level of tolerance. Importantly, because we replicated the HGT by using one LT strain from Almadén (AMp07) and one non- Almadén strain (WSM419), which served as a useful control in the case that Almadén strains could more easily acquire tolerance due to local exposure to Hg in the environment. MerT, the essential transport protein which facilitates the transport of the Hg²⁺ ion from MerP to MerA [80] was one of the highest expressed genes present on the Mer operon in any context, including free-living conditions, nodules and following HGT. MerT is likely one of the most important transporters in the operon, next to MerA and MerB [81]. The Mer operon is unique among heavy metal transport mechanisms in that it can provide a near complete source of cellular detoxification and elimination of the target metal, unlike Cd, Pb or Zn which showed that knocking out several genes is needed to reduce tolerance to these metals in a tolerant strain of *S. meliloti* isolated from a mine [14, 31]. By contrast, in *Mesorhizobium metalladurins*, regulation of Zn in a few strains that were isolated from a heavy metal contaminated site appears to be due to enhanced *cis*-promoter driven regulation of the P_{II}-type ATPase, *cadA* to achieve high tolerance [13, 30]. However, because of the correlated up-regulation of the adjacent genes in Mer operon in the Hg-tolerant Almadén strains, including hypothetical genes with no homology to known Hg transporters, there is the potential to assess incremental losses in tolerance by doing knock-outs of these genes. Such studies would identify the role of each gene linked within the operon, and the value of their up-regulation in connection with a tolerance phenotype. To summarize, the complete Mer operon is necessary for HT based on the strains we have isolated from Almadén, but the syntenic and structural diversity (variation in gene order

and presence/absence variation) among α -proteobacteria, combined with extremely high upregulation of the Mer operon genes suggest that *cis*-regulation within the operon also may also play a very large role in achieving the level of tolerance that we observe in the collection of rhizobia from Almadén (i.e., MIC up to 250 μ M of Hg) which may have evolved further due to exposure and contact with Hg in this highly toxic environment. We therefore speculate the operon likely evolved by a succession of horizontal gene transfer followed by *cis*-regulatory evolution of Mer genes. The presence of a Hg tolerance adaptation in symbiotic α -proteobacteria strains which can be used to fix nitrogen for host plants, improve nitrogen in soils at the entire plant community/population scale, while also contributing to heavy metal detoxification, may be an advantage to future efforts in removing toxins from plants [82], and also in bioremediation using legume-rhizobia interactions [15].

Abbreviations

Mer	Mercury reductase operon
merA	Mercury reductase A
Hg	Mercury
HT	High tolerant
LT	Low tolerant
MIC	Minimum inhibitory concentration
HGT	Horizontal gene transfer
IMG	Integrated Microbial Genome database ()

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s12866-024-03391-5>.

Supplementary Material 1

Supplementary Material 2

Acknowledgements

We would like to thank Katy Heath, Alvaro Hernandez and Chris Wight at the University of Illinois at Urbana-Champaign for consulting, oligo design, and rRNA cleanup and sequencing for the nodule samples. We thank Sanhita Chakraborty at the Institute for Advancing Health through Agriculture at Texas A&M for useful comments and interpretation of data.

Author contributions

TP and JJP conceived and designed the study and isolated DNA for whole genome sequencing. BE and PT performed genome assembly. TP, BE performed gene annotation. KD performed protein structure analysis. TP, AM, RB and TW conducted synteny analysis. MML conducted free-living rhizobia experiment and extracted RNA. AB and RS conducted nodule experiment and extracted RNA. TP and AB performed differential expression analysis. TP and AB drafted the manuscript and Figures with editorial contributions from coauthors.

Funding

for this project came from the Department of Energy Quantitative Plant Science Initiative SFA, USDA-ARS Project Number 3092-53000-001-000D, and the Agencia Estatal de Investigación, Spain, grant number PID2021-125371OB-I00 awarded to JJP. The work (Community Science Proposal – Functional Genomics: <https://doi.org/10.46936/10.25585/60001331>) conducted by the U.S. Department of Energy Joint Genome Institute (<https://ror.org/04xm1d337>), a DOE Office of Science User Facility, is supported by the Office of Science of the U.S. Department of Energy operated under Contract No. DE-AC02-05CH11231.

Data availability

Raw data for genome assemblies (PacBio reads) <https://www.ncbi.nlm.nih.gov/bioproject/?term=PRJNA565470>. Raw data for free-living rhizobia RNA-seq (150 bp paired end Illumina reads) <https://www.ncbi.nlm.nih.gov/bioproject/?term=PRJNA996196>. Raw data for rhizobia reads from dual-transcriptomics (150 bp paired-end Illumina reads) <https://www.ncbi.nlm.nih.gov/bioproject/?term=PRJNA986885>.

Declarations**Ethics approval and consent to participate**

Plant material was obtained from the Medicago HapMap collection <https://medicago.legumeinfo.org/tools/germplasm/>.

Consent for publication

Not applicable.

Competing interests

The authors declare no competing interests.

Author details

¹Brookhaven National Laboratory, Upton, USA

²Institute of Agricultural Sciences, ICA-CSIC, Madrid, Spain

³Institute for Advancing Health Through Agriculture, Texas A&M, College Station, TX, USA

⁴DOE Joint Genome Institute, Lawrence Berkeley National Laboratory, Berkeley, CA, USA

⁵Plant and Microbial Biology, University of Minnesota, St. Paul, Minnesota, USA

⁶USDA-ARS Children's Nutrition Research Center, Houston, TX, USA

Received: 21 March 2024 / Accepted: 19 June 2024

Published online: 06 July 2024

References

- Driscoll CT, Mason RP, Chan HM, Jacob DJ, Pirrone N. Mercury as a global pollutant: sources, pathways, and effects. *Environ Sci Technol*. 2013;47:4967–83.
- Alloway BJ. Sources of heavy metals and metalloids in Soils. In: Alloway BJ, editor. *Heavy metals in Soils: Trace metals and metalloids in Soils and their bioavailability*. Dordrecht: Springer Netherlands; 2013. pp. 11–50.
- Tchounwou PB, Yedjou CG, Patlolla AK, Sutton DJ. Heavy metals toxicity and the Environment. *EXS*. 2012;101:133–64.
- Rai PK, Lee SS, Zhang M, Tsang YF, Kim K-H. Heavy metals in food crops: Health risks, fate, mechanisms, and management. *Environ Int*. 2019;125:365–85.
- Alengebawy A, Abdelkhalik ST, Qureshi SR, Wang M-Q. Heavy metals and pesticides toxicity in agricultural soil and plants: ecological risks and Human Health implications. *Toxics*. 2021;9:42.
- Poulain AJ, Barkay T. Cracking the Mercury methylation code. *Science*. 2013;339:1280–1.
- Barkay T, Gu B. Demethylation The Other Side of the Mercury methylation Coin: a critical review. *ACS Environ Au*. 2021. acsenvironau.1c00022.
- Parks JM, Johs A, Podar M, Bridou R, Hurt RA, Smith SD, et al. The genetic basis for bacterial Mercury methylation. *Science*. 2013;339:1332–5.
- Barkay T, Miller SM, Summers AO. Bacterial mercury resistance from atoms to ecosystems. *FEMS Microbiol Rev*. 2003;27:355–84.
- Barkay T, Wagner-Döbler I. Microbial transformations of Mercury: potentials, challenges, and achievements in Controlling Mercury Toxicity in the Environment. *Advances in Applied Microbiology*. Elsevier; 2005. pp. 1–52.
- Boyd ES, Barkay T. The Mercury Resistance Operon: from an origin in a geothermal environment to an efficient detoxification machine. *Front Microbiol*. 2012;3.
- Christakis CA, Barkay T, Boyd ES. Expanded diversity and phylogeny of mer genes broadens Mercury Resistance paradigms and reveals an origin for MerA among Thermophilic Archaea. *Front Microbiol*. 2021;12:682605.
- Maynaud G, Brunel B, Yashiro E, Mergeay M, Cleyet-Marel J-C, Le Quére A. CadA of Mesorhizobium metallidurans isolated from a zinc-rich mining soil is a PIB-2-type ATPase involved in cadmium and zinc resistance. *Res Microbiol*. 2014;165:175–89.
- Lu M, Li Z, Liang J, Wei Y, Rensing C, Wei G. Zinc resistance mechanisms of PIB-type ATPases in Sinorhizobium meliloti CCNWSX0020. *Sci Rep*. 2016;6:29355.
- Fagorzi C, Checcucci A, diCenzo G, Debiec-Andrzejewska K, Dziejew L, Pini F, et al. Harnessing Rhizobia to Improve Heavy-Metal phytoremediation by Legumes. *Genes*. 2018;9:542.
- Petrus AK, Rutner C, Liu S, Wang Y, Wiatrowski HA. Mercury Reduction and Methyl Mercury degradation by the Soil Bacterium Xanthobacter autotrophicus Py2. *Appl Environ Microbiol*. 2015;81:7833–8.
- Hong B, Nauss R, Harwood IM, Miller SM. Direct measurement of Mercury(II) removal from Organomercurial Lyase (MerB) by Tryptophan fluorescence: NmerA Domain of Coevolved γ -Proteobacterial Mercuric Ion Reductase (MerA) is more efficient than MerA Catalytic Core or glutathione. *Biochemistry*. 2010;49:8187–96.
- Jones KM, Kobayashi H, Davies BW, Taga ME, Walker GC. How rhizobial symbionts invade plants: the Sinorhizobium–Medicago model. *Nat Rev Microbiol*. 2007;5:619–33.
- Reeve W, Chain P, O'Hara G, Ardley J, Nandesena K, Bräu L, et al. Complete genome sequence of the Medicago microsymbiont Ensifer (Sinorhizobium) medicae strain WSM419. *Stand Genomic Sci*. 2010;2:77–86.
- Reeve W, O'Hara G, Chain P, Ardley J, Bräu L, Nandesena K, et al. Complete genome sequence of Rhizobium leguminosarum Bv. Trifolii strain WSM1325, an effective microsymbiont of annual Mediterranean clovers. *Stand Genomic Sci*. 2010;2:347–56.
- Mendoza-Suárez MA, Geddes BA, Sánchez-Cañizares C, Ramírez-González RH, Kirchelle C, Jorriñ B, et al. Optimizing *Rhizobium* legume symbioses by simultaneous measurement of rhizobial competitiveness and N_2 fixation in nodules. *Proc Natl Acad Sci USA*. 2020;117:9822–31.
- Hawkins JP, Oresnik IJ. The Rhizobium-Legume Symbiosis: co-opting successful stress management. *Front Plant Sci*. 2022;12:796045.
- Shwetha Raghupathy, Arunachalam S. Trends in Legume-Rhizobia Symbiosis in Remediation of Mercury-Contaminated Agricultural soils. *Commun Soil Sci Plant Anal*. 2023;1–15.
- Bobay L-M, Ochman H. The evolution of Bacterial Genome Architecture. *Front Genet*. 2017;8:72.
- Arnold BJ, Huang I-T, Hanage WP. Horizontal gene transfer and adaptive evolution in bacteria. *Nat Rev Microbiol*. 2022;20:206–18.
- Chen X, Du Z, Song X, Wang L, Wei Z, Jia L, et al. Evaluating the occurrence frequency of horizontal gene transfer induced by different degrees of heavy metal stress. *J Clean Prod*. 2023;382:135371.
- Maynaud G, Willems A, Sousou S, Vidal C, Mauré L, Moulin L, et al. Molecular and phenotypic characterization of strains nodulating Anthyllis vulneraria in mine tailings, and proposal of Aminobacter anthyllidis sp. nov., the first definition of Aminobacter as legume-nodulating bacteria. *Syst Appl Microbiol*. 2012;35:65–72.
- Nonnoi F, Chinnaswamy A, de la García VS, Coba, de la Peña T, Lucas MM, Pueyo JJ. Metal tolerance of rhizobial strains isolated from nodules of herbaceous legumes (Medicago spp. and Trifolium spp.) growing in mercury-contaminated soils. *Applied Soil Ecology*. 2012;61:49–59.
- Roszbach S, Mai DJ, Carter EL, Sauviac L, Capela D, Bruand C, et al. Response of Sinorhizobium meliloti to elevated concentrations of Cadmium and Zinc. *Appl Environ Microbiol*. 2008;74:4218–21.
- Maynaud G, Brunel B, Mornico D, Durot M, Severac D, Dubois E, et al. Genome-wide transcriptional responses of two metal-tolerant symbiotic Mesorhizobium isolates to Zinc and Cadmium exposure. *BMC Genomics*. 2013;14:292.
- Lu M, Jiao S, Gao E, Song X, Li Z, Hao X et al. Transcriptome response to Heavy metals in Sinorhizobium meliloti CCNWSX0020 reveals New Metal Resistance determinants that also promote bioremediation by Medicago lupulina in metal-contaminated soil. *Appl Environ Microbiol*. 2017;83.
- Arregui G, Hipólito P, Pallol B, Lara-Dampier V, García-Rodríguez D, Varela HP, et al. Mercury-Tolerant Ensifer medicae strains Display High Mercuric reductase activity and a Protective Effect on Nitrogen fixation in Medicago truncatula nodules under Mercury stress. *Front Plant Sci*. 2021;11:560768.
- García Gomez M, Diego Caballero Klink J, Boffetta P, Espanol S, Sallsten G, Gomez Quintana J. Exposure to mercury in the mine of Almaden. *Occup Environ Med*. 2006;64:389–95.
- Jach ME, Sajnaga E, Ziaja M. Utilization of Legume-Nodule Bacterial Symbiosis in Phytoremediation of Heavy Metal-contaminated soils. *Biology*. 2022;11:676.

35. Tan Y-F, O'Toole N, Taylor NL, Millar AH. Divalent Metal Ions in Plant Mitochondria and their role in interactions with proteins and oxidative stress-Induced damage to respiratory function. *Plant Physiol.* 2010;152:747–61.
36. Jozefczak M, Remans T, Vangronsveld J, Cuypers A. Glutathione is a key player in Metal-Induced oxidative stress defenses. *IJMS.* 2012;13:3145–75.
37. Sharma P, Jha AB, Dubey RS, Pessaralki M. Reactive oxygen species, oxidative damage, and Antioxidative Defense Mechanism in Plants under stressful conditions. *J Bot.* 2012;2012:1–26.
38. Alvarez-Rivera G, Sanz A, Cifuentes A, Ibáñez E, Paape T, Lucas MM, et al. Flavonoid Accumulation varies in *Medicago truncatula* in response to Mercury stress. *Front Plant Sci.* 2022;13:933209.
39. Villar I, Larrainzar E, Milazzo L, Pérez-Rontomé C, Rubio MC, Smulevich G, et al. A Plant Gene Encoding one-heme and two-heme hemoglobins with Extreme Reactivities toward Diatomic gases and Nitrite. *Front Plant Sci.* 2020;11:600336.
40. Bobik C, Meilhoc E, Batut J. FixJ: A Major Regulator of the Oxygen Limitation Response and late symbiotic functions of *Sinorhizobium meliloti*. *J Bacteriol.* 2006;188:4890–902.
41. Ruiz-Díez B, Fajardo S, Del Rosario De Felipe M, Fernández-Pascual M. Characterization of rhizobia from legumes of agronomic interest grown in semi-arid areas of Central Spain relates genetic differences to soil properties. *J Basic Microbiol.* 2012;52:66–78.
42. Chin C-S, Alexander DH, Marks P, Klammer AA, Drake J, Heiner C, et al. Nonhybrid, finished microbial genome assemblies from long-read SMRT sequencing data. *Nat Methods.* 2013;10:563–9.
43. Nelson M, Guhlin J, Epstein I, Tiffin P, Sadowsky MJ. The complete replicons of 16 *Ensifer meliloti* strains offer insights into intra- and inter-replicon gene transfer, transposon-associated loci, and repeat elements. *Microb Genomics.* 2018;4.
44. Hunt M, Silva ND, Otto TD, Parkhill J, Keane JA, Harris SR. Circlator: automated circularization of genome assemblies using long sequencing reads. *Genome Biol.* 2015;16:294.
45. Brettin T, Davis JJ, Disz T, Edwards RA, Gerdes S, Olsen GJ, et al. RASTtk: a modular and extensible implementation of the RAST algorithm for building custom annotation pipelines and annotating batches of genomes. *Sci Rep.* 2015;5:8365.
46. Tatusova T, DiCuccio M, Badretdin A, Chetvernin V, Nawrocki EP, Zaslavsky L, et al. NCBI prokaryotic genome annotation pipeline. *Nucleic Acids Res.* 2016;44:6614–24.
47. Emms DM, Kelly S. OrthoFinder: phylogenetic orthology inference for comparative genomics. *Genome Biol.* 2019;20:238.
48. Minkin I, Patel A, Kolmogorov M, Vyahhi N, Pham S, Sibelia. A scalable and comprehensive synteny block generation tool for closely related microbial genomes. 2013. <https://doi.org/10.48550/ARXIV.1307.7941>.
49. Katoh K, Standley DM. MAFFT multiple sequence alignment Software Version 7: improvements in performance and usability. *Mol Biol Evol.* 2013;30:772–80.
50. Price MN, Dehal PS, Arkin AP. FastTree 2 – approximately maximum-likelihood trees for large alignments. *PLoS ONE.* 2010;5:e9490.
51. Mirdita M, Schütze K, Moriawaki Y, Heo L, Ovchinnikov S, Steinegger M. ColabFold: making protein folding accessible to all. *Nat Methods.* 2022;19:679–82.
52. Emsley P, Cowtan K. Coot: model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr.* 2004;60:2126–32.
53. Schrödinger LLC. The PyMOL Molecular Graphics System, version 1.8. 2015.
54. Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, et al. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics.* 2013;29:15–21.
55. Andrews S, FastQC: A Quality Control Tool for High Throughput Sequence Data. 2010.
56. Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics.* 2014;30:2114–20.
57. Love MI, Huber W, Anders S. Moderated estimation of Fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* 2014;15:550.
58. Gruber-Vodicka HR, Seah BKB, Pruesse E. phyloFlash: Rapid Small-Subunit rRNA profiling and targeted Assembly from Metagenomes. *mSystems.* 2020;5:e00920–20.
59. Robinson MD, Oshlack A. A scaling normalization method for differential expression analysis of RNA-seq data. *Genome Biol.* 2010;11:R25.
60. Robinson MD, McCarthy DJ, Smyth GK. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics.* 2010;26:139–40.
61. Raudvere U, Kolberg L, Kuzmin I, Arak T, Adler P, Peterson H, et al. G:profiler: a web server for functional enrichment analysis and conversions of gene lists (2019 update). *Nucleic Acids Res.* 2019;47:W191–8.
62. Garg B, Dogra RC, Sharma PK. High-Efficiency Transformation of *Rhizobium leguminosarum* by Electroporation. *Appl Environ Microbiol.* 1999;65:2802–4.
63. Geddes BA, Kearsley JVS, Huang J, Zamani M, Muhammed Z, Sather L, et al. Minimal gene set from *Sinorhizobium (Ensifer) meliloti* pSymA required for efficient symbiosis with *Medicago*. *Proc Natl Acad Sci USA.* 2021;118:e2018015118.
64. Ghosh P, Adolphsen KN, Yurgel SN, Kahn ML. *Sinorhizobium medicae* WSM419 genes that improve symbiosis between *Sinorhizobium meliloti* Rm1021 and *Medicago truncatula* Jemalong A17 and in other Symbiosis systems. *Appl Environ Microbiol.* 2021;87.
65. Young JPW, Crossman LC, Johnston AW, Thomson NR, Ghazoui ZF, Hull KH, et al. The genome of *Rhizobium leguminosarum* has recognizable core and accessory components. *Genome Biol.* 2006;7:R34.
66. Ledwidge R, Patel B, Dong A, Fiedler D, Falkowski M, Zelikova J, et al. NmerA, the Metal Binding Domain of Mercuric Ion Reductase, removes hg²⁺ from proteins, delivers it to the Catalytic Core, and protects cells under glutathione-depleted conditions. *Biochemistry.* 2005;44:11402–16.
67. Baek Y, Kim J, Ahn J, Jo I, Hong S, Ryu S, et al. Structure and function of the hypochlorous acid-induced flavoprotein RclA from *Escherichia coli*. *J Biol Chem.* 2020;295:3202–12.
68. Schiering N, Kabsch W, Moore MJ, Distefano MD, Walsh CT, Pai EF. Structure of the detoxification catalyst mercuric ion reductase from *Bacillus* sp. strain RC607. *Nature.* 1991;352:168–72.
69. Bafana A, Khan F, Suguna K. Structural and functional characterization of mercuric reductase from *Lysinibacillus sphaericus* strain G1. *Biometals.* 2017;30:809–19.
70. Argyrou A, Blanchard JS. Flavoprotein disulfide reductases: advances in Chemistry and function. *Progress in Nucleic Acid Research and Molecular Biology.* Elsevier; 2004. pp. 89–142.
71. Shearer HL, Loi VV, Weiland P, Bange G, Altegoer F, Hampton MB, et al. MerA functions as a hypothiocyanoic acid reductase and defense mechanism in *Staphylococcus aureus*. *Mol Microbiol.* 2023;119:456–70.
72. Lafrance-Vanasse J, Lefebvre M, Di Lello P, Sygusch J, Omichinski JG. Crystal structures of the Organomercurial Lyase MerB in its free and Mercury-bound forms. *J Biol Chem.* 2009;284:938–44.
73. Prabhakaran P, Ashraf MA, Aqma WS. Microbial stress response to heavy metals in the environment. *RSC Adv.* 2016;6:109862–77.
74. Sharma R, Bhat A, Xie M, Pueyo JJ, Paape T. Above and belowground responses to heavy metals show complex tissue specific patterns of gene expression responses to cadmium or mercury in *Medicago truncatula*. unpublished
75. Naguib MM, El-Gendy AO, Khairalla AS. Microbial diversity of mer Operon genes and their potential roles in Mercury Bioremediation and Resistance. *Open Biotechnol J.* 2018;12:56–77.
76. Msaddak A, Mars M, Quiñones MA, Lucas MM, Pueyo JJ. Lupin, a Unique Legume that is nodulated by multiple microsymbionts: the role of horizontal gene transfer. *IJMS.* 2023;24:6496.
77. Barrett R, Schluter D. Adaptation from standing genetic variation. *Trends Ecol Evol.* 2008;23:38–44.
78. Woods LC, Gorrell RJ, Taylor F, Connallon T, Kwok T, McDonald MJ. Horizontal gene transfer potentiates adaptation by reducing selective constraints on the spread of genetic variation. *Proc Natl Acad Sci USA.* 2020;117:26868–75.
79. Sugawara M, Epstein B, Badgley BD, Unno T, Xu L, Reese J, et al. Comparative genomics of the core and accessory genomes of 48 *Sinorhizobium* strains comprising five genospecies. *Genome Biol.* 2013;14:R17.
80. Morby AP, Hobman JL, Brown NL. The role of cysteine residues in the transport of mercuric ions by the *tn501* MerT and MerP mercury-resistance proteins. *Mol Microbiol.* 1995;17:25–35.
81. Hamlett NV, Landale EC, Davis BH, Summers AO. Roles of the Tn21 merT, merP, and merC gene products in mercury resistance and mercury binding. *J Bacteriol.* 1992;174:6377–85.
82. Narayanan M, Ma Y. Metal tolerance mechanisms in plants and microbe-mediated bioremediation. *Environ Res.* 2023;222:115413.
83. Paape T, Heiniger B, Santo Domingo M, Clear MR, Lucas MM, Pueyo JJ. Genome-Wide Association Study Reveals Complex Genetic Architecture of Cadmium and Mercury Accumulation and Tolerance traits in *Medicago truncatula*. *Front Plant Sci.* 2022;12:806949.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.