Supplementary Information for the manuscript:

A plasmid-chromosome crosstalk in multidrug resistant enterobacteria

Laura Toribio-Celestino[1¶], Alicia Calvo-Villamañán[1¶], Cristina Herencias[2,3], Aida Alonso-del Valle[1], Jorge Sastre-Dominguez[1], Susana Quesada[1], Didier Mazel[4], Eduardo PC Rocha[5], Ariadna Fernández-Calvet[1*], Alvaro San Millan[1,6*]

[1] Centro Nacional de Biotecnología (CNB-CSIC), Madrid, Spain

[2] Servicio de Microbiología, Hospital Universitario Ramón y Cajal and Instituto Ramón y Cajal de Investigación Sanitaria, Madrid, Spain

[3] Centro de Investigación Biológica en Red de Enfermedades Infecciosas, Instituto de Salud Carlos III, Madrid, Spain

[4] Institut Pasteur, Université de Paris Cité, CNRS UMR3525, Bacterial Genome Plasticity, Paris, France

[5] Institut Pasteur, Université de Paris Cité, CNRS UMR3525, Microbial Evolutionary Genomics, Paris, France

[6] Centro de Investigación Biológica en Red de Epidemiología y Salud Pública, Instituto de Salud Carlos III, Madrid, Spain

*Corresponding authors

E-mail: asanmillan@cnb.csic.es (ASM)

E-mail: ariadna.fernandez@cnb.csic.es (AFC)

¶These authors contributed equally to this work.

Supplementary Results and Discussion

Expression of other resident plasmid's genes

We analyzed the differential expression (DE) of genes of other resident plasmids (excluding pOXA-48) in the 11 MDR strains. To find groups of genes with similar functions that tend to be DE in the same direction, we performed gene set enrichment analysis (GSEA) on the sets of plasmid genes. Transposases or insertion sequences (ISs) were enriched in three *Klebsiella* spp. strains (Supplementary Data 5). In J57, transposition processes tended to be upregulated with pOXA-48 carriage, while in KPN07 and KPN10, downregulated. The IS families involved in these patterns were diverse in the three strains (Supplementary Data 5). Interestingly, we previously observed a similar pattern with the pOXA-48-encoded IS*1*, where IS*1* was downregulated in strains encoding multiple copies of IS*1*, like in KPN07 and KPN10 due to high PCN, and upregulated in strains encoding few copies of IS*1*, like J57, which does not encode IS*1* in the chromosome or other plasmids[1]. This suggests that the IS*1* of pOXA-48 could be influencing the expression of other IS families. Additionally, processes involved in response to antibiotics were upregulated in KPN15 (Supplementary Data 5).

Analysis of the conservation of $LysR_{pOXA-48}$

If overexpressing the operon is important for pOXA-48-carrying strains, $LysR_{pOXA-48}$ could be under selective pressure to maintain its regulatory function, and thus its protein sequence should be conserved. To study this, we analyzed the protein sequences of $LysR_{pOXA-48}$ in the RefSeq database of Proteobacteria, using the $LysR_{pOXA-48}$ sequence of pOXA-48_K8—the most frequent variant isolated in our hospital—as query. We filtered BLASTp searches by 90% identity and full alignment length (303 amino acids) to discard matches with other LysR regulators (Supplementary Data 12). This way, we identified $LysR_{pOXA-48}$ only in Gammaproteobacteria (288 strains). Importantly, there were no hits in the range 90-96% of identity, indicating high sequence conservation. Furthermore, when lowering the alignment length threshold to >250 amino acids, only five additional hits were obtained, supporting the high sequence conservation.

We used ABRicate v1.0.1 (https://github.com/tseemann/abricate) with the PlasmidFinder database to determine whether LysRs with high identity with $LysR_{pOXA-48}$ were encoded on plasmids or chromosomes. We performed the analyses separating hits in two groups: (i) LysRs with 100% identity

with $LysR_{pOXA-48}$, and (ii) LysRs with <100% (and >96%) identity with $LysR_{pOXA-48}$. Most $LysR_{pOXA-48}$ (71%) were encoded on plasmids, specially on pOXA-48-like plasmids (n = 172). Chromosomally-encoded $LysR_{pOXA-48}$ were mainly present in *Shewanella* spp., from where the bla_{OXA-48}-*lysR* genes were originally transposed to the IncL plasmid backbone. Most plasmid-encoded $LysR_{pOXA-48}$ (93%) shared 100% with the reference sequence, and 90.6% were encoded in pOXA-48-like plasmids (Supplementary Data 12). Then, only 13 $LysR_{pOXA-48}$ shared less than 100% identity with the reference, and 38% of these were encoded in other plasmid backbones (Supplementary Data 12).

To analyze amino acid substitutions in LysRs encoded on plasmids, we aligned the protein sequences of these LysRs with MAFFT v7.453. Alignments were visualized in ESPript v3.0[2]. We used the SIFT[3] and Mutpred2[4] webservers (May 2024) to assess the impact of amino acid substitutions on protein function. The 13 $LysR_{pOXA-48}$ with less than 100% identity with the reference shared the same amino acid substitution (S112G) compared to the reference $LysR_{pOXA-48}$ (Supplementary Fig. 21), and it was predicted to not affect protein function (SIFT score = 0.45, Mutpred2 score = 0.36), supporting the protein sequence of $LysR_{pOXA-48}$ is under high selective pressure.

Next, we calculated the ratio of non-synonymous to synonymous nucleotide substitutions (N/S) of $LysR_{pOXA-48}$ among the 172 pOXA-48-like plasmids. For this, we used the nucleotide sequence of pOXA-48_K8 (accession number MT441554), the most common pOXA-48 variant found in clinical isolates from the Hospital Universitario Ramón y Cajal[5], as a common reference to analyze the variants. We employed the NCBI PGAP pipeline v.2022-12-13.build6494 to annotate the reference sequence. We used Snippy v.4.6.0 (https://github.com/tseemann/snippy) to identify the variants from each of the plasmid sequences downloaded from RefSeq using the contigs as input. We merged the mutations by nucleotide position for each of the genes (i.e. different n plasmids with the same SNP were identified as 1 event). We then classified the identified variants by their effect as Non-synonymous ("missense_variant") or Synonymous ("synonymous"). Finally, we calculated the ratio of Non-synonymous over Synonymous SNPs (N/S) for each of the plasmid genes as a proxy for the strength of selection. We identified a total of 599 different variants (Supplementary Data 12). Only one of these variants affected $lysR_{pOXA-48}$ (Non-synonymous: Ser112Gly), which correspond to the eight pOXA-48-like plasmids encoding for the $LysR_{pOXA-48}$ with the S112G amino acid substitution

(Supplementary Data 12). Given the low number of mutations in *lysR*$_{pOXA-48}$, the N/S (1/0) could be misleading (Supplementary Data 12).

Thus, we further tested whether mutations were equally distributed along the plasmid genes by performing a permutation test (n permutations per gene = 10,000). First, we corrected the number of mutations per gene by the gene length. Then, we used as test statistic the absolute difference between the number of mutations per base for each gene (Ng) and the mean Ng for the rest of the genes of the plasmid (Ng-Np). In each permutation, an Ng was randomly sampled and compared with the mean of the rest of the genes. The *p*-value would be the probability of a random value being more extreme than the observed value for each gene. We found four genes significantly enriched ($P < 0.05$) in the number of mutations per base (*traR*, a helix-turn-helix transcriptional regulator, and two hypothetical proteins, Supplementary Data 12), but we did not detect any significant differences for *lysR*$_{pOXA-48}$ ($P = 0.13$, Supplementary Data 12). However, when analyzing the number of mutations per base for all pOXA-48 genes, *lysR*$_{pOXA-48}$ was the second least targeted gene, after *traY* (Supplementary Data 12).

In addition to the high sequence conservation of *lysR*$_{pOXA-48}$, knocked-out versions of *lysR*$_{pOXA-48}$ have rarely been identified—only two of seven versions of the Tn*1999* transposon (Tn*1999.4* and Tn*1999.5*, Supplementary Fig. 22) encode a *lysR*$_{pOXA-48}$ disrupted by a transposon insertion, and these Tn*1999* versions are rare compared to the widely disseminated Tn*1999.2*, invTn*1999.2* and Tn*1999.1* versions[5–9]. Moreover, *lysR*$_{pOXA-48}$ was the only complete gene transposed along with *bla*$_{OXA-48}$ from the chromosome of *Shewanella* spp. to pOXA-48.

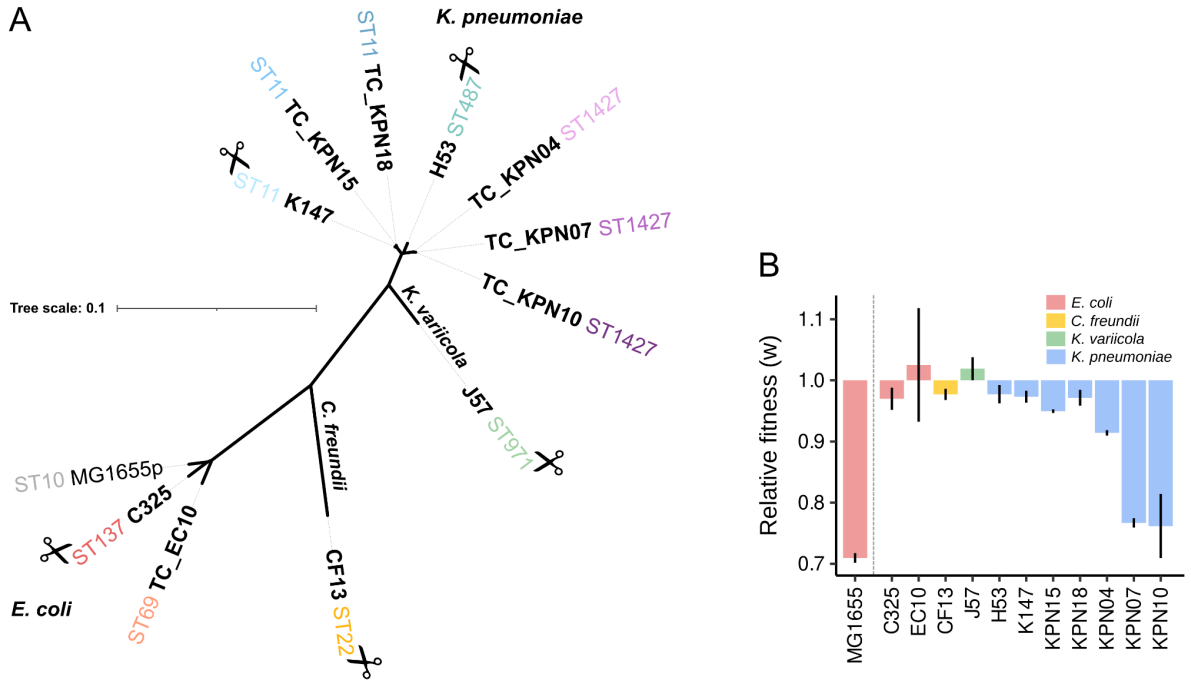
Taken together, these observations suggest that the function of LysR$_{pOXA-48}$ is important for pOXA-48 carriers and thus is under selective pressure to retain its regulatory function, possibly over the *pfp-ifp* operon genes.

References

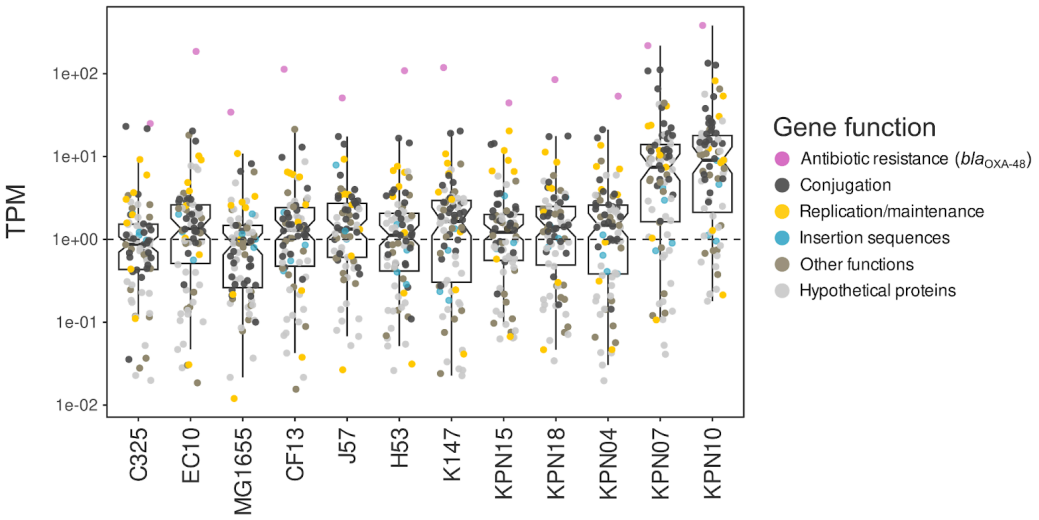
1. Sastre-Dominguez, J. *et al.* Plasmid-encoded insertion sequences promote rapid adaptation in clinical enterobacteria. *Nat. Ecol. Evol.* 1–16 (2024).

2. Robert, X. & Gouet, P. Deciphering key features in protein structures with the new ENDscript server. *Nucleic Acids Res*. **42**: W320–W324 (2014).
3. Ng, P.C. & Henikoff, S. SIFT: predicting amino acid changes that affect protein function. *Nucleic Acids Res*. **31**: 3812–3814 (2003).
4. Pejaver, V. *et al.* Inferring the molecular and phenotypic impact of amino acid variants with MutPred2. *Nat. Commun*. **11**: 5918 (2020).
5. León-Sampedro, R. *et al.* Pervasive transmission of a carbapenem resistance plasmid in the gut microbiota of hospitalised patients. *Nat. Microbiol*. **6**, 606–616 (2021).
6. Pitout, J. D. D., Peirano, G., Kock, M. M., Strydom, K.-A. & Matsumura, Y. The global ascendency of OXA-48-type carbapenemases. *Clin. Microbiol*. Rev. **33**, e00102-19 (2019).
7. Sattler J. *et al*. Emergence of Tn*1999.7*, a new transposon in *bla*$_{OXA-48}$-harboring plasmids associated with increased plasmid stability. *Antimicrob. Agents. Chemother.* **66**:e00787-22 (2022).
8. Potron, A., Poirel, L., Rondinaud, E. & Nordmann, P. Intercontinental spread of OXA-48 beta-lactamase-producing Enterobacteriaceae over a 11-year period, 2001 to 2011. *Euro Surveill*. **18**(31):20549 (2013).
9. Lerminiaux, N. *et al.* Plasmid genomic epidemiology of carbapenem-hydrolysing class D β-lactamase (CDHL)-producing Enterobacterales in Canada, 2010−2021. *Microb. Genom*. **10**(6):001257 (2024).
10. Alonso-del Valle, A. *et al.* Variability of plasmid fitness effects contributes to plasmid persistence in bacterial communities. *Nat. Commun*. **12**, 2653 (2021).
11. Fernández-Calvet, A. *et al*. The distribution of fitness effects of plasmid pOXA-48 in clinical enterobacteria. *Microbiology* **169**, 001369 (2023).
12. Rodríguez-Beltrán, J. *et al.* Translational demand is not a major source of plasmid-associated fitness costs. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* **377**, 20200463 (2022).
13. Ishikawa, S. A., Zhukova, A., Iwasaki, W. & Gascuel, O. A fast likelihood method to reconstruct and visualize ancestral scenarios. *Mol. Biol. Evol.* **36**, 2069–2085 (2019).

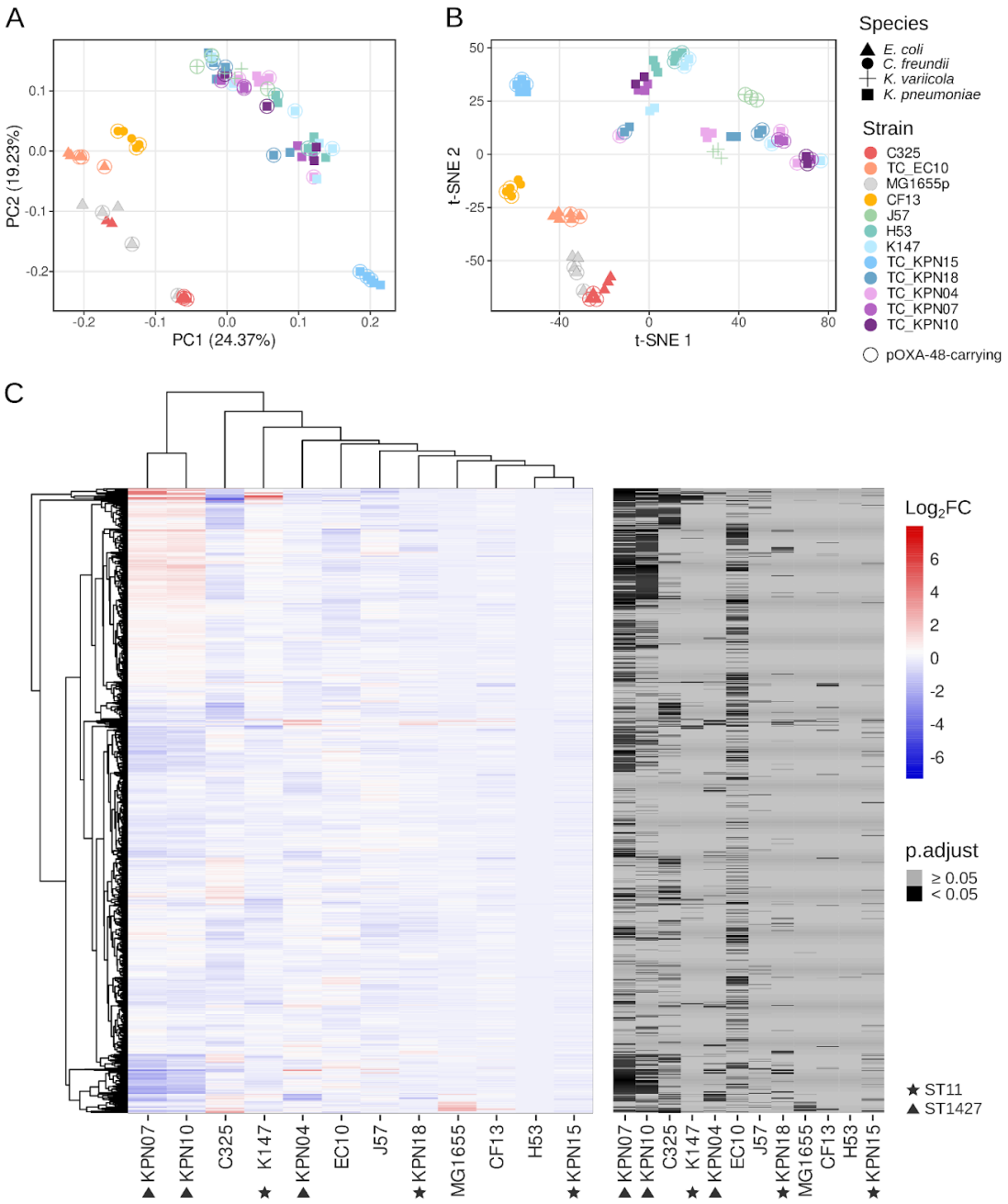
Supplementary Figures



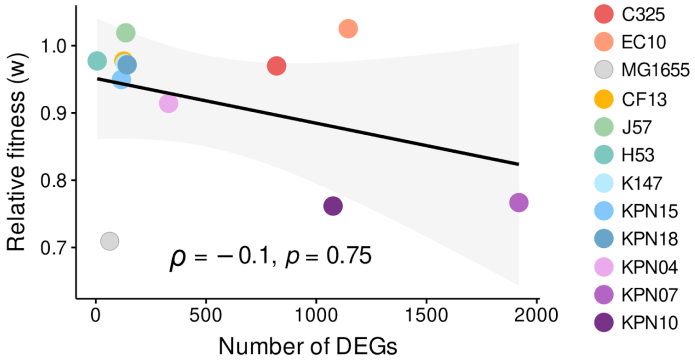
Supplementary Fig. 1. Enterobacterial strains. (A) Unrooted phylogeny of mash distances between the 11 clinically relevant MDR enterobacterial strains, marked in bold, and the laboratory *E. coli* strain MG1655. Trees were constructed with the closed genomic sequences of the pOXA-48-carrying strains. Strains that were cured from pOXA-48 using CRISPR-Cas9 technology are indicated with scissor icons; strains for which pOXA-48 transconjugants were obtained are denoted with the prefix TC. Sequence Types (STs) are indicated. (B) Distribution of fitness effects of pOXA-48-carrying strains. Relative fitness (*w*) values were previously obtained from competition experiments between pOXA-48-carrying and their respective pOXA-48-free strains[10–12]. Values >1 indicate that pOXA-48 carriage is associated with a fitness advantage, while values <1 indicate a reduction in fitness associated with pOXA-48. Error bars indicate the standard error of the mean.



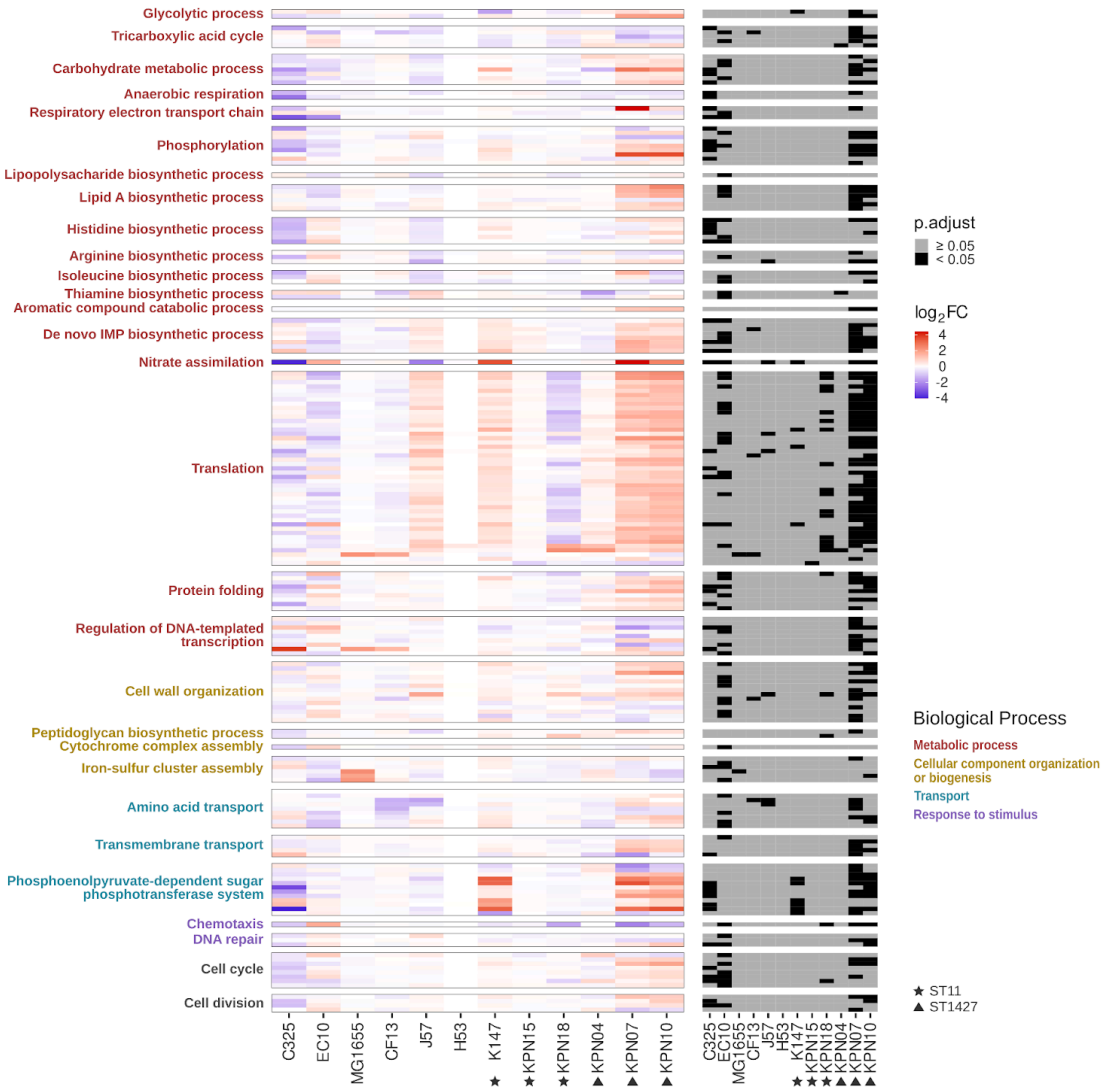
Supplementary Fig. 2. Expression of pOXA-48 genes. Boxplots of the expression of pOXA-48 genes of each strain, measured as the median TPM (Transcripts Per Million) between replicates, normalized by median chromosomal TPM values of the corresponding strain (horizontal dashed line). The *y* scale is $\log_{10}$-transformed for visualization. TPM > 1 indicates that pOXA-48 genes are expressed more than the median expression of chromosomal genes, while TPM < 1 indicates that pOXA-48 gene expression is below the median chromosomal TPM values. Median TPM values are indicated as horizontal lines inside the boxes and the notches represent the 95% confidence interval for the median. The upper and lower box hinges correspond to the 25th and 75th percentiles and the whiskers extend to observations within 1.5x the interquartile range.



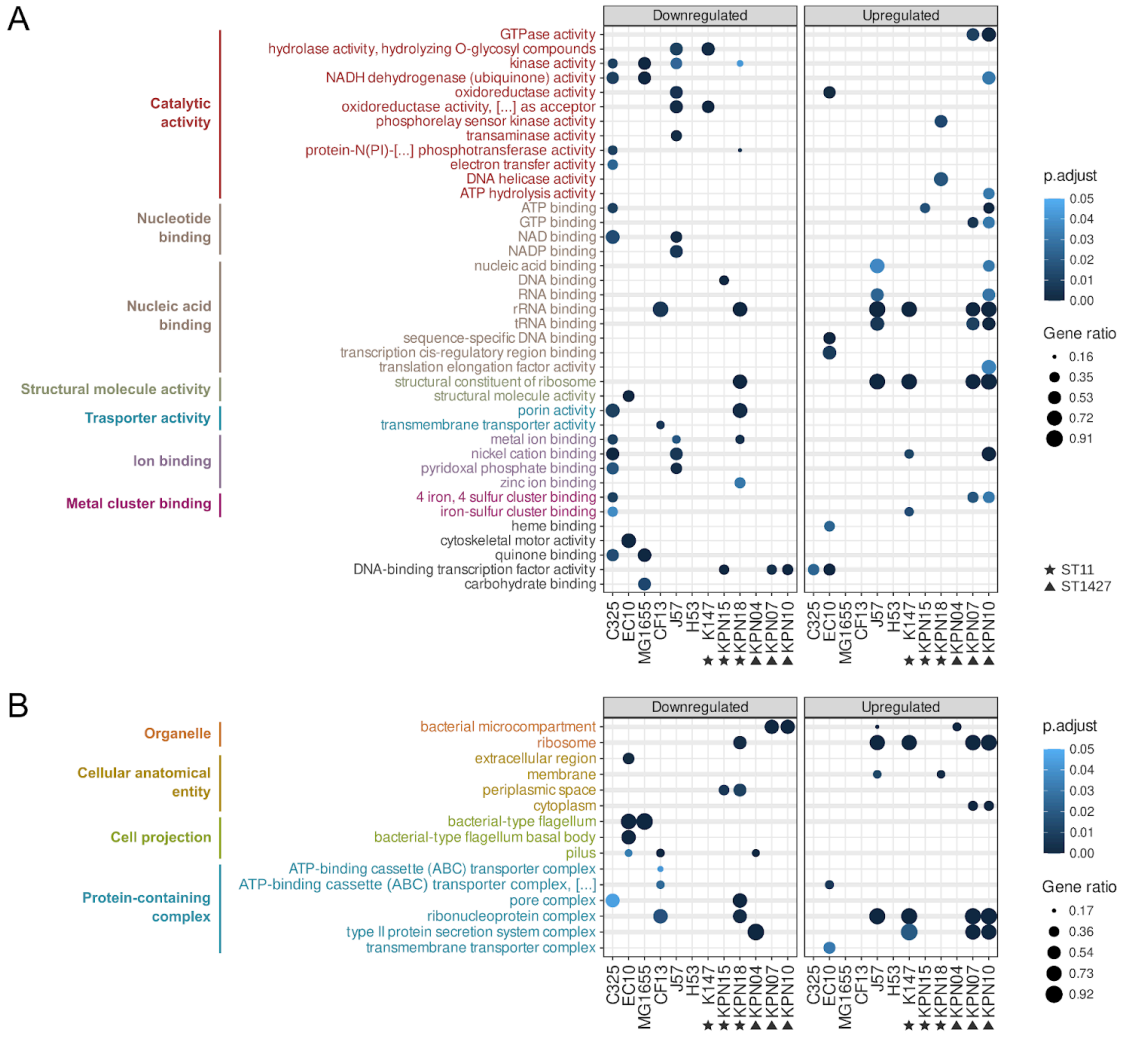
Supplementary Fig. 3. Strain-specific expression of 2,488 single copy orthologs (SCO). (A) Principal component analysis and (B) t-distributed stochastic neighbor embedding of pOXA-48-free and -carrying (indicated with circles) enterobacterial replicates. Both panels represent gene expression of 2,488 single copy orthologs (see Methods). Shapes and colors indicate the species and the strain of each sample, respectively. (C) The heatmap on the left panel represents the differential expression (log_2FC) of 2,488 SCO calculated by comparing pOXA-48-carrying strains to their pOXA-48-free counterparts. The dendrogram represents the hierarchical clustering of strains (complete linkage method). Black tiles in the heatmap of the right panel indicate genes whose differential expression is significant (adjusted *p* value < 0.05). Strains belonging to the same Sequence Type (ST) are indicated with shapes.



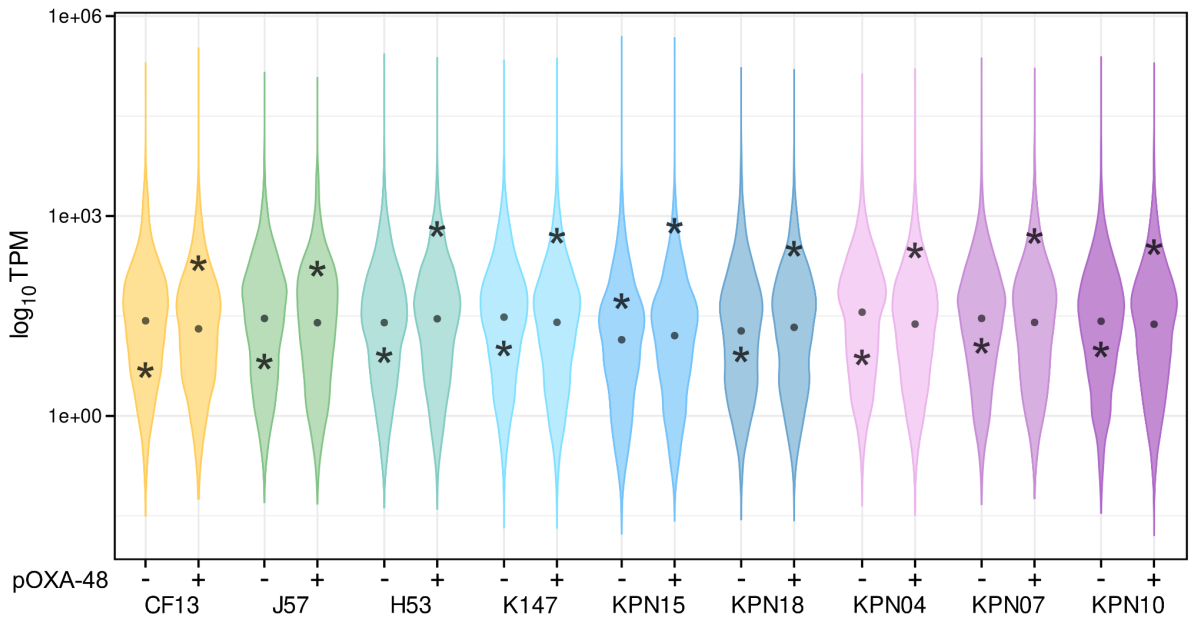
Supplementary Fig. 4. Spearman's rank correlation between the number of DEGs and the relative fitness (*w*) of pOXA-48-carrying enterobacteria. Relative fitness (*w*) values were previously obtained from competition experiments between pOXA-48-carrying and their respective pOXA-48-free strains [38–40]. Both variables did not follow a normal distribution (Shapiro test, $P < 0.05$), so Spearman's rank correlation (ρ = -0.1, P = 0.75) was used. Gray shading indicates the 95% confidence interval of the regression line (R^2 = 0.14, $F(1,10)$ = 1.59, P = 0.24; number of DEGs $\beta = -6.65 \times 10^{-5}$, P = 0.24). Strains are differentiated by colors.



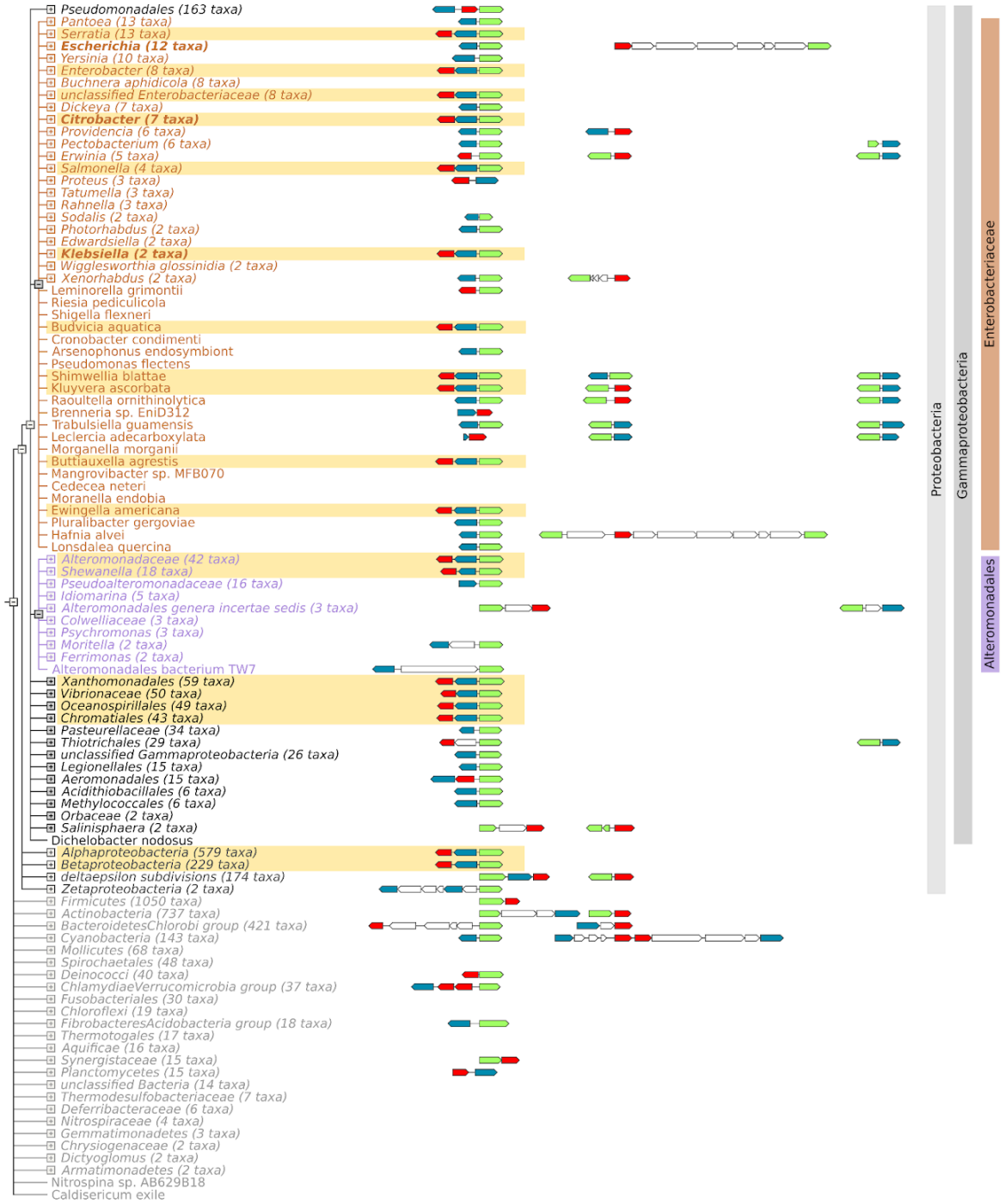
Supplementary Fig. 5. Differential expression of genes from enriched Biological Processes. The heatmap on the left panel represents the differential expression (log_2FC, calculated by comparing pOXA-48-carrying strains to their pOXA-48-free counterparts) of genes annotated with Biological Processes that were enriched in the GSEA. Genes were matched across strains by RefSeq ID, Gene and Product annotations. Black tiles in the heatmap of the right panel indicate genes whose differential expression is significant (adjusted *p* value < 0.05). Strains belonging to the same Sequence Type (ST) are indicated with shapes.



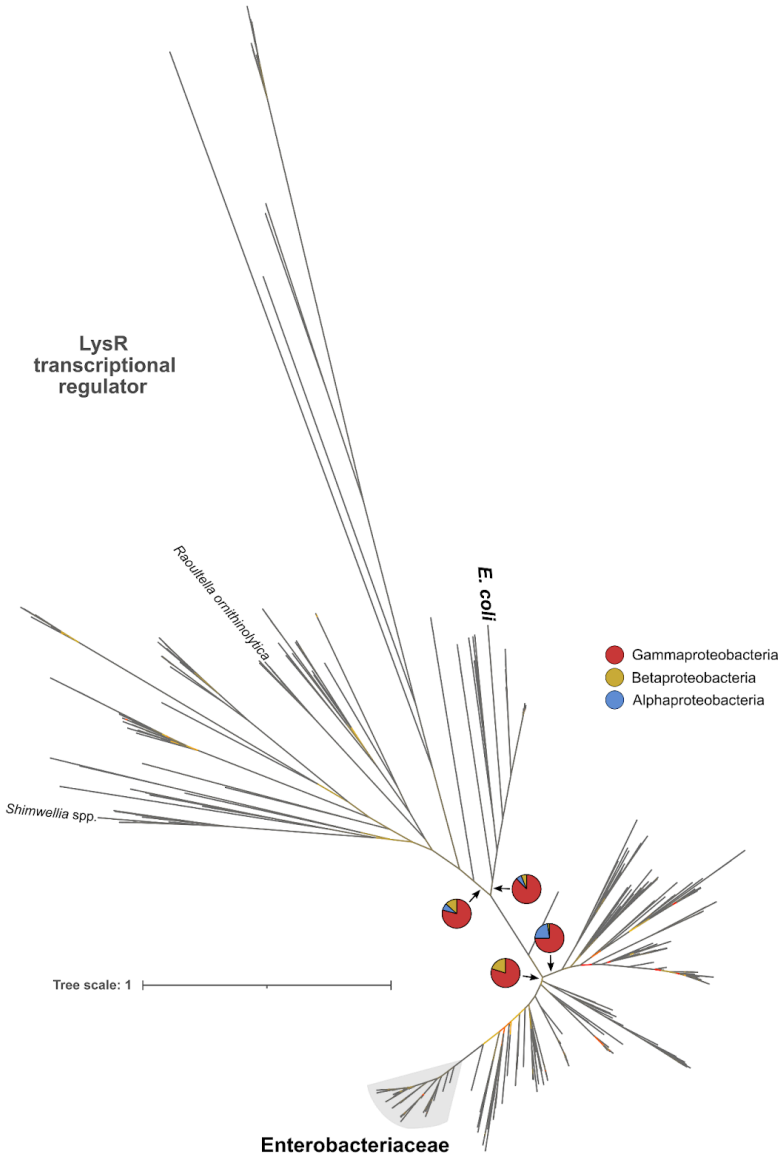
Supplementary Fig. 6. Additional Gene Set Enrichment Analysis (GSEA) results. Molecular Functions (MF) (A) and Cellular Components (CC) (B) enriched in pOXA-48-carrying enterobacteria. Gene Set Enrichment Analysis (GSEA) was performed on the lists of raw DEGs annotated with Gene Ontology terms for MF and CC (see Methods). Downregulated and upregulated enriched MF and CC are separated in two panels, and are represented by circles. Thicker horizontal lines indicate terms enriched in more than one strain. The size of the circles indicates the ratio of the number of enriched genes of a specific MF or CC by the number of total genes annotated with that term. The adjusted *p* value for each enriched term and strain is represented in a color gradient. Strains belonging to the same Sequence Type (ST) are indicated with shapes.



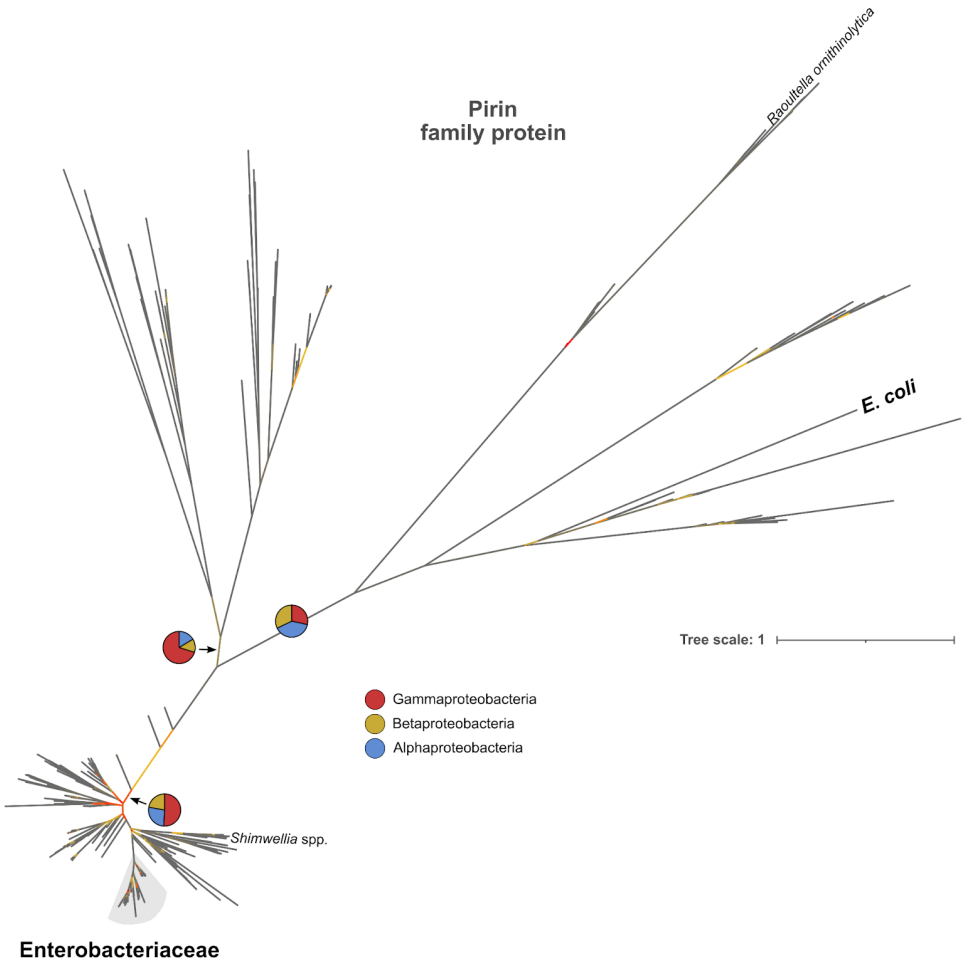
Supplementary Fig. 7. Expression of the *pfp-ifp* operon genes. Violin plots show the distribution of the expression of chromosomal genes in TPM (Transcripts Per Million) for each pOXA-48-free and -carrying strain pairs. Dots indicate the median chromosomal gene expression. Asterisks indicate the level of expression of the *pfp-ifp* operon. To aid visualization, the *y* axis is represented in $\log_{10}$ scale without value transformation.



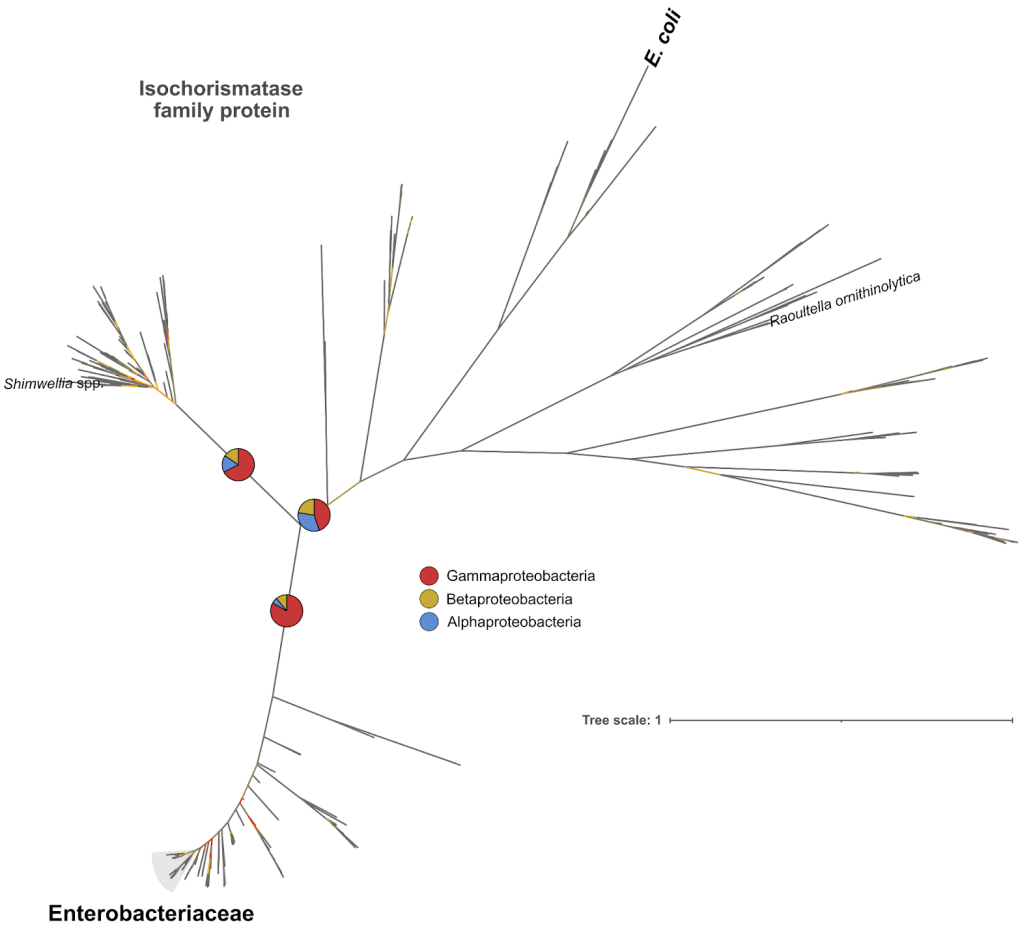
Supplementary Fig. 8. STRING conserved neighborhood view of the *lysR-pfp-ifp* cluster in bacteria. The *lysR*, *pfp* and *ifp* genes are represented as red, blue and green arrows, respectively. The representative neighborhood of the three genes is shown for each bacterial genera/species. Bacterial species or genera encoding the *lysR-pfp-ifp* cluster are highlighted in yellow. Species classification is indicated at the right. The dendrogram does not show phylogenetic distances.



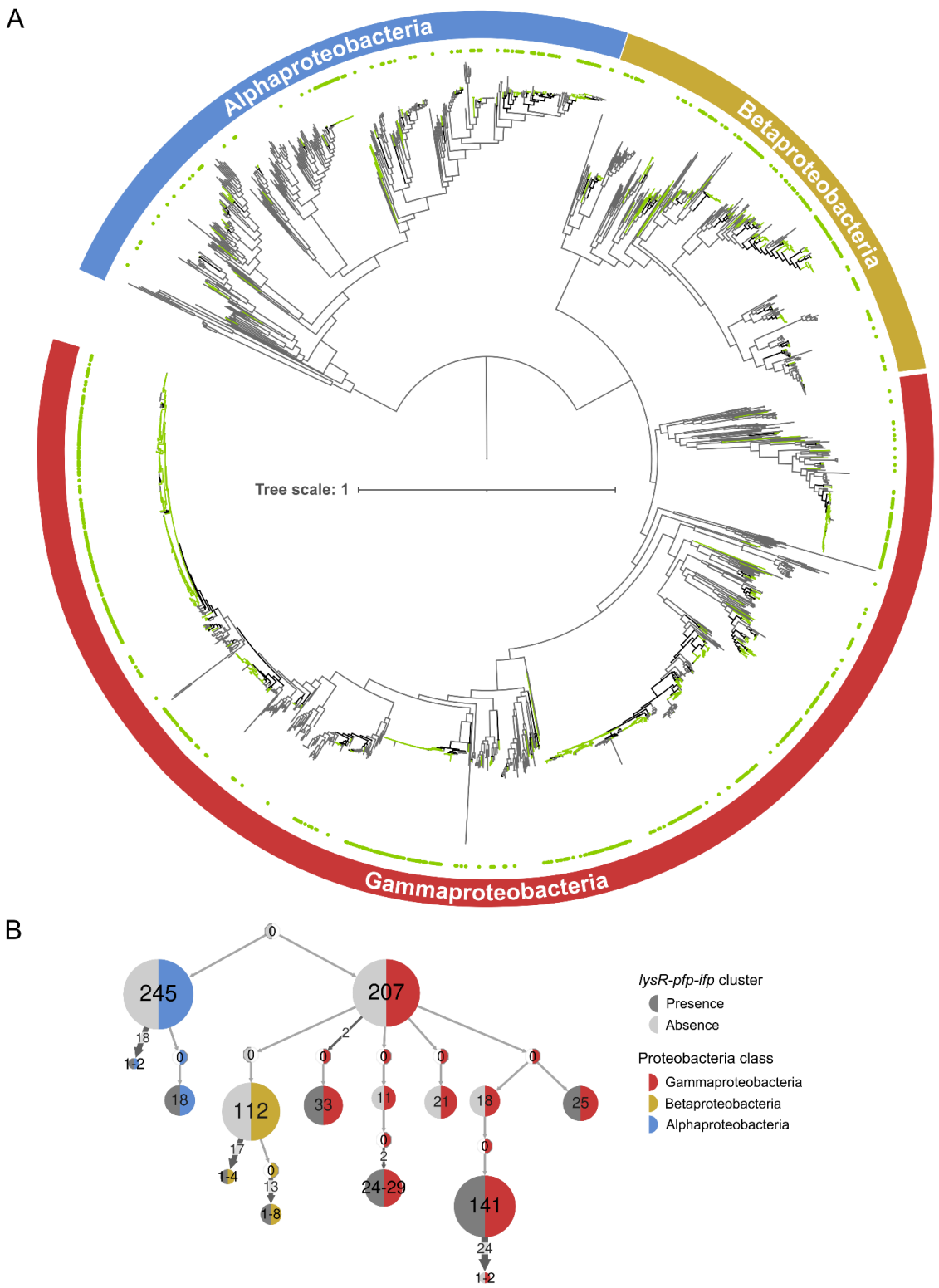
Supplementary Fig. 9. Phylogeny of the LysR transcriptional regulator. Unrooted phylogenetic tree of Proteobacteria (n = 698) constructed from the protein sequences of the LysR that putatively regulates the small chromosomal operon (see Methods). Pie charts show the proportion of Proteobacteria strains belonging to the Gamma- (red), Beta- (yellow) or Alphaproteobacteria (blue) classes included in each of the indicated internal branches. Enterobacteriaceae strains that branch from the same node are shaded in gray; other Enterobacteriaceae species not branching from that node are indicated. Bootstrap values are represented with a color gradient in the branches of the tree: low (aprox. 3), medium (aprox. 48.5) and high (aprox. 100) bootstrap values are colored in red, yellow and gray, respectively. Tree scale represents the average number of substitutions per site.



Supplementary Fig. 10. Phylogeny of the pirin family protein. Unrooted phylogenetic tree of Proteobacteria (n = 698) constructed from the protein sequences of the pirin family protein of the small chromosomal operon (see Methods). Pie charts show the proportion of Proteobacteria strains belonging to the Gamma- (red), Beta- (yellow) or Alphaproteobacteria (blue) classes included in each of the indicated internal branches. Enterobacteriaceae strains that branch from the same node are shaded in gray; other Enterobacteriaceae species not branching from that node are indicated. Bootstrap values are represented with a color gradient in the branches of the tree: low (aprox. 3), medium (aprox. 48.5) and high (aprox. 100) bootstrap values are colored in red, yellow and gray, respectively. Tree scale represents the average number of substitutions per site.



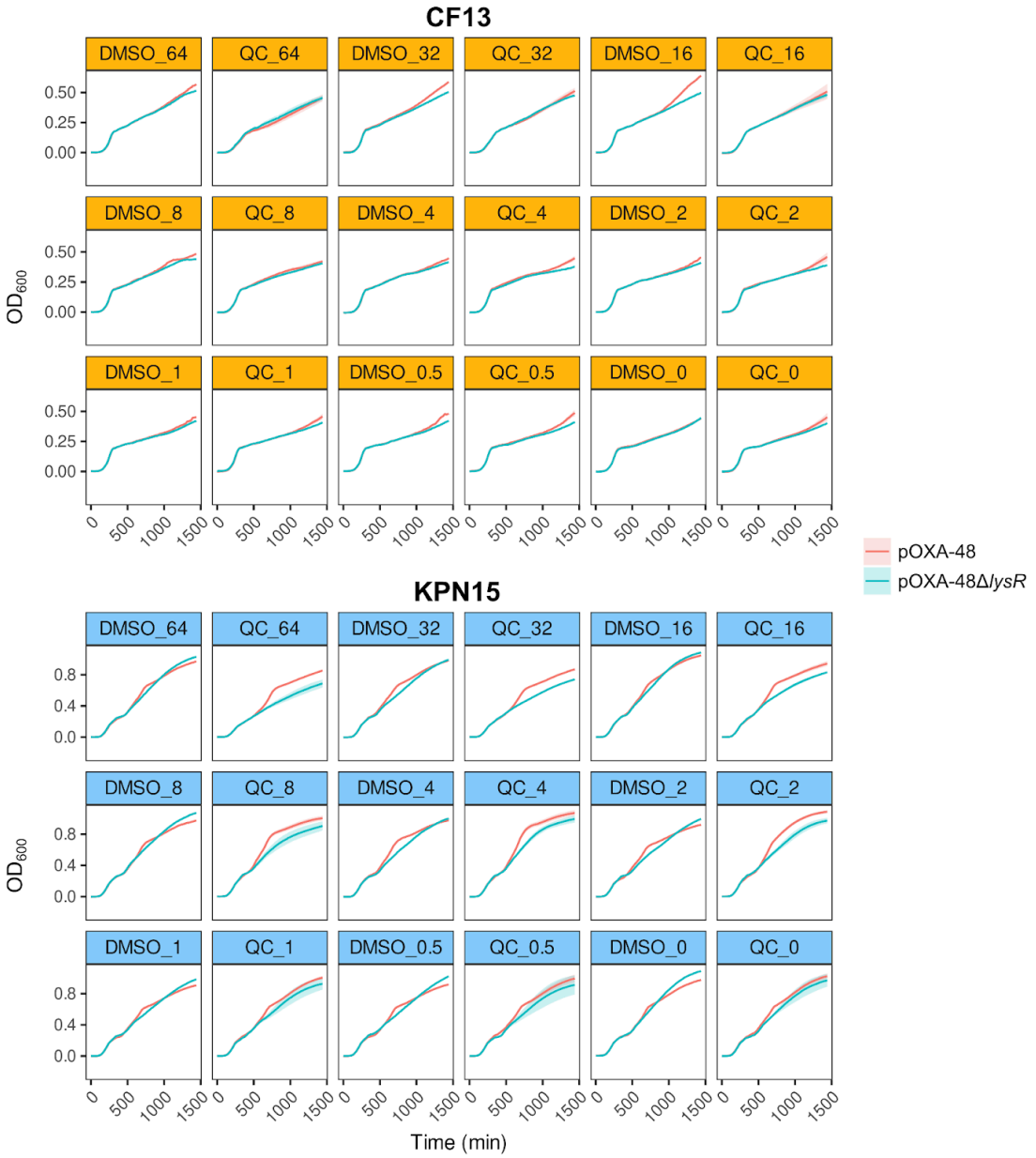
Supplementary Fig. 11. Phylogeny of the isochorismatase family protein. Unrooted phylogenetic tree of Proteobacteria (n = 698) constructed from the protein sequences of the isochorismatase family protein of the small chromosomal operon (see Methods). Pie charts show the proportion of Proteobacteria strains belonging to the Gamma- (red), Beta- (yellow) or Alphaproteobacteria (blue) classes included in each of the indicated internal branches. Enterobacteriaceae strains that branch from the same node are shaded in gray; other Enterobacteriaceae species not branching from that node are indicated. Bootstrap values are represented with a color gradient in the branches of the tree: low (aprox. 1), medium (aprox. 49.5) and high (aprox. 100) bootstrap values are colored in red, yellow and gray, respectively. Tree scale represents the average number of substitutions per site.



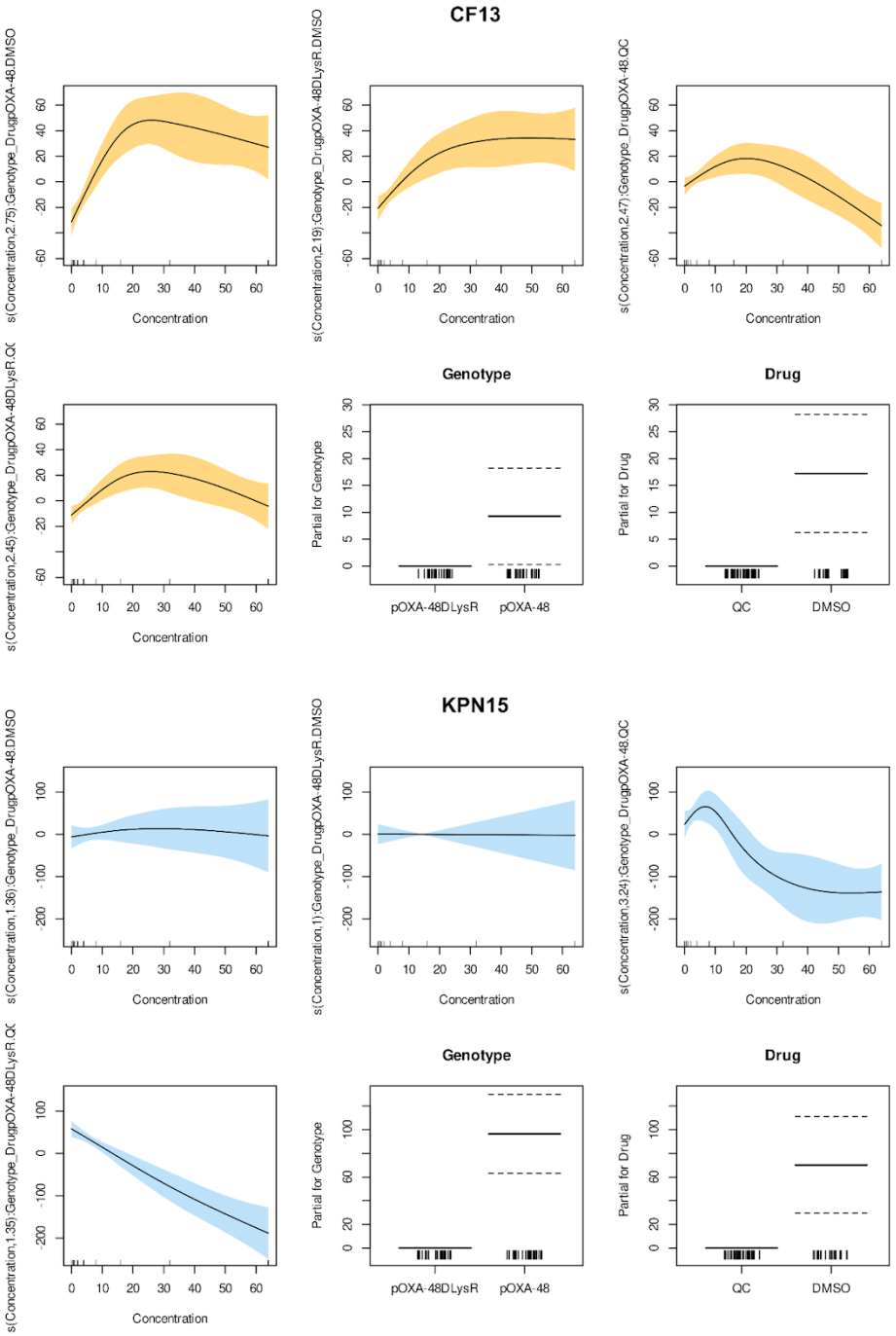
Supplementary Fig. 12. Ancestral character reconstruction of the *lysR-pfp-ifp* cluster. (A) Phylogenetic tree of 1,686 strains of Gamma-, Beta- and Alphaproteobacteria showing the ancestral character reconstruction (ACR) of the *lysR-pfp-ifp* cluster, rooted at the Epsilonproteobacteria outgroup (removed from the tree). The tree was constructed from the concatenated protein sequences of 128 conserved bacterial single-copy genes (see

Methods). Green dots at the tips of the tree and green nodes/branches indicate presence of the *lysR-pfp-ifp* cluster. Gray and black nodes/branches indicate absence or an undefined state (after ACR) of the *lysR-pfp-ifp* cluster, respectively. See Supplementary Fig. 20 (unpruned tree) for bootstrap values. Tree scale represents the average number of substitutions per site. (B) Compressed ACR of the *lysR-pfp-ifp* cluster in Proteobacteria (see Methods). The Proteobacteria phylogenetic tree is compressed vertically and horizontally based on the Proteobacteria class (right halves of circles) and the presence/absence of the *lysR-pfp-ifp* cluster (dark and light gray, respectively; left halves of circles). White circle halves indicate that the state of the respective character could not be inferred, which occurs in the internal nodes of the tree, indicated with 0s. Numbers inside circles indicate the number of tree tips that possess the corresponding characters after vertical compression. Numbers in arrows indicate the number of identical subtrees horizontally merged. For a visual explanation of the interpretability of the ACR compression, visit the PastML help page (https://pastml.pasteur.fr/help)[13].

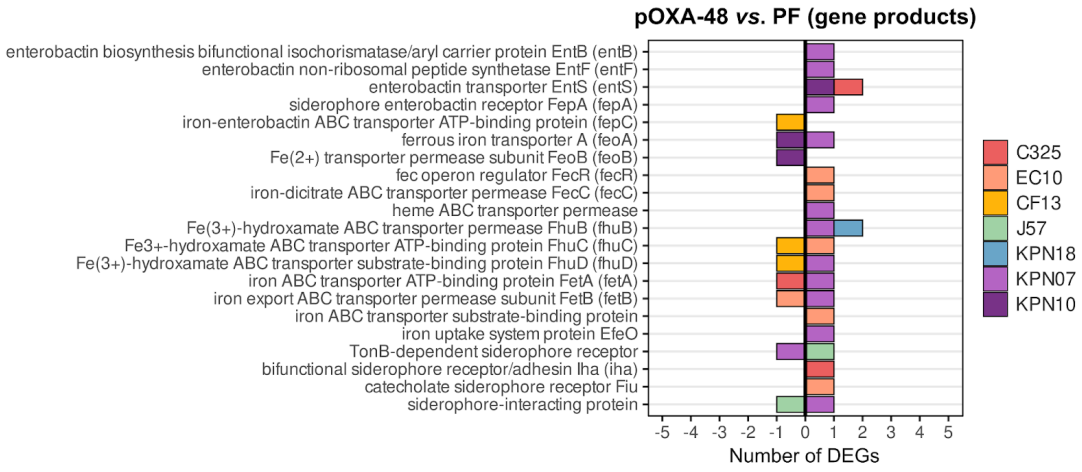
Supplementary Fig. 13. Genomic neighborhood of the *lysR-pfp-ifp* cluster. A representative subset of Proteobacteria strains was selected (see Methods). Genes are color-coded as indicated at the bottom. Homologous genes between strains and their percentage of sequence identity are indicated with a gray shading.



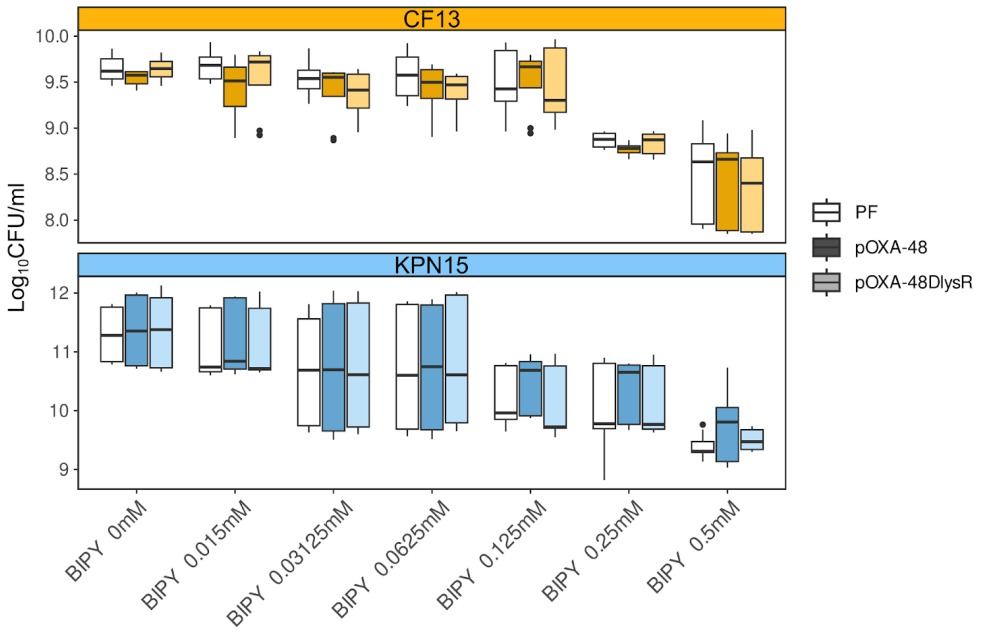
Supplementary Fig. 14. Growth curves of MDR enterobacteria grown under different concentrations of quercetin (QC). OD_{600} measured every 10 minutes for 24 hours for strains CF13 and KPN15 carrying pOXA-48 or pOXA-48Δ*lysR*. Strains were grown under concentrations ranging from 64 to 0.5 µg/mL of DMSO (as control, 1 replicate per concentration and strain) or DMSO+QC (2 replicates per concentration and strain). Shading indicates the standard error of the mean. Source data are provided as a Source Data file.



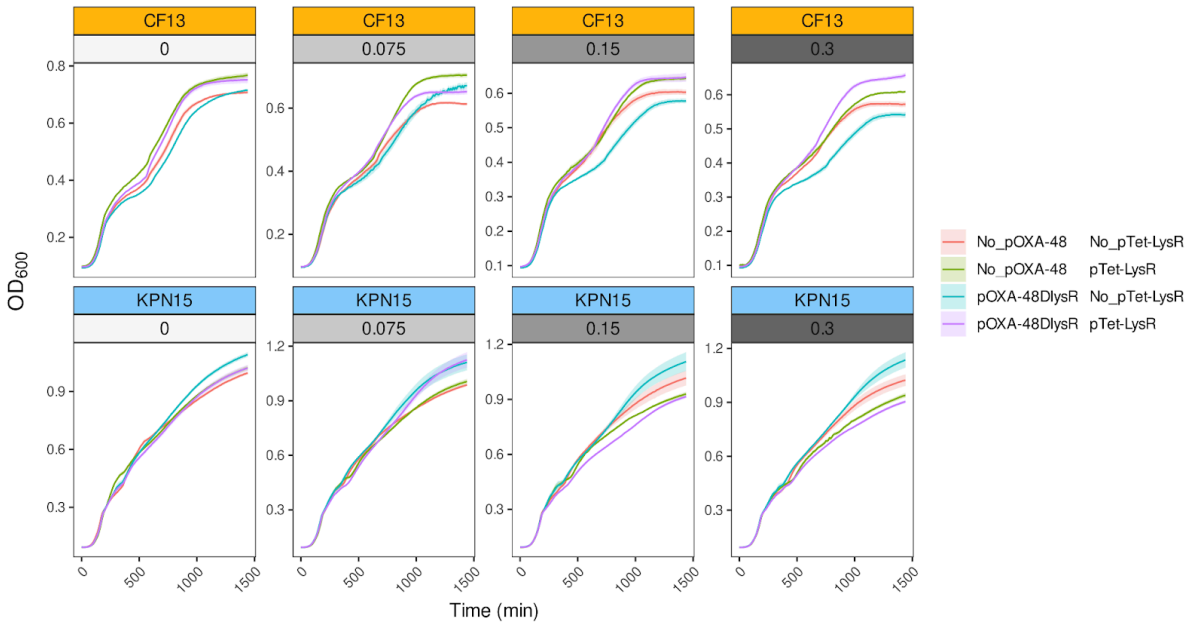
Supplementary Fig. 15. Generalized Additive Model (GAM) results. The GAMs are fitted to model bacterial growth (AUC of growth curves, Supplementary Fig. 14) as a function of the concentration of the drugs (DMSO and DMSO+QC), for CF13 (top) and KPN15 (bottom) separately. Four smooth functions are fitted for each genotype (pOXA-48 or pOXA-48Δ*lysR*) and drug (DMSO and DMSO+QC) interaction combinations, with the shaded area indicating the 95% confidence interval. Interaction effects of genotype and drug are incorporated into the model. Genotype pOXA-48Δ*lysR* and drug DMSO+QC are set as reference levels. Rug plots along the x-axis indicate the observed data points. Models fitted using the restricted maximum likelihood (REML) method (see Methods).



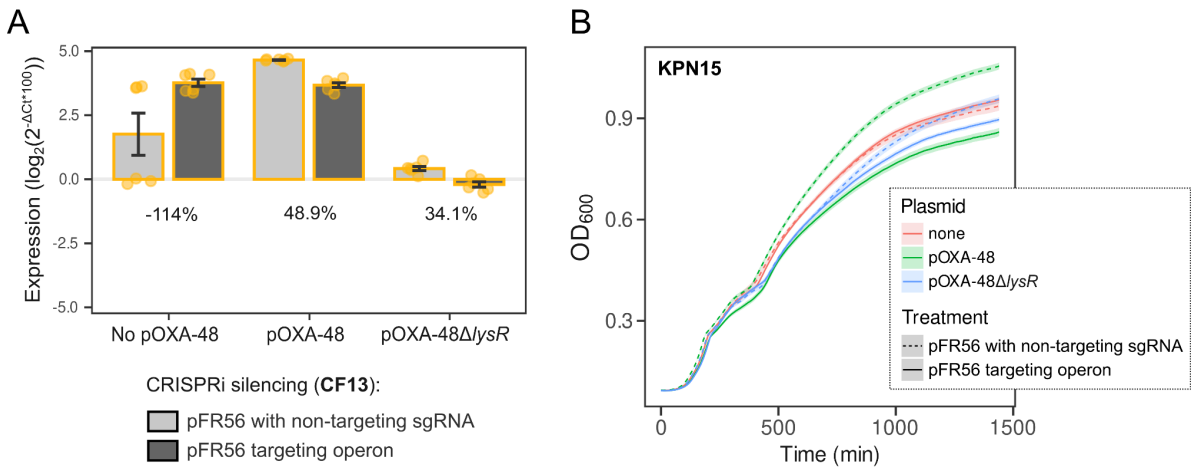
Supplementary Fig. 16. Number of significant DEGs involved in iron transport. DE was analyzed by comparing pOXA-48-carrying against pOXA-48-free strains (first RNA-Seq dataset). Values >0 represent the count of upregulated genes, and values <0 are downregulated genes.



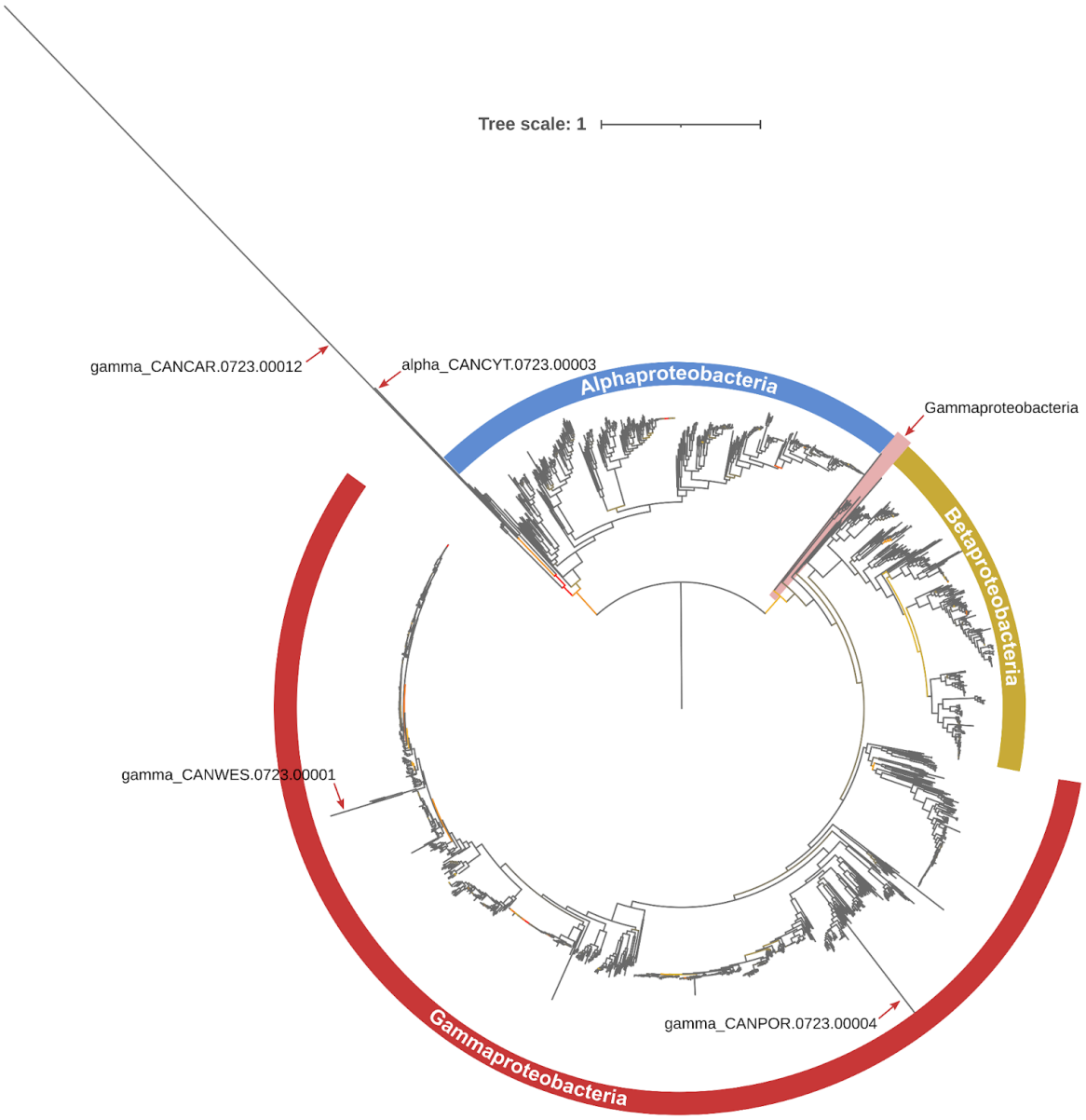
Supplementary Fig. 17. Growth of MDR enterobacteria under increasing concentrations of the iron chelator bipyridine. Colony forming units (CFU) per mL of strains CF13 (top figure) and KPN15 (bottom figure) not carrying or carrying pOXA-48 or pOXA-48Δ*lysR* grown under increasing concentrations of the iron chelator bipyridine (n = 3 replicates per condition and strain). Median CFU/mL values are indicated as horizontal lines inside the boxes, the upper and lower box hinges correspond to the 25th and 75th percentiles and the whiskers extend to observations within 1.5x the interquartile range. Source data are provided as a Source Data file.



Supplementary Fig. 18. Growth curves of strains overexpressing $LysR_{pOXA-48}$ at different concentrations of inducer. Effect of the overexpression of $LysR_{pOXA-48}$ on the growth of CF13 and KPN15. Curves performed for plasmid-free and pOXA-48Δ*lysR*-carrying CF13 and KPN15, with expression of $LysR_{pOXA-48}$ in *trans* from the pTet-LysR. Experiment performed with increasing concentrations of inducer (aTc). Shading indicates the standard error of the mean (n = 5 replicates of each strain/induced pairing). Source data are provided as a Source Data file.

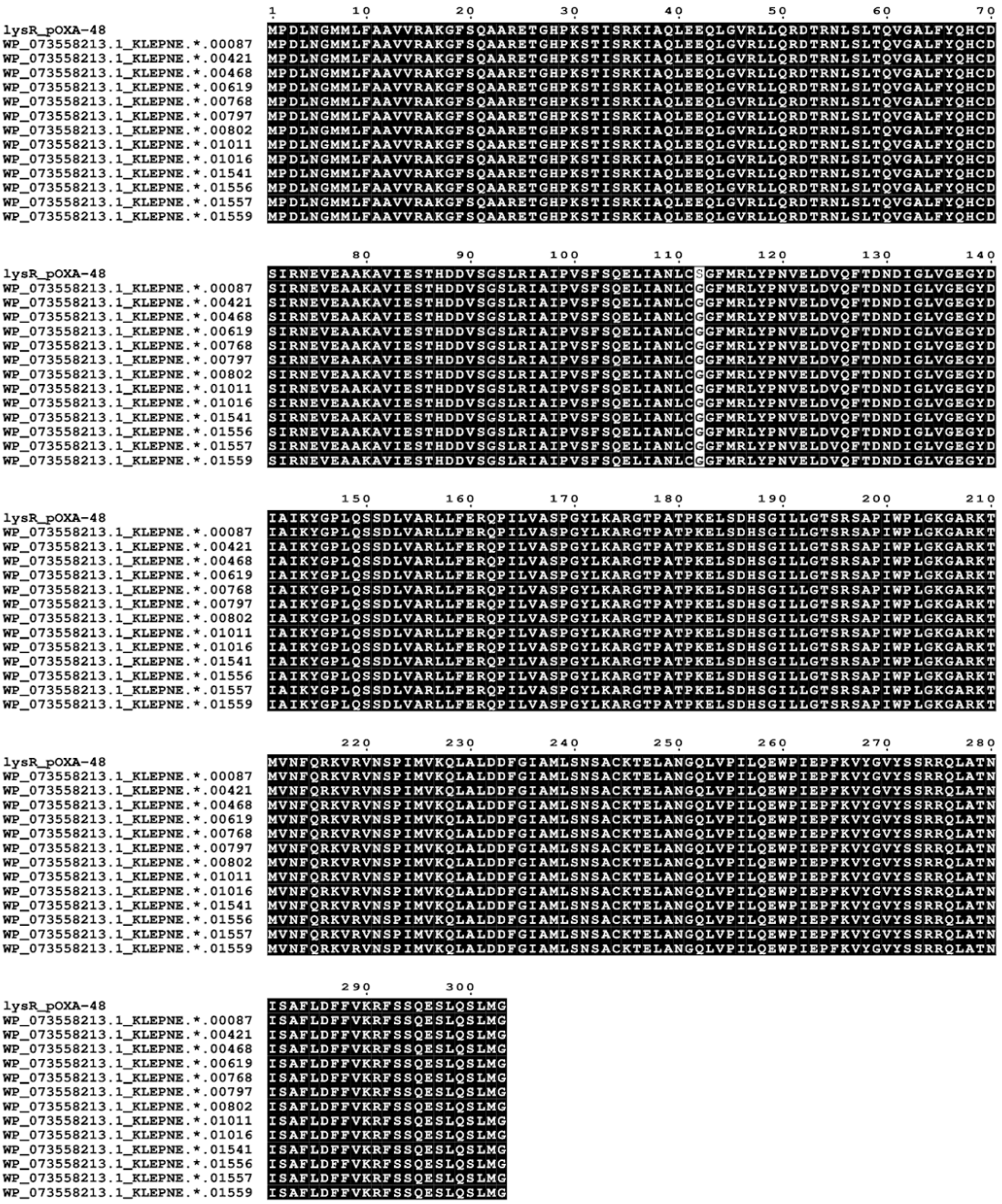


Supplementary Fig. 19. Repression of the *pfp-ifp* operon using CRISPRi. (A) RT-qPCR results for the level of repression of the *pfp-ifp* operon using CRISPRi in CF13. Levels of expression of the operon are calculated using RT-qPCR results for *pfp* as proxy for the whole operon and are normalized to the endogenous control *rpoB*. Percentages indicate the percentage of repression of the operon comparing the expression levels of cells carrying the non-targeting version of pFR56apm and cells carrying pFR56apm targeting the operon (n = 2 biological replicates, each as the median of 3 technical replicates). (B) Growth curves (optical density at 600 nm wavelength) of strain KPN15 with CRISPRi silencing of the *pfp-ifp* operon, used to calculate AUC values for Fig. 7C. OD_{600} was measured every 10 minutes during 24 hours for 16 replicates per group. Error bars (A) and shading (B) indicate the standard error of the mean. Source data are provided as a Source Data file.

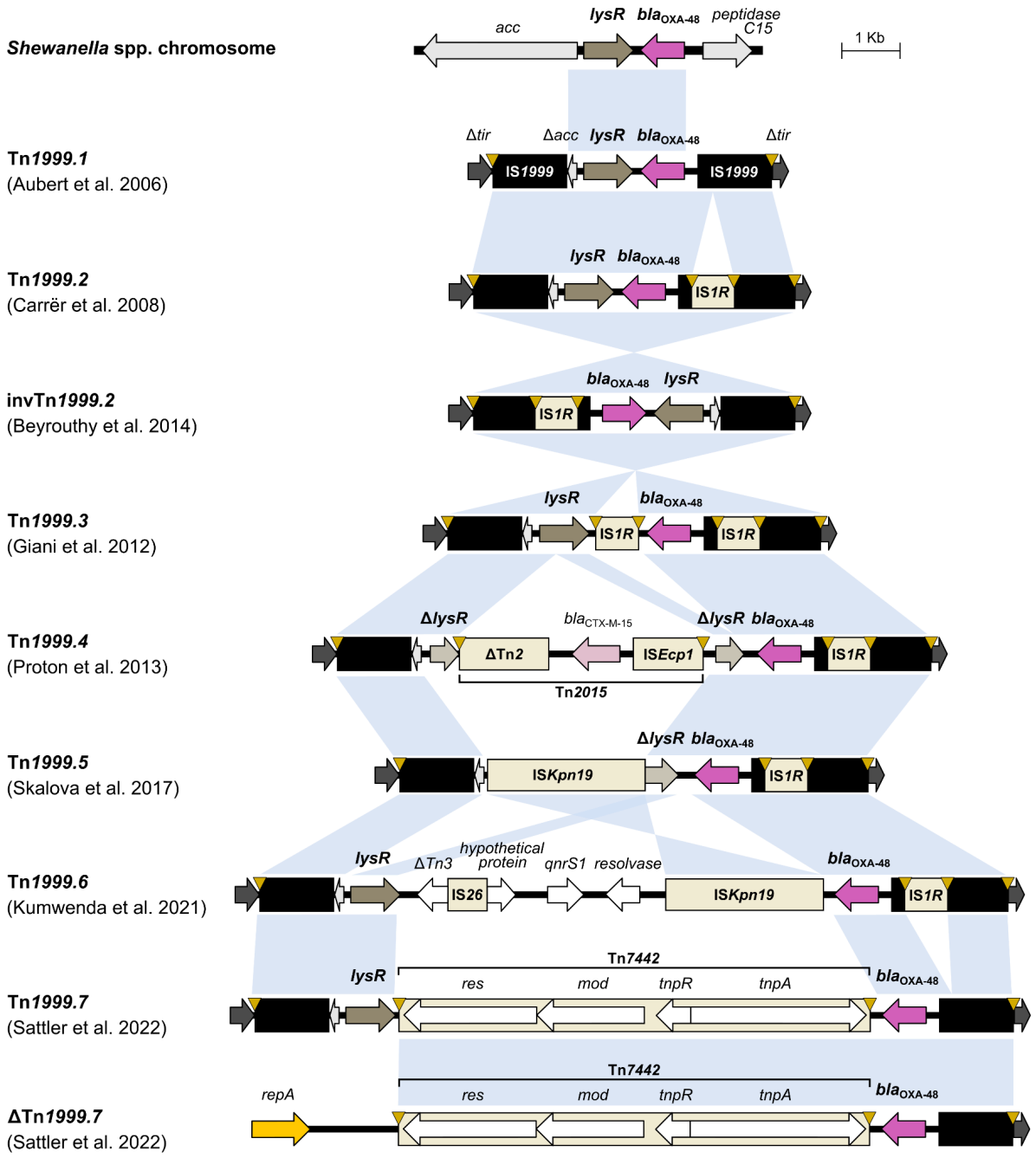


Supplementary Fig. 20. Unpruned phylogenetic tree of Proteobacteria. Phylogenetic tree of 1,704 strains of Gamma-, Beta- and Alphaproteobacteria, rooted at the Epsilonproteobacteria outgroup (removed from the tree). The tree was constructed from the concatenated protein sequences of 128 conserved bacterial single-copy genes (see Methods). The four outlier taxa identified by TreeShrink (Candidatus *Cytomitobacter primus* [GCF_951805245.1], Candidatus *Carsonella ruddii* [GCF_016593055.1], Candidatus *Westeberhardia cardiocondylae* [GCF_001242845.1], Candidatus *Portiera aleyrodidarum* BT-B-HRs [GCF_000300075.1]) and the clade of Gammaproteobacteria branching outside of the Gammaproteobacteria clade (*Sulfurifustis variabilis* [GCF_002355415.1], Candidatus *Ruthia magnifica* str. Cm [GCF_000015105.1], *Thiosulfatimonas sediminis* [GCF_011398355.1], *Hydrogenovibrio crunogenus* [GCF_004786015.1], *Ignatzschineria rhizosphaerae* [GCF_022655595.1], *Francisella frigiditurris* [GCF_001880225.1], *Francisella* sp. Scap27 [GCF_013394105.1],

Francisella adeliensis [GCF_012224465.1], *Francisella halioticida* [GCF_002211785.1], *Francisella persica* ATCC VR-331 [GCF_001653955.1], *Francisella opportunistica* [GCF_003347095.1], *Francisella philomiragia* [GCF_000833315.1], *Francisella noatunensis* subsp. *noatunensis* FSC774 [GCF_014844275.1], Candidatus *Comchoanobacter bicostacola* [GCF_024172705.1]) are indicated. These were pruned from the tree prior to ACR (Supplementary Fig. 12). Bootstrap values are represented with a color gradient in the branches of the tree: low (aprox. 8), medium (aprox. 46) and high (aprox. 100) bootstrap values are colored in red, yellow and gray, respectively. Tree scale represents the average number of substitutions per site.



Supplementary Fig. 21. MAFFT alignment of plasmid-encoded LysR transcriptional regulators. The protein sequences of 13 plasmid-encoded LysRs that share less than 100% (and >96%) identity with $LysR_{pOXA-48}$ are aligned to $LysR_{pOXA-48}$ (top of alignment). Matching and mismatching amino acids are highlighted in black and white, respectively. Of the 13 LysRs, eight are encoded in IncL/Ms (pOXA-48 plasmid variants): WP_073558213.1 in KLEPNE.0723.00087, WP_073558213.1 in KLEPNE.0723.00421, WP_073558213.1 in KLEPNE.0723.00468, WP_073558213.1 in KLEPNE.0723.00619, WP_073558213.1 in KLEPNE.0723.00768, WP_073558213.1 in KLEPNE.0723.01011, WP_073558213.1 in KLEPNE.0723.01016 and WP_073558213.1 in KLEPNE.0723.01556.



Supplementary Fig. 22. Versions of the Tn*1999* transposon in pOXA-48 plasmids. The chromosome of *Shewanella* spp. encodes the bla_{OXA-48}-*lysR* genes between the genes encoding for the peptidase C15 and the acetyl coenzyme A (*acc*). bla_{OXA-48}-*lysR* were captured by a Tn*1999* composite transposon (delimited by black boxes) and inserted into an IncL plasmid backbone. The different versions of Tn*1999* (from 1 to 7) within pOXA-48 plasmids are represented, including the reference where they were first described. Light yellow boxes represent insertion sequences and transposons; yellow triangles indicate target site duplications. Figure adapted from Pitout *et al.*[6] and Sattler *et al.*[7].