

Supplementary information

Microbial Ecology

Comparative genomics of the first plasmid identified in the genus *Halovibrio* revealing a mosaic structure shaped by horizontal gene transfer

Zuzana Lorková Kellnerová¹, Jana Kisková^{1*}, Arancha Peñil-Celis², M. Pilar Garcillán-Barcia², Zuzana Bártová³, Jozef Hančulák³, Peter Pristaš^{1,4}

¹ *Department of Microbiology, Institute of Biology and Ecology, Faculty of Science, Pavol Jozef Šafárik University in Košice, Šrobárova 2, 04154 Košice, Slovakia*

² *Instituto de Biomedicina y Biotecnología de Cantabria (Consejo Superior de Investigaciones Científicas – Universidad de Cantabria), Albert Einstein 22, 39011 Santander, Spain*

³ *Institute of Geotechnics, Slovak Academy of Sciences, Watsonova 45, 04001 Košice, Slovakia*

⁴ *Institute of Animal Physiology, Centre of Biosciences, Slovak Academy of Sciences, Šoltésovej 4-6, 04001 Košice, Slovakia*

* Corresponding author: Jana Kisková, email: jana.kiskova@upjs.sk

Content

Table S1 Bioinformatics tools used for the annotation of plasmid pHv20-59a

Table S2 Complete annotation of plasmid pHv20-59a

Table S3 List of genetic metal tolerance determinants identified in the pHv20-59a

Table S4 List of genetic metal tolerance determinants identified in the chromosome of the *Halovibrio* sp. HP20-59

Table S5 List of plasmids used in the comparative analyses based on BLAST results

Table S6 List of plasmids used in the comparative analyses downloaded from the IMG/PR database and associated information

Fig. S1 PLACNETw reconstruction of the *Halovibrio* sp. HP20-59 draft genome. The chromosome contigs are shown as the blue nodes and the size of nodes is proportional to the contig length. The orange nodes represent the reference genomes or plasmids. The yellow-green node represents contig with the replication initiation protein – RIP and the red node represents the contig with relaxase – REL. This reconstruction results in only one plasmid. The full lines represent the homology to the reference and the dashed line represents the scaffold link

Fig. S2 Circular visualization of plasmid pHv20-59b illustrating the distribution of coding sequences

Fig. S3 COG category distribution for pHv20-59a

Fig. S4 EasyFig comparative analysis illustrating shared genetic regions among pHv20-59a and genetically closest plasmids from halophilic microorganisms publicly available in the GenBank database (AP019515.1, NZ_CP130611.1, NZ_CP158485.1, NZ_CP011053.1, NZ_CP022738.1) and IMG plasmids (IMGPR_plasmid_2905650138_000002, IMGPR_plasmid_3300003108_000005,

IMGPR_plasmid_3300027834_000114, IMGPR_plasmid_3300031976_000086, IMGPR_plasmid_3300032083_000208). The alignments span the complete sequence of all compared plasmids, from coordinate 1 to the full plasmid length. Homologous regions are indicated by shaded gray blocks, with shading intensity reflecting the percentage of sequence identity (as shown in the scale bars). Gene annotations are displayed as orange arrows along each sequence. For simplicity, plasmids are labeled using only the last two or three digits of their original identifiers

Table S1 Bioinformatics tools used for the annotation of plasmid pHv20-59a

Bioinformatic tool	Version	Parameters	Reference
Bakta	1.12.0	database v6.0; default	Schwengers et al [1]
RAST		default	Aziz et al [2]; Overbeek et al [3]; Brettin et al [4]
BLAST	June 2026	BlastP; database: non-redundant protein sequences; min. query cover 90%, min. identity 90%; e-value cutoff 0.05	Altschul et al [5]
PGAP	02/22/2025	Best-placed reference protein set; GeneMarkS-2+	Li et al [6]
HHpred	57c8707149031cc9f8edceba362c71a3762bdbf8	databases: PDB_mmCIF70_20_Feb, Pfam-A_v38.2, SCOPe70_2.08, ECOD_F70_20230309; min. coverage of MSA hits 50%, min. identity of MSA hits with query 50%; e-value cutoff for MSA generation 1e-6	Zimmermann et al [7]
HMMER (HMMScan)	3.4	database: pfam 37.2; e-value cutoff 0.05	Finn et al [8]; Potter et al [9]
InterProScan	5.78-109.0	default	Blum et al [10]; Jones [11]
CDD	3.21	e-value cutoff 0.05	Wang et al [12]; Marchler-Bauer [13]
EggNOG-mapper	5.0.0	default	Huerta-Cepas et al [14]
CARD	4.0.1	default	Alcock et al [15]
ResFinder	4.7.2	default	Bortolaia [16];
ISEScan	1.7.3	default	Xie [17]
MobileElementFinder	1.0.3	default	Johansson [18]

Bakta - Rapid & standardized annotation of bacterial genomes, MAGs & plasmids server, RAST - Rapid Annotation using Subsystem Technology server, BLAST - Basic Local Alignment Search Tool, PGAP - Prokaryotic Genome Annotation Pipeline (NCBI), HHpred - server for protein remote homology detection and 3D structure prediction, HMMER - tool for fast and sensitive homology searching based on profile Hidden Markov models, InterProScan - software package that allows sequences to be scanned against InterPro's member database signatures, CDD – Conserved Domain Database (NCBI), EggNOG-mapper - tool for functional annotation of large sets of sequences based on fast orthology assignments, CARD - The Comprehensive Antibiotic Resistance Database, ResFinder – tool for identification of acquired antibiotic resistance genes, ISEScan – pipeline for identification of insertion sequence elements in prokaryotic genomes, MobileElementFinder - tool for identifying mobile genetic elements in whole genome shotgun sequence data

References

1. Schwengers O, Jelonek L, Dieckmann MA, Beyvers S, Blom J, Goesmann A (2021) Bakta: rapid and standardized annotation of bacterial genomes via alignment-free sequence identification. *Microb Genom* 7:000685. <https://doi.org/10.1099/mgen.0.000685>
2. Aziz RK, Bartels D, Best AA et al (2008) The RAST Server: Rapid Annotations using Subsystems Technology. *BMC Genomics* 9:75. <https://doi.org/10.1186/1471-2164-9-75>

3. Overbeek R, Olson R, Pusch GD et al (2014) The SEED and the Rapid Annotation of microbial genomes using Subsystems Technology (RAST). *Nucleic Acids Res* 42:D206-214. <https://doi.org/10.1093/nar/gkt1226>
4. Brettin T, Davis JJ, Disz T et al (2015) RASTtk: a modular and extensible implementation of the RAST algorithm for building custom annotation pipelines and annotating batches of genomes. *Sci Rep* 5:8365. <https://doi.org/10.1038/srep08365>
5. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic local alignment search tool. *J Mol Biol* 215:403-410. [https://doi.org/10.1016/S0022-2836\(05\)80360-2](https://doi.org/10.1016/S0022-2836(05)80360-2)
6. Li W, O'Neill KR, Haft DH et al (2021) RefSeq: expanding the prokaryotic genome annotation pipeline reach with protein family model curation. *Nucleic Acids Res* 49:D1020-1028. <https://doi.org/10.1093/nar/gkaa1105>
7. Zimmermann L, Stephens A, Nam SZ, Rau D, Kübler J, Lozajic M, Gabler F, Söding J, Lupas AN, Alva V (2018) A completely reimplemented MPI bioinformatics toolkit with a new HHpred server at its core. *J Mol Biol* 430:2237-2243. <https://doi.org/10.1016/j.jmb.2017.12.007>
8. Finn RD, Clements J, Eddy SR (2011) HMMER web server: interactive sequence similarity searching. *Nucleic Acids Res* 39:W29-37. <https://doi.org/10.1093/nar/gkr367>
9. Potter SC, Luciani A, Eddy SR, Park Y, Lopez R, Finn RD (2018) HMMER web server: 2018 update. *Nucleic Acids Res* 46:W200-204. <https://doi.org/10.1093/nar/gky448>
10. Blum M, Chang H-Y, Chuguransky S et al (2021) The InterPro protein families and domains database: 20 years on. *Nucleic Acids Res* 49:D344-354. <https://doi.org/10.1093/nar/gkaa977>
11. Jones P, Binns D, Chang H-Y et al (2014) InterProScan 5: genome-scale protein function classification. *Bioinformatics* 30:1236-1240. <https://doi.org/10.1093/bioinformatics/btu031>
12. Wang J, Chitsaz F, Derbyshire MK et al (2023) The conserved domain database in 2023. *Nucleic Acids Res* 51:D384-388. <https://doi.org/10.1093/nar/gkac1096>
13. Marchler-Bauer A, Bo Y, Han L et al (2017) CDD/SPARCLE: functional classification of proteins via subfamily domain architectures. *Nucleic Acids Res* 45:D200-203. <https://doi.org/10.1093/nar/gkw1129>
14. Huerta-Cepas J, Szklarczyk D, Heller D et al (2019) eggNOG 5.0: a hierarchical, functionally and phylogenetically annotated orthology resource based on 5090 organisms and 2502 viruses. *Nucleic Acids Res* 47:D309-314. <https://doi.org/10.1093/nar/gky1085>
15. Alcock BP, Huynh W, Chalil R et al (2023) CARD 2023: expanded curation, support for machine learning, and resistome prediction at the Comprehensive Antibiotic Resistance Database. *Nucleic Acids Res* 51:D690-699. <https://doi.org/10.1093/nar/gkac920>
16. Bortolaia V, Kaas RS, Ruppe E et al (2020) ResFinder 4.0 for predictions of phenotypes from genotypes. *J Antimicrob Chemother* 75:3491-3500. <https://doi.org/10.1093/jac/dkaa345>
17. Xie Z, Tang H (2017) ISEScan: automated identification of insertion sequence elements in prokaryotic genomes. *Bioinformatics* 33:3340-3347. <https://doi.org/10.1093/bioinformatics/btx433>
18. Johansson MKH, Bortolaia V, Tansirichaiya S, Aarestrup FM, Roberts AP, Petersen TN (2021) Detection of mobile genetic elements associated with antibiotic resistance in *Salmonella enterica* using a newly developed web tool: MobileElementFinder. *J Antimicrob Chemother* 76:101-109. <https://doi.org/10.1093/jac/dkaa390>

Table S2 Complete annotation of plasmid pHv20-59a

CDS ID	Start (bp)	Stop (bp)	Strand	Product	Gene
ID=0001	10	1,098	+	ATP-dependent RecD2 DNA helicase	<i>recD2</i>
ID=0002	1,139	1,246	+	hypothetical protein	
ID=0003	1,243	1,515	+	DUF2339 domain-containing protein	
ID=0004	1,596	2,615	+	HTH cro/C1-type domain-containing protein	
ID=0005	2,630	2,911	+	hypothetical protein	
ID=0006	2,898	3,302	+	hypothetical protein	
ID=0007	3,299	3,790	+	Restriction alleviation protein%2C Lar family	
ID=0008	3,820	4,341	+	hypothetical protein	
ID=0009	4,401	4,676	+	DNA-binding transcriptional regulator%2C FrmR family/metal-sensing transcriptional repressor	<i>frmR</i>
ID=0010	4,711	4,884	+	Major Facilitator Superfamily protein	
ID=0011	4,885	6,081	+	NreB protein	<i>nreB</i>
ID=0012	6,032	6,316	+	hypothetical protein	
ID=0013	6,433	6,873	-	DUF411 domain-containing protein	
ID=0014	6,887	7,021	-	DUF3149 domain-containing protein	
ID=0015	7,058	7,609	-	Zinc resistance-associated protein	
ID=0016	7,968	8,339	+	CopC domain-containing protein/copper resistance CopC family protein	<i>copC</i>
ID=0017	8,391	9,248	+	Copper resistance protein D	
ID=0018	9,398	9,745	-	Outer membrane protein, beta-barrel	
ID=0019	10,637	11,914	-	Circularly permuted type 2 ATP-grasp protein	
ID=0020	12,448	12,948	+	Blue (type 1) copper domain-containing protein	
ID=0021	12,963	13,628	+	Transcriptional regulator	
ID=0022	13,615	14,949	+	Histidine kinase	
ID=0023	15,018	16,871	+	Copper resistance system multicopper oxidase	
ID=0024	16,881	17,645	+	Copper resistance B	
ID=0025	17,776	18,192	+	Carboxypeptidase regulatory-like domain-containing protein	
ID=0026	18,353	18,682	-	Ferredoxin/our-helix bundle copper-binding protein	
ID=0027	19,103	19,825	-	Cytochrome c-type biogenesis CcdA family protein	
ID=0028	20,072	20,707	+	Peroxiredoxin	
ID=0029	20,779	21,186	+	Transcriptional regulator%2C MerR family DNA-binding protein	<i>merR</i>
ID=0030	21,266	23,740	+	Copper chaperone CopZ/heavy metal translocating P-type ATPase	<i>copZ</i>

ID=0031	23,774	23,902	+	Oxaloacetate decarboxylase%2C gamma chain	
ID=0032	23,968	24,798	+	thiol:disulfide interchange protein DsbG	<i>dsbG</i>
ID=0033	24,861	25,538	+	TVP38/TMEM64 family membrane protein	
ID=0034	25,706	26,155	+	Thioredoxin domain-containing protein	
ID=0035	26,223	27,035	+	Prolipoprotein diacylglyceryl transferase	<i>lgt</i>
ID=0036	27,032	27,574	+	Thiol:disulfide interchange protein DsbD N-terminal domain-containing protein	<i>dsbD</i>
ID=0037	27,575	28,309	+	Urease accessory protein UreH-like transmembrane domain-containing protein/Cytochrome C biogenesis protein transmembrane region/sulfite exporter TauE/SafE family protein	
ID=0038	28,306	28,731	-	HTH merR-type domain-containing protein	
ID=0039	28,978	29,214	+	Pentapeptide MXKDX repeat protein	
ID=0040	29,314	29,754	+	Metal-binding protein	
ID=0041	29,769	30,317	+	Thioredoxin domain-containing protein/TlpA family protein disulfide reductase	
ID=0042	30,330	30,713	+	DUF2933 domain-containing protein	
ID=0043	30,730	31,473	+	Cytochrome C biogenesis protein CcdA	<i>ccdA</i>
ID=0044	31,497	32,363	+	Sortilin N-terminal domain-containing protein/WD40/YVTN/BNR-like repeat-containing protein	
ID=0045	32,356	32,826	+	Cytochrome c domain-containing protein	
ID=0046	33,253	33,489	-	hypothetical protein	
ID=0047	34,167	34,373	+	DUF2933 domain-containing protein	
ID=0048	34,414	34,758	+	Cation transporter	
ID=0049	34,979	36,316	+	Amino acid permease/APC family permease	
ID=0050	36,344	36,802	+	ATPase	
ID=0051	36,795	39,107	+	Cadmium-translocating P-type ATPase/ Lead, cadmium, zinc, mercury and copper translocating P-type ATPase	
ID=0052	39,149	39,379	-	Integrase catalytic subunit	
ID=0053	39,403	39,963	-	IS3 family transposase	
ID=0054	40,020	40,298	-	Transposase family protein	
ID=0055	40,339	40,509	-	adenine nucleotide alpha hydrolase family protein/phosphoadenosine phosphosulfate reductase domain-containing protein	
ID=0056	40,743	41,036	+	Mycothioid-dependent maleylpyruvate isomerase metal-binding domain-containing protein	
ID=0057	41,121	41,369	+	HTH cro/C1-type domain-containing protein	
ID=0058	41,381	41,806	+	hypothetical protein	
ID=0059	41,819	42,505	+	RiboL-PSP-HEPN domain-containing protein	
ID=0060	42,545	43,216	-	Abasic site processing protein/SOS response-associated peptidase	

ID=0061	43,516	44,877	-	Competence protein CoiA/cell envelope integrity TolA family protein	<i>coiA</i>
ID=0062	45,174	45,701	-	N-acetyltransferase domain-containing protein	
ID=0063	45,925	46,377	+	antitoxin Xre/MbcA/ParS toxin-binding domain-containing protein	
ID=0064	46,374	46,838	+	RES domain-containing protein/RES family NAD ⁺ phosphorylase	
ID=0065	46,924	47,793	+	DUF4113 domain-containing protein/Y-family DNA polymerase/ DinB/UmuC family translesion DNA polymerase	
ID=0066	48,298	48,498	+	Replication, recombination and repair/viral genome integration into host DNA	
ID=0067	48,820	49,347	-	sulfite exporter TauE/SafE family protein/urease accessory protein UreH domain-containing Protein	<i>pseudounitary</i>
ID=0068	49,747	50,025	-	PKc_like superfamily	
ID=0069	50,090	50,518	-	TlpA disulfide reductase family protein	<i>tlpA</i>
ID=0070	50,522	50,689	-	TlpA family protein disulfide reductase	
ID=0071	51,134	51,415	-	ATP-binding cassette domain-containing protein	
ID=0072	51,361	51,690	-	ATP-binding cassette domain-containing protein	
ID=0073	51,668	51,811	-	ABC transporter ATP-binding protein	
ID=0074	51,808	52,239	-	ABC transporter permease	
ID=0075	52,381	52,758	-	ABC transporter permease	
ID=0076	52,871	53,044	-	ABC transporter permease	
ID=0077	53,041	53,340	-	ABC transporter permease	
ID=0078	53,370	53,888	-	Cytochrome C biogenesis protein CcdA	<i>ccdA</i>
ID=0079	54,074	54,361	-	Glutaredoxin/Thioredoxin family protein	
ID=0080	54,358	54,792	-	Heavy metal-responsive transcriptional regulator	
ID=0081	55,082	56,314	+	Multidrug efflux MFS transporter	
ID=0082	56,393	56,659	-	VTT domain-containing protein/TVP38/TMEM64 family protein	
ID=0083	56,668	57,147	-	TVP38/TMEM64 family protein	
ID=0084	57,319	59,775	-	Cation-translocating P-type ATPase/heavy metal translocating P-type ATPase/ Lead, cadmium, copper, zinc and mercury transporting ATPase	
ID=0085	59,926	60,150	+	heavy metal-associated domain-containing protein	
ID=0086	60,187	60,486	+	heavy metal-associated domain-containing protein	
ID=0087	60,470	60,607	-	hypothetical protein	
ID=0088	60,781	61,107	-	Four-helix bundle copper-binding protein	
ID=0089	61,301	61,528	+	Heavy-metal-associated domain-containing protein	
ID=0090	62,054	62,506	+	DUF4145 domain-containing protein	
ID=0091	62,626	62,847	-	hypothetical protein	

ID=0092	62,898	63,407	-	Surface antigen domain-containing protein	
ID=0093	63,657	63,908	+	Cadherin-like	
ID=0094	63,892	64,212	+	hypothetical protein	
ID=0095	64,218	64,700	+	PH domain-containing protein	
ID=0096	64,708	65,076	+	hypothetical protein	
ID=0097	65,172	65,489	-	TrfB transcriptional repressor protein domain-containing protein	<i>trfB</i>
ID=0098	65,659	67,317	+	Methyltransferase domain containing protein	
ID=0099	67,402	67,620	+	hypothetical protein	
ID=0100	67,794	68,102	+	hypothetical protein	<i>pseudounitary</i>
ID=0101	68,099	68,707	+	Thioredoxin domain-containing protein	
ID=0102	68,707	69,351	+	OmpA-like domain-containing protein	
ID=0103	69,348	70,127	+	hypothetical protein	
ID=0104	70,138	70,551	+	Transglycosylase SLT domain-containing protein	
ID=0105	70,548	70,910	+	Conjugal transfer protein TrbC/TrbC/VirB2 family protein	<i>trbC</i>
ID=0106	70,928	71,221	+	Type IV conjugative transfer system protein TraL	<i>traL</i>
ID=0107	71,221	71,778	+	Type IV conjugative transfer system protein TraE	<i>traE</i>
ID=0108	71,775	72,467	+	TraK protein	<i>traK</i>
ID=0109	72,460	73,947	+	Conjugal transfer protein TraB/TrbI/VirB10 family type IV secretion system	<i>traB</i>
ID=0110	73,952	74,449	+	Type IV conjugative transfer system protein TraV	<i>traV</i>
ID=0111	74,451	77,063	+	TraG P-loop domain-containing protein/TraG/VirB4 family ATPase	<i>traG</i>
ID=0112	77,073	77,855	+	Peptidase M48 domain-containing protein	
ID=0113	77,852	80,833	+	HD/PDEase domain-containing protein	
ID=0114	80,930	81,349	+	Antitoxin Xre/MbcA/ParS-like toxin-binding domain-containing protein	
ID=0115	81,389	82,012	+	RES domain protein/RES family NAD ⁺ phosphorylase	
ID=0116	82,259	82,465	-	hypothetical protein	
ID=0117	82,458	82,922	-	Resolvase domain-containing protein/recombinase family protein	
ID=0118	83,658	84,599	+	CHASE4 domain-containing protein	
ID=0119	84,590	85,144	+	GGDEF domain-containing protein	
ID=0120	85,166	85,489	+	EAL domain-containing protein	
ID=0121	85,532	86,056	+	EAL domain-containing protein	
ID=0122	86,423	87,034	-	YeeE/YedE family protein	
ID=0123	87,388	87,816	-	TIGR01244 family phosphatase	
ID=0124	87,837	88,748	-	MBL fold metallo-hydrolase	

ID=0125	89,036	89,182	-	hypothetical protein	
ID=0126	89,623	89,745	+	Oxaloacetate decarboxylase	
ID=0127	90,413	90,529	-	hypothetical protein	
ID=0128	90,877	91,503	+	Class I SAM-dependent methyltransferase	
ID=0129	91,919	92,113	+	Aminotransferase class III	
ID=0130	92,306	92,581	+	Cytochrome c oxidase subunit IV -like protein	
ID=0131	92,586	92,708	+	hypothetical protein	
ID=0132	92,825	93,667	-	N-acetyltransferase	
ID=0133	93,809	94,165	+	hypothetical protein	
ID=0134	94,158	94,568	+	DUF6573 domain-containing protein	
ID=0135	94,793	95,383	+	DUF6927 domain-containing protein	
ID=0136	95,567	96,019	+	hypothetical protein	
ID=0137	96,269	97,501	+	hypothetical protein	
ID=0138	97,690	98,151	+	hypothetical protein	
ID=0139	98,304	98,645	+	Single-stranded DNA-binding protein	
ID=0140	98,710	99,273	+	Type-F conjugative transfer system protein (TrbI_Ftype)	
ID=0141	99,368	100,183	+	Thiol:disulfide interchange protein/DsbC family protein	
ID=0142	100,256	100,636	+	hypothetical protein	
ID=0143	100,633	101,091	+	hypothetical protein	
ID=0144	101,088	101,423	+	hypothetical protein	
ID=0145	101,891	102,832	+	Recombination-associated protein RdgC	<i>rdgC</i>
ID=0146	103,028	103,804	+	Chromosome partitioning protein/ParA family protein	
ID=0147	103,806	104,975	+	ParB/RepB/Spo0J family partition protein	
ID=0148	104,981	105,262	+	HTH iclR-type domain-containing protein	
ID=0149	105,351	106,364	+	DUF2167 domain-containing protein	
ID=0150	106,473	107,534	+	YaaC-like Protein	
ID=0151	107,586	108,317	-	HEPN domain-containing protein	
ID=0152	108,668	109,624	-	Integrase family protein/tyrosine-type recombinase	
ID=0153	109,617	109,940	-	hypothetical protein	
ID=0154	110,081	113,053	-	Tn3 family transposase	
ID=0155	113,569	114,033	-	Hg(II)-responsive transcriptional regulator	<i>merR</i>

ID=0156	114,230	114,697	+	Transposase IS204/IS1001/IS1096/IS1165 DDE domain-containing protein	
ID=0157	114,698	115,060	+	ISL3 family transposase	
ID=0158	115,226	115,387	+	hypothetical protein	
ID=0159	115,378	115,755	-	DUF1232 domain-containing protein/YkvA family protein	
ID=0160	115,774	116,253	-	Major facilitator superfamily (MFS) profile domain-containing protein/MFS transporter	
ID=0161	116,281	117,096	-	Putative zinc-finger domain-containing protein	
ID=0162	117,089	117,673	-	RNA polymerase subunit sigma-24	
ID=0163	117,747	119,243	-	Cytochrome c domain-containing protein/cytochrome D1 domain-containing protein	
ID=0164	119,430	120,104	-	Mercuric resistance operon regulatory protein/MerR family transcriptional regulator	
ID=0165	120,269	120,631	+	Mercuric ion transporter MerT	<i>merT</i>
ID=0166	120,645	120,941	+	Mercury resistance system periplasmic binding protein MerP	<i>merP</i>
ID=0167	120,947	122,350	+	Mercury(II) reductase	<i>merA</i>
ID=0168	122,296	122,583	+	Alkylmercury lyase MerB	
ID=0169	122,635	122,745	-	hypothetical protein	
ID=0170	122,745	123,734	-	Site-specific Integrase	
ID=0171	123,929	124,984	+	KfrA N-terminal DNA-binding domain-containing protein	<i>kfrA</i>
ID=0172	125,054	125,437	+	DUF1232 domain-containing protein/YkvA family protein	
ID=0173	125,516	125,635	-	hypothetical protein	
ID=0174	126,455	127,486	+	Plasmid replication protein	
ID=0175	128,647	129,201	-	DNA repair protein RadC	<i>radC</i>
ID=0176	129,345	130,247	+	Antirestriction protein/ArdC family protein	
ID=0177	130,587	132,047	+	Group II intron reverse transcriptase/maturase	<i>ltrA</i>
ID=0178	132,383	133,411	+	DNA-binding protein	
ID=0179	133,819	134,337	+	Integrase catalytic domain-containing protein/IS3 family transposase	
ID=0180	134,566	134,727	+	DUF411 domain-containing protein	
ID=0181	134,825	135,151	+	hypothetical protein	
ID=0182	135,103	136,023	-	Bile acid:sodium symporter family protein	
ID=0183	136,228	137,142	+	Cation efflux system protein/ Cobalt/zinc/cadmium resistance protein CzcD	<i>czcD</i>
ID=0184	137,256	137,933	+	Thioredoxin domain-containing protein/DsbA family protein	
ID=0185	137,933	138,922	+	Peptidase M23 domain-containing protein	
ID=0186	139,016	139,423	-	Cd(II)/Pb(II)-responsive transcriptional regulator	<i>cadR</i>
ID=0187	139,522	140,418	+	Sodium:proton antiporter/Cation efflux system protein	
ID=0188	140,560	141,321	+	Peptidoglycan-binding lysin domain-containing protein/ZIP family metal transporter	<i>zupT</i>

ID=0189	141,757	142,215	-	HTH-type transcriptional regulator ZntR/MerR family transcriptional regulator	<i>zntR</i>
ID=0190	142,302	144,488	+	Cadmium-translocating P-type ATPase/ heavy metal translocating P-type ATPase	<i>zntA</i>
ID=0191	144,485	144,988	+	Signal peptidase II	<i>lspA</i>
ID=0192	144,993	146,525	+	Apolipoprotein N-acyltransferase	<i>lnt</i>
ID=0193	146,528	147,601	+	LysM peptidoglycan-binding domain-containing protein	<i>lysM</i>
ID=0194	147,876	148,148	-	Phage gp6-like head-tail connector protein	
ID=0195	148,287	149,162	-	DUF6685 domain-containing protein	
ID=0196	149,252	149,935	+	DUF192 domain-containing protein	
ID=0197	149,998	150,276	+	Addiction module killer protein/type II toxin-antitoxin system RelE/ParE family Toxin	
ID=0198	150,411	150,695	+	HigA family addiction module antitoxin	<i>higA</i>
ID=0199	150,879	151,031	-	hypothetical protein	
ID=0200	151,262	151,810	-	hypothetical protein	
ID=0201	152,275	153,213	+	DNA-binding protein/Abi family protein/Abortive infection system protein AbiD/AbiF-like	
ID=0202	153,278	153,442	+	hypothetical protein	
ID=0203	153,479	154,384	+	Phage_DpoZ_2 superfamily	
ID=0204	155,367	155,780	+	Tyrosine specific protein phosphatases domain-containing protein	
ID=0205	155,929	156,141	+	hypothetical protein	
ID=0206	156,333	157,403	-	HEPN domain-containing protein	
ID=0207	157,643	157,957	-	hypothetical protein	
ID=0208	158,040	158,291	-	hypothetical protein	
ID=0209	158,496	160,193	-	DUF2326 domain-containing protein	
ID=0210	160,184	160,420	-	ABC-three component system middle component 6	
ID=0211	160,407	161,000	-	C-terminal domain 10 of the ABC-three component (ABC-3C) systems	
ID=0212	161,000	161,314	-	TMEM214, C-terminal, caspase 4 activator	
ID=0213	161,558	163,345	-	DNA 3'-5' helicase/ATP-dependent helicase	
ID=0214	163,338	165,131	-	AAA family ATPase/ATP-dependent nuclease	
ID=0215	165,595	165,966	+	hypothetical protein	
ID=0216	166,290	167,690	+	DNA (cytosine-5-)-methyltransferase	
ID=0217	167,900	168,991	+	Phage P4 alpha zinc-binding domain-containing protein	
ID=0218	169,124	169,354	-	hypothetical protein	
ID=0219	169,533	170,186	+	ERF family protein	
ID=0220	170,234	170,713	+	Membrane-bound metal-dependent hydrolase	

ID=0221	170,710	171,729	+	TraF-like protein/conjugal transfer protein TraF	
ID=0222	171,729	173,288	+	Conjugal transfer protein TraH	<i>traH</i>
ID=0223	173,288	176,392	+	TraG N-terminal Proteobacteria domain-containing protein	<i>traG</i>
ID=0224	176,422	176,766	-	2TM domain-containing protein	
ID=0225	176,851	177,102	-	hypothetical protein	
ID=0226	177,112	178,782	-	TraD/TraG TraM recognition site domain-containing protein/type IV secretory system conjugative DNA	<i>traM</i>
ID=0227	178,779	179,192	-	hypothetical protein	
ID=0228	179,192	179,782	-	hypothetical protein	
ID=0229	179,770	181,989	-	Conjugal transfer mating pair stabilization protein TraN	<i>traN</i>
ID=0230	181,993	183,360	-	Conjugal transfer protein TraN	<i>traN</i>
ID=0231	183,362	184,339	-	type-F conjugative transfer system pilin assembly protein TrbC	<i>trbC</i>
ID=0232	184,324	185,331	-	TraU family protein	<i>traU</i>
ID=0233	185,328	186,029	-	Conjugal transfer protein TraW	<i>traW</i>
ID=0234	186,007	186,474	-	Homeodomain-like domain-containing protein	
ID=0235	186,581	187,123	+	XRE family transcriptional regulator	
ID=0236	187,129	187,464	-	DUF2274 domain-containing protein	
ID=0237	187,605	188,222	+	Helix-turn-helix domain protein	
ID=0238	189,554	190,078	-	hypothetical protein	
ID=0239	190,276	192,630	-	DNA topoisomerase	
ID=0240	192,943	193,404	-	hypothetical protein	

Table S3 List of genetic metal tolerance determinants identified in the pHv20-59a

Mechanism	Gene	Function	Accession no. of proteins
copper resistance	<i>copA</i>	copper resistance protein CopA	MEA2118734
	<i>copB</i>	copper resistance protein CopB	MEA2118735
	<i>copC</i>	copper-binding protein CopC	MEA2118728
	<i>copD</i>	copper resistance CopD family protein	MEA2118729
	<i>copI</i>	cupredoxin domain-containing protein CopI	MEA2118731
	<i>copZ</i>	copper chaperone CopZ	MEA2118792, MEA2118793, MEA2118795
mercury resistance	<i>merA</i>	mercury (II) reductase	MEA2118863
	<i>merB</i>	alkylmercury lyase MerB	MEA2118864
	<i>merP</i>	mercury resistance system periplasmic binding protein MerP	MEA2118862
	<i>merT</i>	mercuric ion transporter MerT	MEA2118861
	<i>merR</i>	mercuric resistance operon transcriptional regulator	MEA2118855
Co/Zn/Cd resistance	<i>czcD</i>	cation diffusion facilitator (CDF) family transporter	MEA2118874, MEA2118878
zinc transport	<i>zupT</i>	zinc transporter ZupT	MEA2118879
ATPases		heavy metal translocating P-type ATPase	MEA2118741, MEA2118761, MEA2118881
		cation-transporting P-type ATPase	MEA2118791
efflux transport	<i>bcr/cflA</i>	multidrug resistance (MDR) MFS transporter	MEA2118788
transcriptional regulators	<i>nreA</i>	metal-sensing transcriptional repressor in CsoR-like superfamily	MEA2118724
	<i>cadR</i>	Cd(II)/Pb(II)-responsive transcriptional regulator	MEA2118877
		HTH DNA binding domain of heavy metal resistance transcription regulators	MEA2118740, MEA2118749, MEA2118787
		HTH DNA binding domain of transcription regulators from the MerR superfamily	MEA2118860
		HTH DNA binding domain of the CadR and PbrR transcription regulators	MEA2118880

Table S4 List of genetic metal tolerance determinants identified in the chromosome of the *Halovibrio* sp. HP20-59

Contig	ID-CDS	Start (bp)	Stop (bp)	Product	Gene
NODE_1_length_460873_cov_226.719807	JGLKGA_01215	247,046	248,449	mercury(II) reductase	<i>merA</i>
NODE_1_length_460873_cov_226.719807	JGLKGA_01290	261,909	264,074	Lead, cadmium, zinc and mercury-transporting ATPase	
NODE_1_length_460873_cov_226.719807	JGLKGA_00610	127,340	129,121	Cation acetate symporter	
NODE_1_length_460873_cov_226.719807	JGLKGA_01815	382,546	384,051	Cation:proton antiporter	
NODE_1_length_460873_cov_226.719807	JGLKGA_01995	421,122	421,835	Trimeric intracellular cation channel family protein	
NODE_1_length_460873_cov_226.719807	JGLKGA_00990	205,125	205,565	Zinc resistance-associated protein	
NODE_1_length_460873_cov_226.719807	JGLKGA_01025	213,959	214,969	Zinc-type alcohol dehydrogenase-like protein	
NODE_1_length_460873_cov_226.719807	JGLKGA_01280	260,551	261,327	Zinc transporter, ZIP family	
NODE_1_length_460873_cov_226.719807	JGLKGA_02050	439,010	439,978	Putative metal ion transporter YfjQ	<i>yfjQ</i>
NODE_1_length_460873_cov_226.719807	JGLKGA_00260	49,642	49,971	Small multidrug resistance protein	<i>emrE</i>
NODE_1_length_460873_cov_226.719807	JGLKGA_01220	248,781	249,215	Hg(II)-responsive transcriptional regulator	<i>merR</i>
NODE_1_length_460873_cov_226.719807	JGLKGA_01285	261,361	261,840	MerR family transcriptional regulator	<i>merR</i>
NODE_1_length_460873_cov_226.719807	JGLKGA_00200	39,304	40,338	Nickel/cobalt efflux system	
NODE_1_length_460873_cov_226.719807	JGLKGA_00235	44,751	46,055	MATE family efflux transporter	
NODE_2_length_435090_cov_172.356776	JGLKGA_02290	19,677	21,614	Copper-transporting P-type ATPase	
NODE_2_length_435090_cov_172.356776	JGLKGA_03700	319,159	319,662	Copper-bind domain-containing protein	
NODE_2_length_435090_cov_172.356776	JGLKGA_03715	321,732	323,573	Copper resistance system multicopper oxidase	
NODE_2_length_435090_cov_172.356776	JGLKGA_03720	323,584	324,354	Copper resistance protein B	
NODE_2_length_435090_cov_172.356776	JGLKGA_03745	328,086	330,311	Copper-exporting P-type ATPase A	
NODE_2_length_435090_cov_172.356776	JGLKGA_02935	162,692	163,579	DNA replication terminus site-binding protein	
NODE_2_length_435090_cov_172.356776	JGLKGA_04180	422,658	425,372	Cation-transporting P-type ATPase	
NODE_2_length_435090_cov_172.356776	JGLKGA_02275	18,431	18,982	Zinc resistance-associated protein	
NODE_2_length_435090_cov_172.356776	JGLKGA_02450	54,888	56,711	Zinc chelation protein SecC	<i>secC</i>
NODE_2_length_435090_cov_172.356776	JGLKGA_03740	327,412	327,984	Zinc resistance-associated protein	
NODE_2_length_435090_cov_172.356776	JGLKGA_02240	14,589	14,864	Metal-sensing transcriptional repressor	
NODE_2_length_435090_cov_172.356776	JGLKGA_03790	339,636	340,379	ZIP family metal transporter	
NODE_2_length_435090_cov_172.356776	JGLKGA_02235	13,629	14,546	CDF family Co(II)/Ni(II) efflux transporter DmeF	<i>dmeF</i>
NODE_3_length_340243_cov_257.234027	JGLKGA_05065	191,500	192,969	Divalent metal cation transporter	
NODE_4_length_325885_cov_216.327966	JGLKGA_05820	22,258	23,235	Cation transporter	
NODE_4_length_325885_cov_216.327966	JGLKGA_06920	255,630	255,911	Monovalent cation/H ⁺ antiporter subunit G	
NODE_4_length_325885_cov_216.327966	JGLKGA_05725	655	1,416	Metallophos domain-containing protein	

NODE_4_length_325885_cov_216.327966	JGLKGA_07215	315,354	316,664	M20/M25/M40 family metallo-hydrolase	
NODE_4_length_325885_cov_216.327966	JGLKGA_06615	188,169	189,440	Multidrug resistance protein mexA	<i>mexA</i>
NODE_4_length_325885_cov_216.327966	JGLKGA_06620	189,440	192,628	Efflux pump membrane transporter	
NODE_5_length_284240_cov_191.739055	JGLKGA_07500	56,249	56,614	MerC mercury resistance protein	<i>merC</i>
NODE_5_length_284240_cov_191.739055	JGLKGA_07455	42,940	44,451	monovalent cation/H ⁺ antiporter subunit D	
NODE_5_length_284240_cov_191.739055	JGLKGA_07465	44,780	47,620	monovalent cation/H ⁺ antiporter subunit A	<i>nha</i>
NODE_5_length_284240_cov_191.739055	JGLKGA_07320	15,607	16,644	putative oxidoreductase (Zinc-containing alcohol dehydrogenase family)	
NODE_5_length_284240_cov_191.739055	JGLKGA_07480	50,441	52,426	ATP-dependent zinc metalloprotease FtsH	<i>ftsH</i>
NODE_5_length_284240_cov_191.739055	JGLKGA_07355	21,809	22,555	Metallophos domain-containing protein	
NODE_5_length_284240_cov_191.739055	JGLKGA_07745	113,313	114,764	metalloprotease TldD	<i>tldD</i>
NODE_5_length_284240_cov_191.739055	JGLKGA_07755	115,415	116,755	metalloprotease PmbA	<i>pmbA</i>
NODE_5_length_284240_cov_191.739055	JGLKGA_07860	132,606	133,367	Nif3-like dinuclear metal center hexameric protein	
NODE_5_length_284240_cov_191.739055	JGLKGA_08500	272,175	272,603	Transcriptional regulator, MerR family	<i>merR</i>
NODE_6_length_238926_cov_247.015236	JGLKGA_08715	27,360	28,292	Copper ABC transporter ATP-binding protein	
NODE_6_length_238926_cov_247.015236	JGLKGA_08720	28,289	29,626	Copper-binding periplasmic protein	
NODE_6_length_238926_cov_247.015236	JGLKGA_09070	103,101	105,698	Copper-translocating P-type ATPase	
NODE_6_length_238926_cov_247.015236	JGLKGA_09320	152,636	153,994	Zinc metalloprotease	
NODE_6_length_238926_cov_247.015236	JGLKGA_09565	207,259	208,485	zinc metallochaperone GTPase ZigA	<i>zigA</i>
NODE_6_length_238926_cov_247.015236	JGLKGA_09120	116,433	117,836	MFS transporter, DHA2 family, multidrug resistance protein	
NODE_8_length_171608_cov_227.581096	JGLKGA_11565	165,697	167,355	multidrug ABC transporter permease/ATP-binding protein	
NODE_8_length_171608_cov_227.581096	JGLKGA_11240	72,520	73,719	Bcr/CflA family efflux transporter	
NODE_9_length_160094_cov_272.828443	JGLKGA_11735	29,170	29,688	Copper chaperone PCu(A)C	
NODE_9_length_160094_cov_272.828443	JGLKGA_11970	84,224	85,417	Ferrous-iron efflux pump FieF	<i>fieF</i>
NODE_10_length_152189_cov_192.772569	JGLKGA_13030	149,365	149,886	Zinc-ribbon containing domain-containing protein	
NODE_11_length_137415_cov_208.601305	JGLKGA_13500	99,032	100,114	MexH family multidrug efflux RND transporter periplasmic adaptor subunit	<i>mexH</i>
NODE_11_length_137415_cov_208.601305	JGLKGA_13505	100,114	103,239	Multidrug transporter AcrB	<i>acrB</i>
NODE_11_length_137415_cov_208.601305	JGLKGA_13075	8,471	9,550	ACR3 family arsenite efflux transporter	<i>arsB</i>
NODE_11_length_137415_cov_208.601305	JGLKGA_13120	17,604	18,791	Bcr/CflA family efflux transporter	
NODE_11_length_137415_cov_208.601305	JGLKGA_13500	99,032	100,114	MexH family multidrug efflux RND transporter periplasmic adaptor subunit	<i>mexH</i>
NODE_12_length_132718_cov_269.566288	JGLKGA_13985	71,447	72,367	MerR family transcriptional regulator	<i>merR</i>
NODE_13_length_104088_cov_247.148734	JGLKGA_14470	28,640	28,906	DksA/TraR family C4-type zinc finger protein	

NODE_14_length_99360_cov_261.926946	JGLKGA_14995	34,620	34,985	Zinc chelation protein SecC	<i>secC</i>
NODE_16_length_85211_cov_199.176639	JGLKGA_15835	28,450	30,651	ATPase, P-type, heavy metal translocating	
NODE_16_length_85211_cov_199.176639	JGLKGA_15880	37,840	38,856	Metallophos domain-containing protein	
NODE_16_length_85211_cov_199.176639	JGLKGA_16090	80,013	81,431	Multidrug resistance protein, MATE family	
NODE_17_length_66127_cov_211.700045	JGLKGA_16190	21,461	21,679	metallothionein	
NODE_17_length_66127_cov_211.700045	JGLKGA_16325	44,408	45,322	Multidrug DMT transporter	
NODE_18_length_63652_cov_261.771986	JGLKGA_16540	21,178	23,142	ATP-dependent zinc metalloprotease FtsH	<i>ftsH</i>
NODE_18_length_63652_cov_261.771986	JGLKGA_16470	5,033	5,485	ferric iron uptake transcriptional regulator	<i>fur</i>
NODE_19_length_62727_cov_204.887167	JGLKGA_16870	24,510	25,493	Hydrolase or metal-binding protein	
NODE_22_length_46039_cov_218.932901	JGLKGA_17575	362	727	Arsenic resistance transcriptional regulator ArsR2	<i>arsR2</i>
NODE_26_length_36372_cov_187.463342	JGLKGA_18400	13,298	14,650	MATE family efflux transporter	
NODE_28_length_32218_cov_202.991755	JGLKGA_18770	22,778	23,182	monovalent cation/H(+) antiporter subunit G	<i>mnhG</i>
NODE_28_length_32218_cov_202.991755	JGLKGA_18775	23,175	23,444	Cation:proton antiporter	
NODE_28_length_32218_cov_202.991755	JGLKGA_18780	23,444	23,917	Cation:proton antiporter	
NODE_28_length_32218_cov_202.991755	JGLKGA_18790	25,506	25,853	Cation:proton antiporter	
NODE_28_length_32218_cov_202.991755	JGLKGA_18800	26,296	28,665	putative monovalent cation/H ⁺ antiporter subunit A	
NODE_29_length_29857_cov_257.392411	JGLKGA_18910	20,012	20,263	YheV family metal-binding protein	<i>yheV</i>

Table S5 List of plasmids used in the comparative analyses based on BLAST results

Plasmid	Host microorganism	Length (bp)	GC content (%)	Ecosystem	Query cover (%)	Identity (%)
pBAA-803-A	<i>Vreelandella sulfidaeris</i> ATCC BAA-803	273,549	53.4	cell culture, metal sulfide rock	70	96.98
unnamed1	<i>Halomonas</i> sp. KO116	313,995	54.8	Nile Delta Mediterranean Sea surface	68	97.49
pHAL1	<i>Halomonas</i> sp. HAL1	193,527	53.5	cell culture, soil	68	98.58
plasA	<i>Vreelandella</i> sp. SM1641	302,095	54.9	cell culture, deep marine sediment	68	97.87

Table S6 List of plasmids used in the comparative analyses downloaded from the IMG/PR database and associated information

Plasmid ID	PTU	Closest reference	Closest reference ANI (%)	MOB genes	T4CP genes
IMGPR_plasmid_2905650138_000002	PTU_00019512	AP019515.1	98.14	MOBF; MOBF	<i>t4cp2</i>
IMGPR_plasmid_3300003108_000005	PTU_00012220	AP019515.1	97.02	MOBF	<i>t4cp2</i>
IMGPR_plasmid_3300027834_000114	PTU_00021634	AP019515.1	94.03	-	-
IMGPR_plasmid_3300031976_000086	PTU_00019512	NZ_CP011053.1	94.31	-	-
IMGPR_plasmid_3300032083_000208	PTU_00019512	NZ_CP011053.1	95.25	-	-
Plasmid ID	T4SS genes	Complete MPF family	Other conjugation genes		
IMGPR_plasmid_2905650138_000002	<i>F_traU</i> ; <i>virb4</i>	MPF_F	<i>F_traW</i> ; <i>F_trbC</i> ; <i>F_traN</i> ; <i>F_traG</i> ; <i>F_traH</i> ; <i>F_traF</i> ; <i>F_traV</i> ; <i>F_traB</i> ; <i>F_traK</i> ; <i>F_traE</i> ; <i>F_traL</i>		
IMGPR_plasmid_3300003108_000005	<i>F_traU</i> ; <i>virb4</i>	MPF_F	<i>F_traW</i> ; <i>F_trbC</i> ; <i>F_traN</i> ; <i>F_traG</i> ; <i>F_traH</i> ; <i>F_traF</i> ; <i>F_traV</i> ; <i>F_traB</i> ; <i>F_traK</i> ; <i>F_traE</i> ; <i>F_traL</i>		
IMGPR_plasmid_3300027834_000114	<i>virb4</i>	-	<i>F_traV</i> ; <i>F_traB</i> ; <i>F_traK</i> ; <i>F_traE</i> ; <i>F_traL</i>		

IMGPR_plasmid_3300031976_000086	virb4	-	F_traL; F_traE; F_traK; F_traB; F_traV
IMGPR_plasmid_3300032083_000208	virb4	-	F_traV; F_traB; F_traK; F_traE; F_traL
Plasmid ID	Source type	Source genome/Metagenome	
IMGPR_plasmid_2905650138_000002	isolate	Halomonas_sulfidaeris SST4	
IMGPR_plasmid_3300003108_000005	metagenome	Groundwater microbial communities from Illinois-Basin Decatur Project, Sample - 1.8km deep Mt. Simon Sandstone pore	
IMGPR_plasmid_3300027834_000114	metagenome	Oil polluted marine microbial communities from Coal Oil Point, Santa Barbara, California, USA - Santa Barbara Oil Seep Sample 6 (SPAdes) (version 2)	
IMGPR_plasmid_3300031976_000086	metagenome	Marine viral communities from the Yellow sea, Qingdao, China - WS0712	
IMGPR_plasmid_3300032083_000208	metagenome	Marine viral communities from the Yellow sea, Qingdao, China - XG0712	
Plasmid ID	Scaffold	Scaffold length (%)	Gene count
IMGPR_plasmid_2905650138_000002	Ga0452674_11	130979	158
IMGPR_plasmid_3300003108_000005	Ga0052255_103175	135411	172
IMGPR_plasmid_3300027834_000114	Ga0209344_10000508	66624	89
IMGPR_plasmid_3300031976_000086	Ga0326496_1000040	101715	155
IMGPR_plasmid_3300032083_000208	Ga0326489_1000189	50999	79
Plasmid ID	Host prediction		
IMGPR_plasmid_2905650138_000002	Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales; Halomonadaceae		
IMGPR_plasmid_3300003108_000005	Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales; Halomonadaceae		
IMGPR_plasmid_3300027834_000114	Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales; Halomonadaceae		
IMGPR_plasmid_3300031976_000086	Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales; Halomonadaceae		
IMGPR_plasmid_3300032083_000208	Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales; Halomonadaceae		

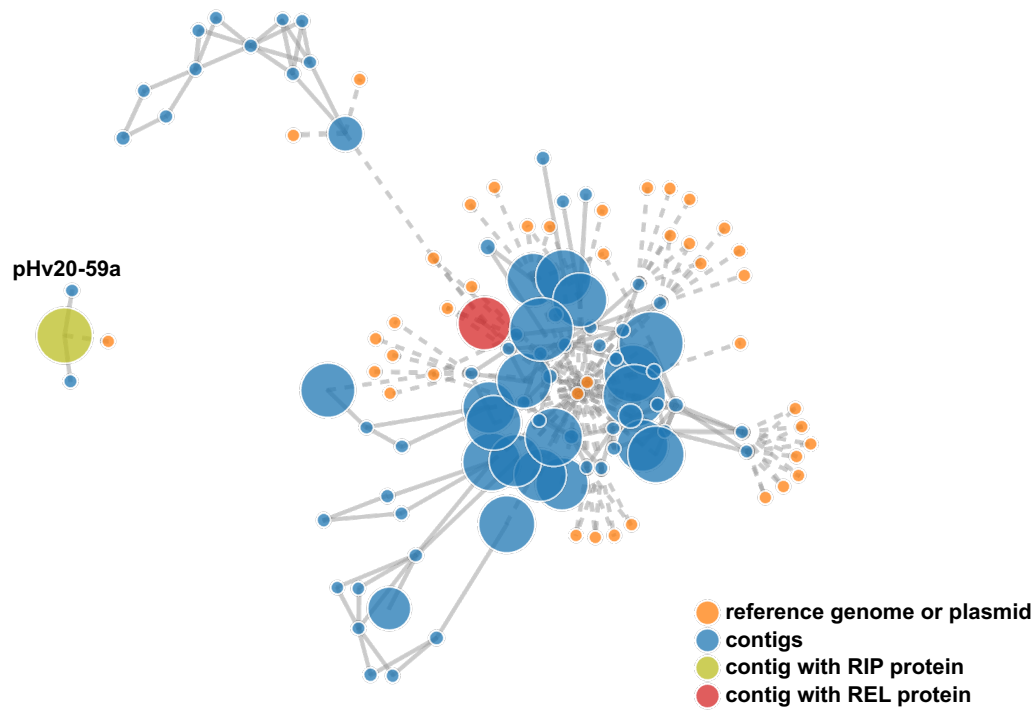


Fig. S1 PLACNETw reconstruction of the *Halovibrio* sp. HP20-59 draft genome. The chromosome contigs are shown as the blue nodes and the size of nodes is proportional to the contig length. The orange nodes represent the reference genomes or plasmids. The yellow-green node represents contig with the replication initiation protein – RIP and the red node represents the contig with relaxase – REL. This reconstruction results in only one plasmid. The full lines represent the homology to the reference and the dashed line represents the scaffold link

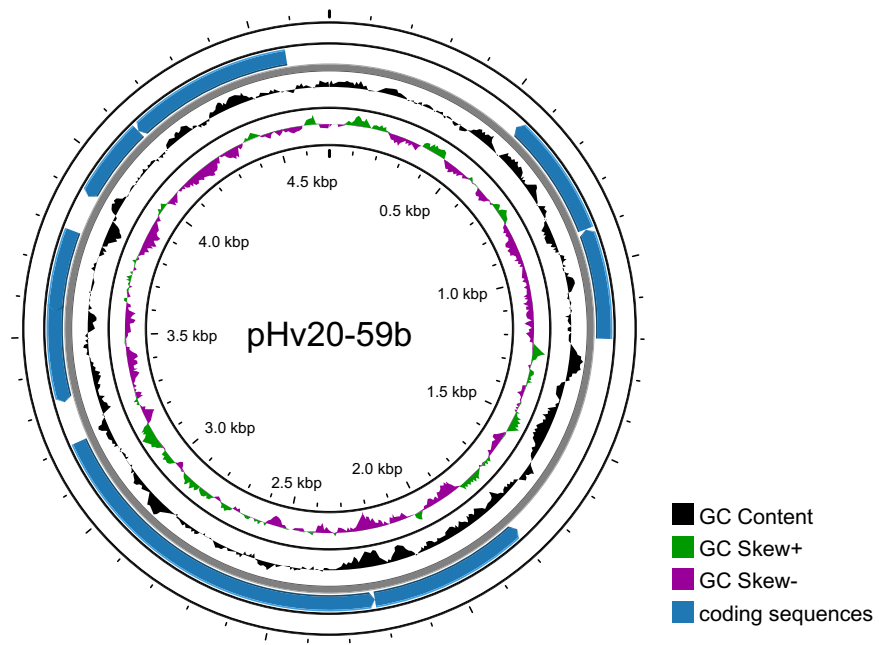


Fig. S2 Circular visualization of plasmid pHv20-59b illustrating the distribution of coding sequences

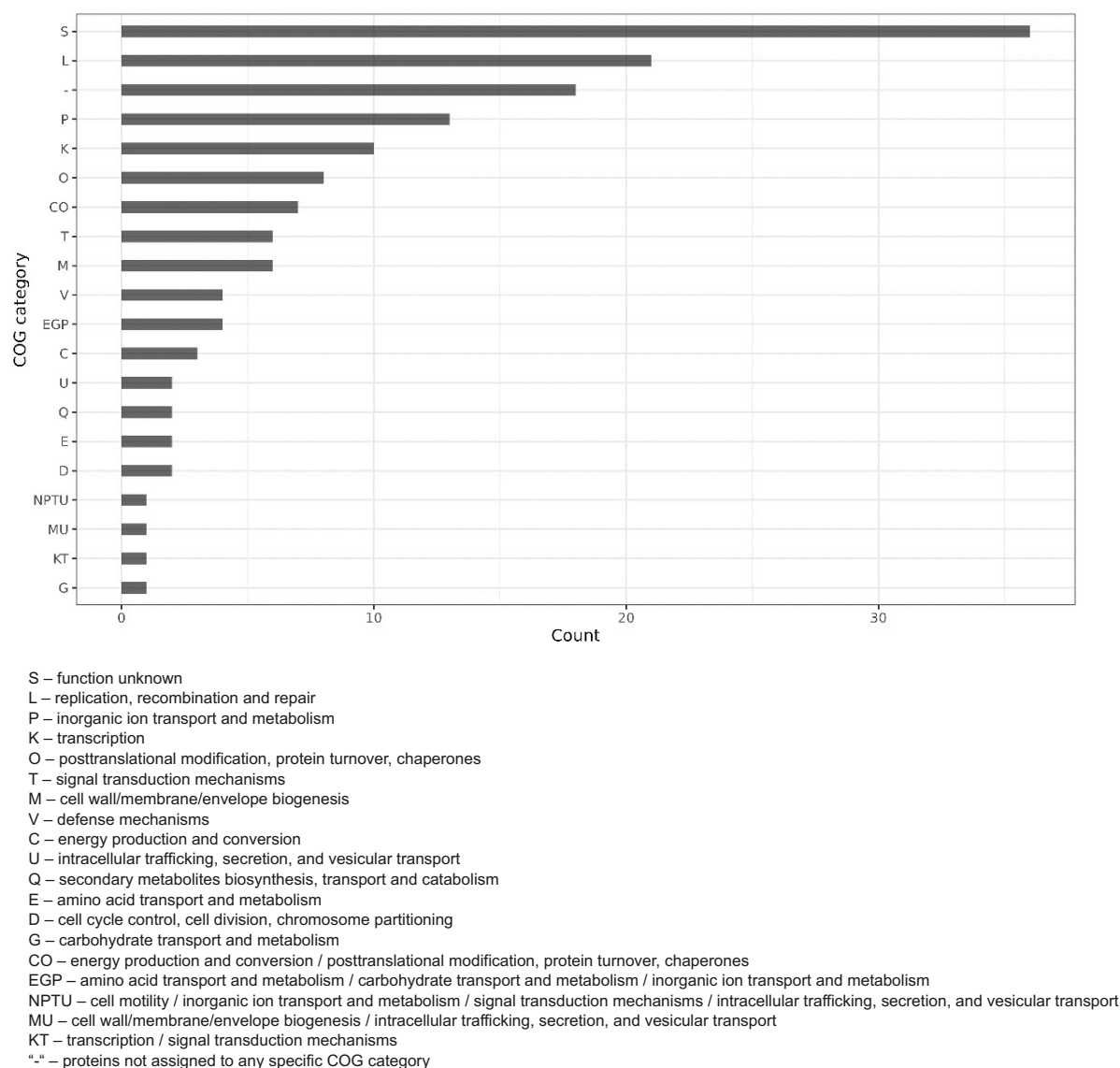


Fig. S3 COG category distribution for pHv20-59a

annotations are displayed as orange arrows along each sequence. For simplicity, plasmids are labeled using only the last two or three digits of their original identifiers