

New mobile genetic elements in *Cupriavidus metallidurans* CH34, their possible roles and occurrence in other bacteria

Rob Van Houdt · Sébastien Monchy ·
Natalie Leys · Max Mergeay

Received: 21 January 2009 / Accepted: 18 March 2009 / Published online: 24 April 2009
© Springer Science+Business Media B.V. 2009

Abstract *Cupriavidus metallidurans* strain CH34 is a β -Proteobacterium that thrives in low concentrations of heavy metals. The genetic determinants of resistance to heavy metals are located on its two chromosomes, and are particularly abundant in the two megaplasmids, pMOL28 and pMOL30. We explored the involvement of mobile genetic elements in acquiring these and others traits that might be advantageous in this strain using genome comparison of *Cupriavidus/Ralstonia* strains and related β -Proteobacteria. At least eleven genomic islands were identified on the main replicon, three on pMOL28 and two on pMOL30. Multiple islands contained genes for heavy metal resistance or other genetic determinants putatively responding to harsh environmental conditions. However, cryptic elements also were noted. New mobile genetic elements (or variations of known ones) were identified through synteny analysis, allowing the detection of mobile genetic elements outside the bias of a selectable marker. Tn4371-like conjugative transposons involved in chemolithotrophy and degradation of aromatic compounds were identified in strain CH34,

while similar elements involved in heavy metal resistance were found in *Delftia acidovorans* SPH-1 and *Bordetella petrii* DSM12804. We defined new transposons, viz., Tn6048 putatively involved in the response to heavy metals and Tn6050 carrying accessory genes not classically associated with transposons. Syntenic analysis also revealed new transposons carrying metal response genes in *Burkholderia xenovorans* LB400, and other bacteria. Finally, other putative mobile elements, which were previously unnoticed but apparently common in several bacteria, were also revealed. This was the case for triads of tyrosine-based site-specific recombinases and for an *int* gene paired with a putative repressor and associated with chromate resistance.

Keywords Genomic islands · Tn4371 · Tyrosine-based site-specific recombinase · Integrase · Heavy metals · Chemolithotrophy · Transposon · Annotation · Synteny

Metal-resistant *Cupriavidus* strains

Strain CH34; from *Alcaligenes eutrophus* to *Cupriavidus metallidurans*

Soon after the isolation of strain CH34, from metallurgical sediments in Belgium (Mergeay et al. 1978, 1985), attention focused on its resistance to

R. Van Houdt · S. Monchy · N. Leys · M. Mergeay (✉)
Unit for Microbiology, SCK•CEN, Boeretang 200, 2400
Mol, Belgium
e-mail: mmergeay@sckcen.be

S. Monchy
Biology Department, Brookhaven National Laboratory,
Upton, NY 11973-5000, USA

multiple heavy metals, especially cadmium, zinc, cobalt, and nickel, for which resistance determinants were unknown, and for the presence of two large plasmids named pMOL28 and pMOL30 (Mergeay et al. 2009; Monchy et al. 2007; Nies et al. 1987, 1989; Taghavi et al. 1997). Further research focused on unraveling the determinants of metal resistance, on its genetics, especially since CH34 was a good recipient of foreign genes, on biotechnological- and ecological-applications of these determinants (Diels and Mergeay 1990; Mergeay 1991, 2000; Mergeay et al. 2003), even on the frequently changing taxonomy (Brim et al. 1999; Goris et al. 2001; Sato et al. 2006; Vandamme and Coenye 2004; Yabuuchi et al. 1995), and finally on genome sequencing.

Genomic approaches in the *Cupriavidus/Ralstonia* genus

To date, six genomes are fully sequenced in the closely related *Cupriavidus/Ralstonia* genera: the plant pathogen *R. solanacearum* GMI1000 (Salanoubat et al. 2002), the β -rhizobium legume symbiont *C. taiwanensis* LMG19424 (Amadou et al. 2008), the facultative chemolithotroph *C. eutrophus* H16 (Pohlmann et al. 2006; Schwartz et al. 2003), the catabolic strain *C. pinatubonensis* JMP134 (GenBank Accession CP000090-CP000094), the opportunistic pathogen *R. pickettii* 12J (GenBank Accession CP0001068-CP0001070), and the multiple metal resistant strain *C. metallidurans* CH34 (Mergeay 2000; Genbank Accession CP000352-CP000355). At least three other strains of *R. solanacearum* are being sequenced as well as another strain of *R. pickettii* (strain 12D). There is an updated annotation of the different replicons of *C. metallidurans* CH34 available on GenoScope's MaGe system (Vallenet et al. 2006; www.genoscope.cns.fr/agc/mage/wwwpkgdb/Login/log.php?pid=33).

All of these strains typically carry two chromosomes; the larger (3–4 Mb) shelters most of the housekeeping genes, while the second (2–3 Mb) seemingly has plasmid-like traits for its replication; hence sometimes it is called a megaplasmid. This designation might, however, be confusing since other native plasmids can be very large, and the second chromosome generally is considered to be indispensable since it carries essential genes. This two-chromosome arrangement is also characteristic of the *Burkholderia* genus, including many strains with

(multiple) large plasmids and in some up to nine coexisting replicons (Fricke et al. 2009). The native plasmids have often specific traits that are linked to the very different ecological niches occupied by these closely related bacteria. *C. eutrophus* H16 has a large plasmid pHG1 carrying genes for chemolithotrophy and anaerobic growth on nitrates (Schwartz et al. 2003), plasmid pRALTA in *C. taiwanensis* LMG19424 is involved in legume symbiosis and nitrogen fixation, while plasmids pMOL28 and pMOL30 in *C. metallidurans* CH34 carry many genetic determinants for high resistance to heavy metals (Mergeay et al. 2009; Monchy et al. 2007).

Mobile genetic elements and niche specialization of *Cupriavidus/Ralstonia* strains

Mobile genetic elements such as composite, unit, and conjugative transposons, plasmids, and genomic islands often carry key determinants ensuring survival in environmental niches, such as hydrogenotrophy, degradation of chloroaromatic compounds, heavy metal resistance, or nitrogen fixation. *In silico* comparisons of the *Cupriavidus/Ralstonia* genomes can shed light on the role of mobile genetic elements in adaptation to specific environmental niches, such as symbiotic association with plants, chemolithotrophy, plant pathogenesis, or the colonization of industrial sites. Furthermore, we can assess the evolution and dissemination of different mobile genetic elements in the genera *Cupriavidus/Ralstonia* and beyond. Our goal was to make an inventory of the genomic islands and associated mobile genetic elements in *C. metallidurans* CH34 using the following tools: the MAGE GenoScope platform (Vallenet et al. 2006) to evaluate the synteny between the analyzed genomes and other genomes or replicons; the ACLAME database for mobile genetic elements (<http://aclame.ulb.ac.be/>; Leplae et al. 2004, 2006); the IS Finder database for IS elements (<http://www.is.biotoul.fr/>; Siguier et al. 2006); and IslandPath (<http://www.pathogenomics.sfu.ca/islandpath/>). The latter graphically presents the genome of an organism, highlighting features such as GC content, dinucleotide bias, proximity to tRNA or other structural RNA genes. Also, it denotes the presence of mobility genes (transposases and integrases) to identify regions likely to be acquired through horizontal gene transfer, since these features commonly are

associated with genomic islands (Hsiao et al. 2003). This inventory can confer a better understanding of the role of genomic islands in bacteria, and how bacteria recruit advantageous genes. In addition, it can illuminate the origin and dispersion of these genes, and the evolutionary forces shaping this bacterium.

Defining genomic islands

Genomic islands were mainly looked for in regions with extensive lack of synteny and in prokaryotic genomes these islands are commonly defined as clusters of genes with one or a set of following properties:

- a tyrosine-based site-specific recombinase gene and an adjacent tRNA gene at one extremity
- flanking insertion sequence elements
- a base composition and/or phylogeny differing from the bulk of the genome, indicating a foreign origin and, hence, acquisition through horizontal transfer
- a higher content of hypothetical genes than neighboring regions
- clustering of genes characteristic of mobile genetic elements: recombinase genes, IS elements
- conservation of the genomic island between different (unrelated) hosts
- a high concentration of genes specialized for resistance, catabolism, unusual metabolisms

Most often, however, as for most pathogenicity islands (reviewed by Hacker et al. 2004), intercellular mobility has not been demonstrated experimentally. Under this broad definition, the term genomic island covers a wide spectrum of integrated mobile genetic elements that are not well-defined. They therefore may be defective, or have lost their determinants of intercellular mobility, but are inferred to be or to have been mobile (or mobilized in trans by a proficient mobile genetic element; E.M. Top and A. Toussaint, personal communication). Thus, the term genomic island is broader than the integrative and conjugative elements that are transferred through conjugation, integrate into and replicate with the host genome. In addition, the term genomic island does not necessarily include structures defined as regions of genomic plasticity that do not reflect their evolutionary origin or genetic basis (Mathee et al. 2008).

Islands in chromosome 1

CMGI-1, a genomic island identical to PAGI-2C of *P. aeruginosa* clone C

The first observation attracting attention to the genomic islands in *C. metallidurans* CH34 was that this strain, as well as some related strains had a large genomic island comparable to the one of *P. aeruginosa* clone C strains isolated from cystic fibrosis patients (Klockgether et al. 2006; Larbig et al. 2002). This 109 kb PAGI-2C island that begins with an *int* gene (tyrosine-based site-specific recombinase) is fully conserved in *C. metallidurans* CH34 (with only a few base pairs different; Table 1). The CH34 version, adjacent to a tRNA-Gly gene, contains an additional endogenic IS element IS1088 (IS30 family), and its *int* gene is inactivated by transposon Tn6049 (see below) although no spontaneous excision from *P. aeruginosa* clone C has been demonstrated and most of the PAGI-2C genes are transcriptionally silent in *P. aeruginosa* clone C (Klockgether et al. 2008). However, in CH34 expression of some genes (*cdfX*, *pbrR2*, *cadA*, *pbrC2* see below) has already been demonstrated in the presence of metals and putatively transcription of the *int* gene might differ in CH34 (and other genetic backgrounds). The island could therefore be prone to excision and insertion of Tn6049 fixed the island in the genome. Parts of the CMGI-1/PAGI-2C genomic island are highly conserved in various β -Proteobacteria (e.g., *Burkholderia xenovorans*, *Delftia acidovorans*, and *Herminiimonas arsenicoxydans*) and in an α -Proteobacterium (*Parvibaculum lavamentivorans* strain DS-1; Schleheck et al. 2004b, 2007) with protein identities higher than 80%. These regions are even conserved in γ -Proteobacteria with significant identities (>35% identity). Synteny reveals that blocs of 10–20 genes, mostly encoding hypothetical proteins, are highly conserved (throughout α -, β -, and even γ -Proteobacteria) and separated by identifiable functions, such as the metal-resistance or cytochrome biosynthesis genes, that are present in CMGI-1 but lacking in other bacteria. Nevertheless, the presence of an *int* gene (tyrosine-based site-specific recombinase) near a tRNA gene still provides the hallmark of a genomic island. Figure 1 shows an example for CH34 versus *R. pickettii* 12D, *H. arsenicoxydans* HEAR, *B. xenovorans* LB400, and *P. lavamentivorans* DS-1 indicating that the CMGI-1/PAGI-2C family of

Table 1 Characteristics of the genomic islands on chromosome 1 of *C. metallidurans* CH34

Island	ORFs (Rmet_)	Coordinates length (bp)	Features
CMGI-1	2287–2408	2,508,632–2,618,230 (109,598)	Integrase (disrupted by Tn6049), direct repeats (include 2 bp of 3' end tRNA-Gly, GACTCGT(T/C)TCCCGTCCAA)
CMGI-2	1236–1351	1,359,263–1,460,900 (101,637)	Almost identical to PAGI-2C of <i>P. aeruginosa</i> clone C Tn4371-family, integrase, direct repeats (TTTTTGTT)
CMGI-3	1465–1560	1,585,489–1,682,531 (97,042)	Hydrogenotrophy, metabolism of aromatic compounds 2 RIT elements (recombinases in trio/triad)
CMGI-4	2987–3045	3,245,127–3,301,656 (56,529)	Tn4371-family, integrase, no direct repeats CO ₂ fixation and hydrogenotrophy
CMGI-5	2824–2847	3,071,473–3,096,896 (25,423)	Tn4371-family, integrase, direct repeats (TTTTTCAT) No conjugative module, Tn6048
CMGI-6	1997–2020	2,167,319–2,184,957 (17,638)	Integrase, direct repeats (63 bp of 3' end tRNA-Leu, TGGTGCCCGGGTCCGACTCGAACCACACGCCCTTGGCGGGGGGATTTTGAGTCCCTG) Plasmid remnants, Tn6049 inserted in a gene involved in plasmid mobilisation (encoding a MobA relaxase)
CMGI-7	0317–0333	335,757–351,119 (15,362)	Integrase (disrupted by Tn6049), direct repeats (include 46 bp of 5' end tRNA-Met, TTGGTTGCGGGGCAGGACTTGAACCTGGACCTCGGGTTATGAGC) No defined function for accessory genes, additional copy of Tn6049
CMGI-8	2549–2561	2,773,806–2,786,063 (12,257)	Integrase, no DR Integrase is part of a four-gene structure from which the second is disrupted by Tn6049
CMGI-9	2156–2172	2,357,123–2,377,771 (20,648)	Arsenic resistance Integrase, direct repeats (16 bp of 3' end tRNA-Arg, CCTGTCCGGCGGGACCA)
CMGI-10	3347–3368	3,624,125–3,645,072 (20,947)	No defined function for accessory genes, Tn6049, IS1087 Integrase, no direct repeats No defined function for accessory genes
CMGI-11	1660–1668	1,794,838–1,805,662 (10,824)	No integrase, no direct repeats, three IS elements No integrase, no direct repeats, flanked by two IS elements (ISCme7) Putative fimbrial operon

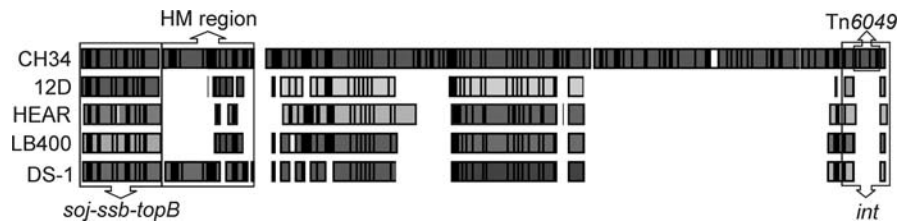


Fig. 1 The 109 kb genomic island CMGI-1 on chromosome 1 of strain CH34 and synteny with *R. picketti* 12D, *H. arsenicoxydans* HEAR, *B. xenovorans* LB400 and *P. lavamentivorans* DS-1 (HM: genes involved in heavy metal resistance). Only conserved gene clusters or synteny groups

are shown and highlighted in different shades of gray. These groups are defined as clusters of orthologous genes (with a minimum threshold of 35% protein sequence identity on 80% of the length of the smallest protein)

islands consists of conserved (with many hypothetical genes up to now refractory to annotation) and variable modules with various putative functions. One conserved module at the end of the genomic island (opposite the *int* gene) seemingly contained replicon maintenance genes such as *soj* (plasmid partitioning), *ssb* (single strand DNA binding), and *topB* (DNA topoisomerase).

Importantly, nosocomial infections and environmental strains from industrial sites share in common this large piece of mobile genetic material, which suggests a recent horizontal gene transfer, and also implies that the barriers between (industrial) environmental reservoirs and clinical environments can be crossed. Acquiring genes by horizontal transfer may thus be enhanced significantly by potential co-inhabitation as well as by genetic relatedness. Mathee et al. (2008) showed that 42% of the open reading frames (ORFs) in the core genome of *P. aeruginosa* had best hits outside the genus *Pseudomonas* and these hits belonged to the genus *Azotobacter*, a closely related γ -Proteobacteria, and *Burkholderia* and *Cupriavidus*, two β -Proteobacteria. Strains of the β -Proteobacterial genera *Cupriavidus*, *Ralstonia*, and *Burkholderia* have been identified in cystic fibrosis patients. Similarly, *P. aeruginosa* and *C. metallidurans* may coexist in industrial sites with high levels of heavy metals.

Metal-response and resistance genes in CMGI-1 and PAGI-2C

CMGI-1 contains metal-resistance genes supporting the recognizable functions carried by this genomic island. However, the particular selective pressures involved in acquiring the island (in *Cupriavidus* or in *Pseudomonas*) are unknown, as are its selective

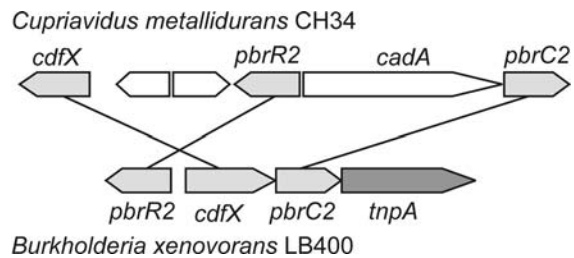


Fig. 2 Local synteny between CMGI-1 of CH34 and *B. xenovorans* LB400. For LB400 the structure consists of a four-gene based transposon with eight copies in the genome of strain *B. xenovorans* LB400. Orthologous genes are connected with a line (with a minimum threshold of 35% protein sequence identity on 80% of the length of the smallest protein). Genes that are shown encode for a cation diffusion factor (*cdfX*), a MerR family regulator (*pbrR2*), an ATPase (*cadA*), a lipoprotein/peptidase (*pbrC2*), and a transposase (*tnpA*)

advantage. Figure 2 shows the gene *cdfX* (Rmet_2299) encoding for a putative metal-efflux system that is distantly related to cation diffusion factors, such as CzcD (Anton et al. 1999) and CnrT, and the *pbrR2 cadA pbrB2* genes (Rmet_2302–2304). The latter resemble three genes in the lead resistance operon, *pbr*, of pMOL30 (Borremans et al. 2001), which, respectively encode for a MerR family regulator (Chen et al. 2005, 2007), a metal-efflux ATPase (for lead, zinc, cadmium; Monchy et al. 2006) and a lipoprotein/peptidase. Microarray data (Gene Expression Omnibus accession number GSE7272) revealed the overexpression of *cdfX*, *pbrR2*, *cadA*, and *pbrC2* after metallic challenge, implying a synergy between CdfX and CadA for metal detoxification from the cytoplasm to the periplasm, although little is known about CdfX. Other accessory CMGI-1 genes include mercury resistance genes (the mercury reductase is likely not operational), efflux ATPases, and the *ccm* and *coo*

genes involved in the biosynthesis of major cytochromes.

New transposons carrying metal response genes found in CMGI-1/PAGI-2C

Syntenic analysis of various β - and γ -Proteobacteria identified the presence of a locus similar to that found in CMGI-1/PAGI-2C but in which *cdfX* replaced *cadA* forming *pbrR2 cdfX pbrC2*. Next, to this conserved gene arrangement a transposase (ISL3 family) was located and this complete four-gene set was identified in *Pseudomonas putida* plasmids as IS*Ppu12* (Li et al. 2004; Weightman et al. 2002; Williams et al. 2002) and in *B. xenovorans* LB400 (Chain et al. 2006). The latter contained eight copies of this transposon, registered here as Tn6052, with one disrupted by Tn6053, a Tn6049-equivalent in *B. xenovorans* LB400 (see below). In *B. xenovorans* LB400 this transposon apparently neighbors catabolic functions. Furthermore, two slightly modified copies were identified in *Idiomarina loihiensis* L2TR (a marine bacterium found in a deep sea hydrothermal vent near Hawaii; Hou et al. 2004) and two almost identical copies were identified in *Acinetobacter baumannii* AYE. The phenotype associated with this mobile genetic element is undefined, and at least for *Pseudomonas putida* mt-2 the wild-type strain carrying IS*Ppu12* on plasmid pWW0 and plasmid-free derivatives (no IS*Ppu12*) showed no differences in resistance to Hg, Zn, Cd, Cu, and Pb (P.A. Williams, personal communication). However, this could perhaps point to the origin of the element since many β -Proteobacterial determinants of heavy metal resistance are not or only slightly expressed in γ -Proteobacteria.

Finally, this listing of transposons clearly extends the set of metal-response genes associated to transposons, now supposedly mainly populated by mercury- and arsenic resistance genes. This example adequately illustrates the power of synteny for detecting new mobile genetic elements outside the bias of a selectable marker, similar to other examples described below (see CMGI-4).

CMGI-2 and CMGI-3: genomic islands of the Tn4371 family: a major contribution to strain CH34, chemolithotrophy and catabolism of aromatic compounds

Cupriavidus metallidurans CH34 contains islands belonging to the Tn4371 family. The first identified

member of this family, Tn4371 from the strain *C. oxalaticus* A5 (formerly *R. eutropha* A5), was described as a catabolic transposon carrying genes that degraded biphenyl- and chlorobiphenyl-compounds (Merlin et al. 1997, 1999). Soon similar transposons were found in other bacteria including *R. solanacearum* (but not in *Cupriavidus* strains H16 or JMP134; Toussaint et al. 2003).

Analysis showed that Tn4371-like elements comprised a large family with a quadripartite structure corresponding to an integrated conjugative plasmid with basic genes (*parAB*, *repA*, *traF*, *virD2*, *traR*, and *traG*), the transfer genes (*trb*), accessory genes systematically present between *virD2* and *traR*, and a generally small *int* module at the *parB* side. Previous descriptions of the Tn4371 family considered the *int* gene as part of the *parB-virD2* module but synteny studies indicated there were variable, poorly conserved *int* modules within the Tn4371 islands. Undoubtedly, the whole family is defined by a quadripartite structure, consisting of an integrase module, a region with plasmid/phage/genomic island maintenance genes, the accessory genes, and the conjugative genes. These conjugative genes strongly resembled those of the broad-host-range plasmids Ti/IncP1, and while the transfer regions of these genomic islands were conserved, the *parAB virD2* sections were much more variable. At the extremity of this sector (with the maintenance genes), lied an *int* gene that likely directed the integration of the conjugative plasmid and controls further replicon-based movements of the corresponding genomic island. Mobility of the Tn4371 family of genomic islands was illustrated by Tn4371 itself: in *C. oxalaticus* A5, the biphenyl genomic island could easily move into the broad-host-range IncP1 plasmid RP4, and after transfer into *C. metallidurans* CH34, could stably integrate into its genome. Annotation of the *C. metallidurans* CH34 genome revealed the presence of at least two complete genetic islands belonging to the Tn4371 family: CMGI-2 and CMGI-3. In both cases, *int-traF-virD2* and *trb* modules flank the variable sector with accessory genes (Fig. 3).

CMGI-2, primarily a “catabolic transposon”

CMGI-2 contained 25 genes involved in toluene degradation, underlying CH34's ability to grow using toluene, benzene, or xylene as the sole carbon source (Table 1; Fig. 3). Besides the biphenyl transposon

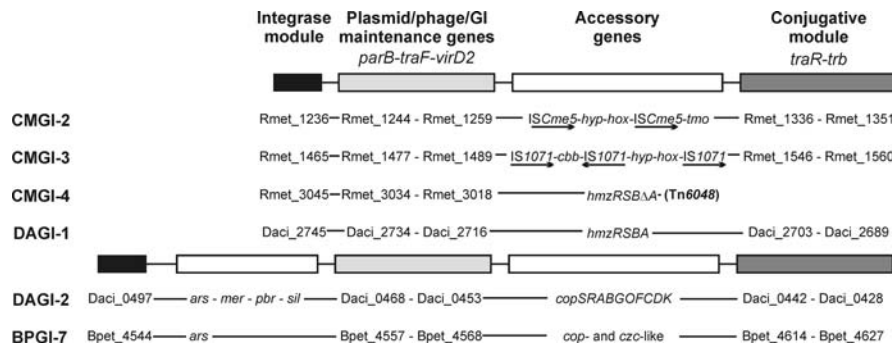


Fig. 3 Quadripartite structure of the Tn4371-family of conjugative transposons, consisting of an integrase module, a region with plasmid/phage/genomic island maintenance genes, the

accessory genes, and the conjugative genes. CMGI-2, CMGI-3 and CMGI-4: *C. metallidurans* CH34 GIs; DAGI-1 and DAGI-2: *D. acidovorans* SPH-1; BPGI-7: *Bordetella petrii* DSM12804

Tn4371 (Toussaint et al. 2003), CMGI-2 is the second catabolic transposon of this family. In addition, other accessory genes were inserted near the catabolic cluster including the hydrogenases involved in chemolithotrophy that were flanked by two copies of IS*Cme5* (four copies in the complete genome), and a bloc of mostly IS elements or genes encoding integrases/tyrosine-based site-specific recombinases with some integrases organized in “trios” (see section below). Both IS*Cme5* copies (IS481 family) in CMGI-2 occurred in a region with multiple transposases and fragments thereof; probably, they were the driving force integrating the *hyp/hox* cluster modulating chemolithotrophy in CMGI-2. An extensive inventory and analysis of the IS elements in CH34 is beyond the scope of this study.

CMGI-3 and chemolithotrophy

CMGI-3, the second Tn4371-family genomic island in CH34, shelters accessory genes involved in

chemolithotrophy between *virD2* and *traR* (Table 1; Fig. 3). They constitute two adjacent blocs flanked by IS1071: one contained genes for fixing CO₂ (*cbb* genes coding for ribulose phosphate kinase and ribulose biphosphate carboxylase), and the other harbored genes for chemolithotrophy (*hyp* and *hox* genes coding for hydrogenases). In fact, all the genes for aerobic chemolithotrophic growth on CO₂ and H₂ via two kinds of hydrogenases (one cytosolic and one membrane-bound, respectively encoded by the *hox* and *hyp* clusters) were located on CMGI-2 and CMGI-3. Together with the numerous genes located on the four replicons involved in the response to heavy metals, they participate in the survival of strain CH34 in mineral niches. The closely related bacteria *C. eutrophus* H16 also grows chemolithotrophically, but the essential genes are organized differently. Table 2 summarizes these organizational differences between H16 and CH34.

Table 2 Organization of the genes involved in chemolithotrophy

	<i>C. eutrophus</i> H16	<i>C. metallidurans</i> CH34
Main support	Plasmid pHG1 (452 kb Schwartz et al. 2003; family pRALTA, pMOL28, pHG1)	Genomic islands CMGI-2 and -3 (Tn4371 family) inserted in chromosome 1
The <i>cbb</i> genes (CO ₂ fixation)	Duplicated on pHG1 and chromosome 2 Pohlmann et al. (2006)	Only one set of Rubisco genes flanked by IS1071 (opposite orientations) in CMGI-3 Much more similar to <i>Nitrosomonas</i> and purple photosynthetic bacteria than to H16
Hydrogenases genes	Strong homology <i>hox</i> , <i>hyp1</i> with CH34, set of <i>hyp2</i> genes	No <i>hyp2</i> genes: replaced by duplication of <i>hyp1</i>
Site-specific recombinases (mainly <i>int</i>)	Hot spots of <i>int/xerD</i> deletions/remnants in pHG1	<i>int/xerD</i> genes organized in “trios” (RIT elements) in CMGI-2 and highly syntenic with a corresponding region of pHG1 (storage of remnants?)

plasmid pMOL28 has a gene core strongly identical to that of pHG1 and later recruited genes involved in the resistance to heavy metals.

This picture of gene mobility again illustrates the bacterial versatility in using the “mobile toolbox” to recruit genes, allowing the colonization of specific niches. Probably the hydrogenotrophy genes in CMGI-3 were recruited via *IS1071* elements. Vice versa, the capability of strain CH34 for autotrophic growth can be lost by *IS1071*-mediated deletion. *IS1071* (Tn3 family) was also detected in several catabolic strains (Peel and Wyndham 1999; Providenti et al. 2006; Wyndham et al. 1994). In all cases, the sequence of *IS1071* was extremely conserved and the dissemination of *IS1071* apparently had infectious traits. This analysis suggests that the hydrogenotrophy determinants were recruited recently, as was the colonization by *C. metallidurans* strains of industrial sites that dates from the modern industrial development. *IS1071* also occurred in plasmid pMOL28 where it also participated to recruit a genomic island carrying determinants for heavy metal resistance (Mergeay et al. 2009; Monchy et al. 2007).

Chromosomally encoded mobilizing capacity in C. metallidurans CH34: a possible intervention of CMGI-2 or CMGI-3

The presence of transfer genes, exhibiting strong similarity to those of the broad-host-range plasmids Ti/IncP1, on the Tn4371-like elements CMGI-2 and CMGI-3 might explain previous observations of horizontal gene transfer between derivatives of *C. metallidurans* CH34. They included the transfer of RSF1010 based plasmids (IncQ, Mob⁺) and pLAFR3 (IncP, Mob⁺) from CH34 and plasmid-free derivative strains at a frequency of 10⁻⁶ or 10⁻⁵ transconjugants per recipient to *C. metallidurans* or *E. coli* without the presence of a Tra⁺ helper plasmid, which implies the presence of functionally active transfer functions on the *C. metallidurans* CH34 chromosome (Taghavi 1996; S. Taghavi and D. van der Lelie, personal communication).

CMGI-4: a large composite genomic island with metal response genes and a full Tn4371-like counterpart in *Delftia acidovorans* SPH-1:

Analysis of CMGI-4 revealed a puzzling observation. Like CMGI-2 and CMGI-3 it contained the *int-parB-*

traF-virD2 region, but lacks a *trb* region with plasmid conjugative genes (Table 1; Fig. 3). Yet synteny studies demonstrated that the putative island would extend much further than the *virD2* gene, suggesting that accessory genes were present. Indeed, three successive blocks were observed: seven genes including an *hmz* gene cluster, the transposon Tn6048 (Van der Auwera et al. 2009), and a more mosaic region. The first block includes a putative ferritin reductase, and a *hmzBA hmzRS* cluster similar to the *czc* metal-efflux genes of the RND family. Genes encoding for HmzB and HmzA are canonical paralogs, respectively of the membrane fusion protein-encoding *czcB* and the efflux pump-encoding *czcA*. The cation pump HmzA was previously allocated to the HME3b family (Nies et al. 2006). However, the *czcC* equivalent, encoding for the outer membrane protein, is absent and is replaced by two small open reading frames. Next to this small *hmz* gene cluster lied Tn6048, a twelve-gene transposon with four genes for transposon maintenance and mobility, and eight genes apparently responding to heavy metals since they were induced by zinc and lead according to microarray data. The last block of eleven genes adjacent to Tn6048 include genes for a truncated transposase, a cluster involved in breaking down phosphite, another IS element, and three genes involved in methionine biosynthesis. This cluster has a 100% identical counterpart in *Delftia acidovorans* strain SPH-1 (Denger et al. 2008; Rosch et al. 2008; Schleheck et al. 2004a). However, strain SPH-1 carries an untruncated transposase (Daci_2679), which is part of a Tn3-related transposon, registered here as Tn6051, together with a resolvase (Daci_2680), and putative transport genes. This element occurred almost 100% conserved in *Pelobacter propionicus* DSM 2379 (3 copies) and *Marinobacter aquaeolei* VT8 (1 copy). The gene cluster involved in phosphite metabolism (Rmet_2992 to Rmet_2996) is syntenic with a region in *Herminiimonas arsenicoxydans* (HEAR3211-3215) (Muller et al. 2006, 2007) where this cluster is flanked by two IS elements (IS30 family), thus defining a potential composite transposon. Furthermore, this complete element is strictly conserved in *Janthinobacterium* sp. Marseille (mma_3039 to mma_3045; Audic et al. 2007). In both species, the neighboring genes apparently modulate heavy metal resistance. Considering precipitation of metal-phosphates, oxidizing phosphite to phosphate could be involved in the resistance to heavy metals. However, microarray data

do not support the induction of these genes under metallic challenge. Instead, constitutive expression, or a slight down regulation are observed. At the moment, the participation of CMGI-4 to the metabolism of CH34 is not clear.

In *D. acidovorans* SPH-1, transposon Tn6051 is close to a complete Tn4371-like island that could be called DAGI-1 (Fig. 3). From *hmzA* (Daci_2709) to *intA* (Daci_2745), DAGI-1 is almost identical to CMGI-4 but it also contains a full *traG-trb* conjugative module like CMGI-2 and CMGI-3 as well as a few accessory genes adjacent to *hmzA* in particular Daci_2708 that is transcriptionally coupled with *hmzA*. The loss of this gene in CH34 may explain the weak expression of *hmzA* in CH34. Therefore, it seems that two events occurred in an undefined order: a transfer of the whole region from Daci_2745 to Daci_2669, and a deletion going from part of Daci_2679 to Daci_2709. This deletion (unspecified in which bacterium it occurred) would have removed some accessory genes, the conjugative module of DAGI-1 and most of Tn6051 in CH34.

Further examination of CMGI-4 and DAGI-1 showed the latter has a full counterpart, CTGI-1, in the β -Proteobacteria *Comamonas testosteroni* KF-1 (Rosch et al. 2008; Schleheck et al. 2004a). Both *D. acidovorans* and *C. testosteroni* were earlier considered to belong to the same genus (Denger et al. 2008; Wen et al. 1999). Furthermore, strains KF-1 and SPH-1 were isolated from the same environment, and their fully sequenced genomes were the first ones to be sequenced in their corresponding genera. The only difference with CTGI-1 is that DAGI-1 contains a four-gene insertion (with two putative integrases) in the *parB* module that might be a mobile element because it also appears elsewhere in the genome of *D. acidovorans* SPH-1.

Other Tn4371-family genomic islands carrying metal resistance genes as revealed by syntenic analysis

DAGI-2 in D. acidovorans SPH1

The *D. acidovorans* SPH1 genome surprisingly contained a new Tn4371-like genomic island with all the accessory genes for metal resistance. Indeed, DAGI-2 carried five clusters of resistance genes, to copper, silver/copper, cadmium/lead, mercury, and arsenic,

with *cop* genes sited between the *parAB* and the *traR-trb* module similar to the biphenyl genes in the original Tn4371. The clusters, *silABCDR*, *pbrRcadApbrC*, *merRTPA*, and *arsRIC2BC1H* were flanked by the *parAB* cluster and the *int* gene (Fig. 3). All these genes have counterparts in strain CH34: *sil* on plasmid pMOL30, *pbr* and *mer* in genomic island CMGI-1, and *ars* in CMGI-7. The assembly of the ten *cop* genes is remarkable: *copK*, *copDC*, *copGFO*, and *copBARS*. Some genes are more closely related to the chromosomal *cop* genes, while others are similar to the plasmid pMOL30 version. DAGI-2 also contains a *czcD/dmeF*-like gene (Anton et al. 1999; Munkelt et al. 2004) inserted with a strongly conserved companion gene in the *parAB* module. Several bacteria contain the *czcD*-like gene (Daci_0464) and its companion gene (Daci_0465). In DAGI-2, this gene pair resides in a module (partition/maintenance module) of the Tn4371 family of genomic islands that normally does not shelter accessory genes (Toussaint et al. 2003). The presence of this *czcD*-like gene would add a supplementary tool to the diversified arsenal of metal resistance genes carried by DAGI-2 (Fig. 3).

BPGI-7 in the soil bacterium Bordetella petrii DSM12804

Bordetella petrii strain DSM12804 was initially isolated from river sediment (Gross et al. 2008). Unlike other *Bordetella* species, it is a facultative anaerobe, not known to be associated with humans or other warm-blooded animals. Analysis of its genome (a 5.3 Mb unique replicon) revealed seven acquired genomic islands (Gross et al. 2008). Although one of them, BPGI-4, was similar to Tn4371 (Gross et al. 2008), the genome clearly contains two Tn4371-like genomic islands displaying a complete quadripartite structure (*int*, *par*, accessory genes and *trb* modules). The second Tn4371-like genomic island is actually BPGI-7 (Bpet4544-Bpet4630) and carries *ars* genes in the *int* module as in *D. acidovorans*, as well as *cop* and *czc*-like genes in the accessory module inserted between *virD2* and *traR* (Fig. 3).

Other small genomic islands on the chromosome: CMGI-5 to -11

The characteristics of these small islands correspond to the main criteria discussed above: lack of synteny,

divergence from the current gene model, location near a tRNA gene, presence of an *int* gene (tyrosine-based site-specific recombinase), high content of hypothetical genes (often small ones), and presence of IS and/or transposons. Their contribution to the metabolism of *C. metallidurans* is not apparent, yet each has a distinctive feature of interest in understanding the diversity of genomic islands. Their characteristics are summarized in Table 1 and their location on chromosome 1 is shown in Fig. 5 (as for other islands).

CMGI-5: a plasmid remnant?

This island comprising 24 ORFs (from Rmet_2847 to Rmet_2824) begins with an *int* gene and is flanked at the other extremity by a tRNA gene. It contains copies of IS1086 (Dong et al. 1992; Taghavi et al. 1997) and Tn6049 (Mergeay et al. 2009). Besides the hypothetical genes, some typical plasmid-related genes are present, such as *repA*, *traY*, *mobA* (likely knocked out by Tn6049), and *mobB*. The *int* gene is coupled transcriptionally with an exonuclease gene that putatively could exert a resolvase-like function. This coupling of *int* genes with a second gene was also observed in CMGI-4, pMOL28 (*bimAB*) and pMOL30 (*bimAB*), but the second gene is very variable and poorly defined. The *repA* gene displays some identity with orthologs of various broad-host-range plasmids like pSa (IncW), and mainly pSB102 (Schneiker et al. 2001), pIPO2T, and pMOL98 (all from the PromA family; Mela et al. 2008; Van der Auwera et al. 2009). Other plasmid-related genes of CMGI-5 have orthologs in different plasmids from various origins, pointing to mosaicism. The limited size of this putative integrated plasmid in CMGI-5 suggests that it probably is a remnant of a larger and more complex genomic island.

CMGI-6, a cryptic genomic island sheltering Tn6049

This island (Rmet_1997 to Rmet_2020) is located near a tRNA-Met gene. It contains two copies of Tn6049, genes coding for hypothetical proteins (mostly small sizes), and a truncated gene encoding for a putative major phage capsid protein. The latter is one of the rare *C. metallidurans* CH34 genes encoding for a bacteriophage structural protein. The terminal *int* gene is disrupted by one of the Tn6049

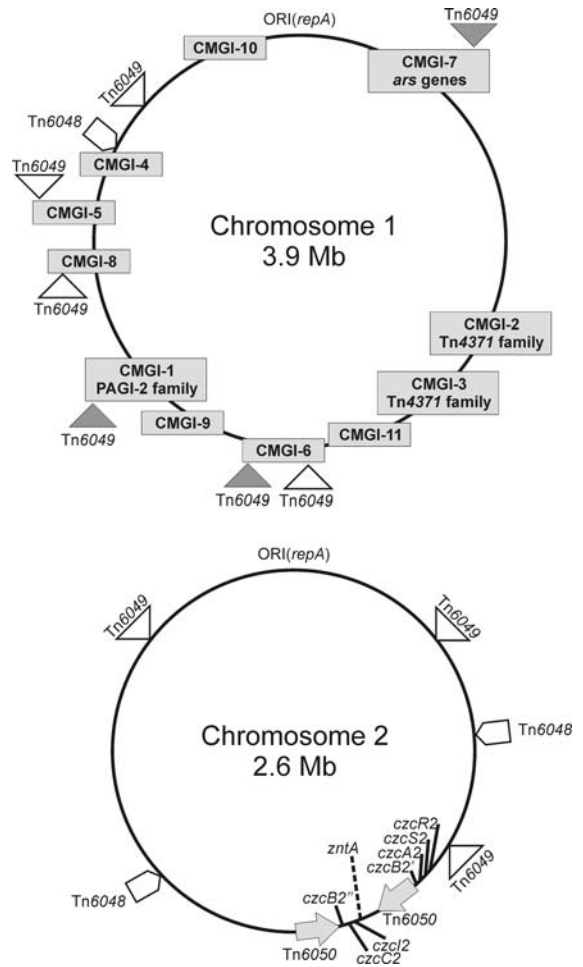


Fig. 5 Schematic presentation (not to scale) indicating the location of the genomic islands CMGI-1 to -11 and the transposons Tn6048, Tn6049, and Tn6050. For Tn6049, filled triangles represent inactivation of an *int* gene

elements and putatively immobilized the island (Fig. 5). Finally, next to the island a copy of IS1087 (IS3/IS911 family) (Collard et al. 1993; Tibazarwa et al. 2000) was recognized.

CMGI-7: an arsenic gene island

Features characterizing this small genomic island are lack of synteny, loss of current gene model, presence of resistance genes, presence of an integrase trio, and one Tn6049 inactivating the central gene of this trio; its extremity, formed by Rmet_0333, is not well-defined. Seven genes (from Rmet_0327 to Rmet_0333) are involved in resistance to arsenic (Zhang et al. 2009) and are highly overexpressed by

arsenate and bismuth exposure and by zinc and lead. Further, expression of the regulatory gene *arsR* is up-regulated by Cd, Hg, and Au. The extremity at the *int* side seems to correspond to a four-gene structure, including a putative trio of recombinases, and is similar to one extremity of CMGI-28b on pMOL28 (further described below). The second CMGI-7 gene is interrupted by a Tn6049, which likely inactivated the mobility as in many other *C. metallidurans* CH34 mobile genetic elements (Fig. 5).

CMGI-8 is also a cryptic genomic island sheltering IS1087 and Tn6049

This small island (13 genes from Rmet_2549 to Rmet_2561) is defined by the presence of a tRNA-Leu gene at one extremity. Besides the *int* gene, there is, in the following order, a gene encoding for a putative serine-based recombinase, a copy of IS1087 and a copy of Tn6049. Other genes present are mostly very small and hypothetical. The linkage between an *int* gene and a *tnpR*-like (resolvase) gene, also observed in other bacteria such as in *P. putida* W619, may correspond to a defined mobility pattern.

CMGI-9 shares features with the int module of CMGI-3

This genomic island with 17 ORFs (from Rmet_2156 to Rmet_2172) fulfills most criteria (i.e., location by a tRNA gene, lack of synteny, divergence from the current gene model, *int* gene at one extremity). The absence of IS elements or copies of Tn6049 suggests that it might still be mobile. Most CMGI-9 genes are hypothetical, except for a four-gene operon that is well conserved through the plasmids of *Deinococcus radiodurans*, α -, β - and γ -Proteobacteria, as well as in the *int* module of CMGI-3. This operon may be involved in repairing or degrading DNA, since it contains genes encoding, respectively for, a nucleoside triphosphate hydrolase, and an endodesoxyribonuclease.

CMGI-10

This island with 22 ORFs (Rmet_3347 to Rmet_3368) was identified mainly by divergence from the gene model and lack of synteny. Besides a gene encoding for an ATPase and hypothetical genes,

it contains three IS elements (ISCme3, ISCme1 and ISCme4). It does not have an *int* gene.

CMGI-11: an island flanked by IS and carrying genes for fimbrial biosynthesis

This small element (9 ORFs) contains an operon of five genes encoding proteins involved in fimbriae biosynthesis. Only one similar structure was found in closely related genomes, namely on the 301 kb megaplasmid pBMC401 of *Burkholderia ambifaria* MC40-6, a soil-borne strain associated with maize roots, where the gene cluster also is flanked by IS elements. However, at least 16 counterparts were uncovered in genomes of γ -Proteobacteria (mostly enterics). Acquiring an extra set of genes involved in the biosynthesis of fimbriae or similar surface components might be relevant to adaptation to harsh soil conditions. The island is flanked by 2 copies of ISCme7 that differ by three base pairs and their transposase genes ended at the stop codon present in the flanking sequences. ISCme7 belongs to the IS6 family. In Firmicutes and γ -Proteobacteria, IS6-related IS appear to be often present as arrays (up to 5 in a 20 kb fragment) flanking captured genes involved in resistance to antibiotics, encoding for toxin/antitoxin, and for pili biosynthesis. These duplicated elements are regularly deleted, probably to fix the captured genes.

Genomic islands of chromosome 2

The structure of the second chromosome in *Cupriavidus*, *Ralstonia*, and *Burkholderia* strains differs markedly from that of the first chromosome. The structure of chromosome 2 is a patchwork, so discarding lack of synteny and divergence from the gene model as significant parameters to define a genomic island. In addition, chromosome 2 contains fewer tRNA genes, only eight compared to 54 in chromosome 1, and very few *xerD/int* genes encoding for tyrosine-based site-specific recombinases. The relative high number of genes involved in the response to heavy metals compared to chromosome 1 and 2 of related *Cupriavidus* and *Ralstonia* species might indicate that their acquisition event occurred long ago. The acclimatization/integration process of these acquired genes might even be completed to the

point of eliminating all clear traces of the lateral acquisition. These regions certainly meet the broader terminology of regions of genomic plasticity but care should be taken when assigning them as genomic islands.

Genomic islands in plasmids pMOL28 and pMOL30 and metal resistance genes

Cupriavidus/Ralstonia strains often carry one or more megaplasmids with very specific traits seemingly linked to niche specialization. For strain CH34, both megaplasmids pMOL28 and pMOL30 carry genetic determinants conferring high resistance to a wide diversity of heavy metals (Mergeay et al. 2009; Monchy et al. 2007). The plasmid backbones and the genomic islands on pMOL28 and pMOL30 carrying these determinants were recently detailed (Mergeay et al. 2009; Monchy et al. 2007).

Plasmid pMOL28 carries three genomic islands. Island CMGI-28a containing all the heavy metal resistance genes of pMOL28, is flanked by IS elements (Tn3 family) inactivated either by partial deletion or by insertion of Tn6049. Island CMGI-28b encompasses at one extremity three *rhs*-like genes of unknown function that code for products rich in tyrosine-aspartate motifs. Finally, CMGI-28c sited adjacent to CMGI-28a, contains mostly hypothetical genes: their proximity suggests that these two might constitute one single genomic island.

In plasmid pMOL30, most genes involved in the response or resistance to heavy metals occur in two large genomic islands, CMGI-30a with the *czc*, *pbr*, and *mer* genes, and CMGI-30b with the *sil* and *cop* genes. Partially deleted IS elements or transposons exist at one extremity of both islands. Furthermore, both contain “nested islands” flanked by (partial) tyrosine-based site-specific recombinases, which delimit specific gene clusters.

New transposons in the genome of *C. metallidurans* CH34

The new transposons reported during this study, both in strain CH34 and other strains, were all registered according to the new revised nomenclature for transposable genetic elements (Roberts et al. 2008).

The transposons on the chromosomes of CH34 are shown in Fig. 5.

Tn6048, containing 8 genes highly induced by zinc and lead

This transposon present at three copies (one in chromosome 1 and two in chromosome 2) contains twelve genes. It has a module of four genes involved in mobility, and eight multiple metal phenotype (*mmf*) genes; although the function of the latter is unknown, they apparently participate in the response to heavy metals as they are strongly induced when zinc and lead are present. The four mobility genes share some similarity with Tn7 genes. Interestingly, thallium- and cesium-resistant clones from a CH34 cosmid library contained one of the Tn6048 copies located on chromosome 2, but this phenotype looks to be linked to genes neighboring Tn6048.

Tn6049: a role in stabilizing genomic islands?

This four-gene-based transposon of 3,517 bp was recently described in plasmids pMOL28 and pMOL30 as TnCme2 (Mergeay et al. 2009). It contains the *tnmB*, *gspA*, *tnmA*, and *tnmC* genes but the functions of *tnmB* and *tnmC* are unknown. The *tnmA* gene codes for a Tn7-like transposase. While *gspA* initially was annotated as an equivalent of a major gene encoding for an ATPase acting in a type II-secretion system, additional analyses via the ACLAME database revealed its possible equivalents in mutator phage B3 and suggested that its gene product shares characteristics with the MuB protein (A. Toussaint, personal communication). The *gspA* gene of Tn6049 therefore was re-annotated as a putative activator of transposition. The genome of strain CH34 has twelve copies of Tn6049. Similar four-gene structure were observed in *B. xenovorans* LB400 (i.e., two identical copies of Tn6053 with around 60% protein identity with their corresponding Tn6049 orthologs, except for *tnmC* where the 72 amino acid gene product of the *B. xenovorans* transposon shares only 24% identity with TnmC of CH34), in *B. phymatum* STMB15 (one copy), in *B. phytofirmans* PsJN (6 copies) and in *B. petrii* (one copy with an insertion in the *gspA*-equivalent). The identities between these *Burkholderia* and *Bordetella*

Tn6049-like transposons are much higher (up to 100% for *tnmC*) than with Tn6049.

Tn6049 appears to play the peculiar role of “stabilizing” genomic islands in strain CH34. Table 3 shows the insertion sites of this transposon. Tn6049 inactivates the terminal *int* gene of CMGI-1, CMGI-7, and CMGI-6. Tn6049 is present in CMGI-5, lies at the extremity of the putative island CMGI-28a and inserted inside IS1071. The latter is a very active IS element often observed in catabolic plasmids. Again, we hypothesize that the inactivation of IS1071 by Tn6049 helps to stabilize the genomic island rich in genes involved in responding to heavy metals.

Tn6050 in chromosome 2: peculiar passenger genes

There are two copies of this peculiar transposon in chromosome 2 lying in opposite orientations and separated by 143 kb. The core (*tnpA tnpR*) is very similar to that of the mercury transposons Tn4378 and Tn4380 located, respectively, on plasmid pMOL28 and pMOL30. As for Tn4378 and Tn4380, Tn6050 is bound by 38 bp inverted repeats (IR) and clearly belongs to the Tn501/Tn3 family. However, the accessory genes are not involved in mercury resistance and encode proteins not generally associated with transposons, namely a sulfate permease, a universal stress protein (UspA), and a DksA-like DnaK suppressor protein. Their role is

unclear but we assume functional associations since these accessory genes (or a part of it) are found in other Tn3-related transposons. In these transposons they are either in combination with antibiotic resistance determinants like for Tn6001, Tn1403, and Tn1404* derived from clinical and environmental *Pseudomonas* strains (Stokes et al. 2007; Tseng et al. 2007), or occur alone like in Tn1013 from the IncP-1 α plasmid pBS228 (Haines et al. 2007).

Closer scrutiny of the chromosome 2 regions flanking the Tn6050 elements revealed the formation of a chromosomal inversion by recombination between the pair of Tn6050 transposons (Fig. 5). The first copy is located upstream, a partial heavy metal cation tri-component efflux system (*czc*-like), comprising a two-component transcriptional regulator *czcR2* (Rmet_4465) and a sensor *czcS2* (Rmet_4466), the efflux pump *czcA2* (Rmet_4468), and the membrane fusion encoding *czcB2'* (Rmet_4469). The latter is truncated by Tn6050, and its apparent 5' part *czcB2''* (Rmet_4597) is found downstream of the second copy of Tn6050 (oriented in the opposite direction). Further downstream of *czcB2''*, the genes *czcC2* and *czcI2* are found, encoding, respectively, for an outer membrane porin, and for a regulator, thus complementing the heavy metal cation tri-complex efflux system (Fig. 5). Additional evidence for this inversion is found through analysis of the direct repeats generated by the transposition of the two Tn6050 copies.

Table 3 Insertion sites Tn6049

Inserted between Rmet_	Location	Gene product inactivated by insertion
0318 and 0323 (fragments)	Chromosome 1	Integrase of CMGI-7
1997 and 2002 (fragments)	Chromosome 1	Integrase of CMGI-6
2006 and 2011	Chromosome 1	NA, located in CMGI-6
2403 and 2408 (fragments)	Chromosome 1	Integrase of CMGI-1
2549 and 2554	Chromosome 1	NA, located in CMGI-8
2834 and 2839	Chromosome 1	Putative relaxase, CMGI-5
3139 and 3144 (fragments)	Chromosome 1	Outer membrane porin
3928 and 3933	Chromosome 2	Bug (extra-cytoplasmic solute receptor)
4421 and 4426 (fragments)	Chromosome 2	Putative MmgE/PrpD family protein
5477 and 5482 (fragments)	Chromosome 2	Gene for conserved hypothetical protein (equivalent of pRALTA_0107)
6333 and 6338 (fragments)	pMOL28	IS1071, extremity of CMGI-28a
5944 and 6180 (fragments)	pMOL30	PbrU, major facilitator permease in CMGI-30a

NA not applicable

Integrases and site-specific recombinases associated with genomic islands

The number of genes (complete or partial) annotated as transposases (with a catalytic site bearing non-adjacent DDE residues) or as integrases, more accurately as tyrosine-based site-specific recombinases, in the genome of *C. metallidurans* CH34 is very striking, especially compared to those in the closely related genomes of *C. eutrophus* H16, *C. pinatubonensis* JMP134 and *C. taiwanensis* LMG19424. The web-interfaced IslandPath software (<http://www.pathogenomics.sfu.ca/islandpath/>) clearly allowed to visualize this difference on the chromosomes. Furthermore, the number of *int/xerD*-like genes (many partial ones) attracts attention, as well as the previously unnoticed diversity of gene combinations around the *int/xerD* genes in association with the genomic islands of *C. metallidurans* CH34. The most notable associations were described here, but indications of their functional significance are lacking. The classification of recombinases and their role in genomic islands is not straightforward because of the confusion between transposases, phage integrases catalytic site, and integrases sensu stricto caused by automatic annotations. Hence, we adopted the denomination, tyrosine-based site-specific recombinase, with *int* most frequently used. A new consensus is needed for denominating *int* genes of genomic islands, especially when they are associated with a transcriptionally or closely linked accompanying gene, or when multiple putative tyrosine-based site-specific recombinases are in tandem. Lastly, we have few clues about the extent to which some of these *int* genes are linked to the well-known integrases of the sensu stricto integrons (Mazel 2006). In this descriptive section, we focus attention on the diversity of *int* genes in *C. metallidurans* CH34 and some related bacteria, without making mechanistic implications.

Integrases of genomic islands (chromosome 1 and plasmids pMOL28 and pMOL30)

When associated with a genomic island, most of the putative *int* genes are at a terminal position, like for chromosomal islands CMGI-1 to CMGI-9, and for the islands CMGI-28b and CMGI-28c on plasmid pMOL28. Other plasmid-nested islands are flanked at

both sides by partially deleted genes encoding for tyrosine-based site-specific recombinases like CMGI-28a on pMOL28 carrying the *bimA1* and *chr* (chromate resistance) genes, CMGI-30a on pMOL30 carrying the *bimA2* and *czc* (cadmium, zinc, cobalt resistance), and CMGI-30b on pMOL30 carrying the *cop* (copper resistance) genes. Among the exceptions are the *rit* genes (recombinase in trio or triad) observed in CMGI-2 and described below. Contrarily, the putative *int/xerD* genes of chromosome 2 are syntenic (Rmet_5821, Rmet_3675 and less clear for Rmet_4514) or partial with no clear clues for a genomic island (Rmet_4534 and Rmet_4556).

rit genes: site-specific recombinases in triads, a poorly noticed, but common element, except in γ -Proteobacteria

Two groups of *int*-like genes in CMGI-2 that are annotated via ACLAME as tyrosine-based site-specific recombinases are defined as trios with 3 bp overlaps. These elements are denominated here as RIT elements, with RIT representing recombinase in trio or triad. One trio (*ritA2B2C2*; Rmet_1239 to Rmet_1241) belongs to the integrase module of CMGI-2 and inactivates a *radC*-like gene (putatively involved in DNA repair). A second trio (*ritA1B1C1*; Rmet_1271 to Rmet_1273) belongs to the complex accessory module of CMGI-2 where it also inactivated a gene, in this case a *mpA* gene of the IS66 family. Thus, both trios seem to behave as mobile genetic elements. The protein identity between them is low but both share higher protein identities with other trios, some being sheltered in bacteria taxonomically very distant. These RIT elements apparently occur in various β -Proteobacteria, such as *Aromatoleum aromaticum* EbN1 and *Polaromonas* sp. JS66, in plasmid pHG1 of *C. eutrophus* H16, and in several α -Proteobacteria, for example, *Mesorhizobium loti* MAFF303099, *Sinorhizobium meliloti* 1021 and *S. medicae* WSM419, *Dinoroseobacter shibae* DFL12, and *Caulobacter* sp. K31. In these two latter species, there are two identical copies of the RIT element, quite distant on the genome, again recalling transposon-like behavior. Furthermore, the RIT elements with substantial protein identities are found in at least one Firmicute of the genus *Heliobacterium*, and in various Actinobacteria such as *Bifidobacterium longum* NCC2705, *Rhodococcus jostii* RHA1,

Frankia sp. EAN1pec, and *Mycobacterium vanbaalenii* PYR-1. Notably, so far, no *rit* trio was detected via synteny in γ -Proteobacteria. We propose to name these elements in a way similar to the nomenclature of transposons: RITCme1 and RITCme2 for *C. metallidurans* CH34, RITMme1 for *Mesorhizobium meliloti*, and so on.

Inside the RIT elements, the recombinases display very low protein identity, but exhibit higher identities with *int* genes of other trios sharing the same position in the trio/triad. Seemingly, each member of the element evolved for a specific role in the mobility behavior of the RIT element. Their lengths are not the same; the central *rit* gene sometimes is shorter. To understand their role in horizontal gene transfer, and bacterial biology, a systematic study is needed, comparing all existing RIT elements and conducting in vivo experiments. A preliminary examination of RIT targets showed that RIT elements lie mainly near transposase genes or remnants, and near *radC*-like genes in at least two other (very different) bacteria.

For the *C. metallidurans* CH34 *rit* genes from CMGI-2 (the genomic island involved in chemolithotrophy that shares various genes with pHG1), the RITCme2 element (Rmet_1239 to Rmet_1241) displays the highest protein identity with an orthologous element in *B. petrii* while the RITCme1 element (Rmet_1271 to Rmet_1273) displays very high homology with its various pHG1 counterparts (Table 2; Schwartz et al. 2003).

rit genes in plasmid pHG1 (*C. eutrophus* H16):
a collection of RIT elements

Plasmid pHG1 contains 21 *rit*-like genes or fragments encoding for putative tyrosine-based site-specific recombinases, mostly organized in trios and linked by 3 bp overlaps. These pHG1 counterparts of *rit* genes lie in two clusters: from PHG034 to PHG048 (with 14 *rit*-like genes or (very small) fragments); and, from PHG129 to PHG141 with seven *rit* genes. In the first region, three full RIT elements perhaps orthologous to RITCme1 are recognizable, but each seems to have a rearrangement (short deletions in PHG035 and PHG044, gene fusion in PHG042; Fig. 6). This region is located not far downstream from the *hox* genes that are involved in the biosynthesis of the soluble hydrogenase, and that are also flanked on both sides by 2–4 remnants of transposase

genes. The second region PHG129 to PHG141 contains two highly identical homologs (more than 92%) of *ritA1B1* (of RITCme1) associated with a more distant *ritC*, and contains a homolog of *ritC1* (of RITCme1) associated to another element RIT element. This region also is flanked by fragments of an IS66-like *tnp* (transposase) as observed for RITCme1 in CMGI-2 in *C. metallidurans* CH34, again intimating possible affinity for specific genes as targets.

At least four varieties of RIT elements seem to be represented in both pHG1 *rit* clusters. It is not clear if one or more of them are functional or if these clusters act as storage of remnants, as it could be the case for pMOL30.

Between both *rit* clusters of pHG1, there also is an isolated RIT element (PHG056 to PHG058) that has orthologs in a variety of bacteria, and especially with the three identical copies of a *B. petrii* RIT-like element that are flanked by fragments of non-identical DNA-cytosine methyl-transferase genes (Gross et al. 2008). Again, this arrangement suggests that the RIT elements may recognize some genes as preferential targets.

Both CMGI-2 and pHG1 have in common RIT elements and genes involved in chemolithotrophy but the role of RIT in the mobility of the *hyp* and *hox* genes clusters remains unclear since other mobile elements are associated with these gene clusters.

CMGI-28b of pMOL28: a four-gene element
with a regulatory gene

Genomic island CMGI-28b, located on pMOL28, mainly contains *rhs* genes (Mergeay et al. 2009). CMGI-28b contains a combination of four closely linked genes including three putative ones encoding for *int*-like site-specific recombinases (Rmet_6252 to Rmet_6254) and a regulatory gene (Rmet_6255) from to the TetR family of repressors. This four-gene structure recalls the beginning of the small *ars* element (CMGI-7) for which the second gene is knocked out by Tn6049. Therefore, these four-gene structures might play a role in the mobility of some genetic elements. They do not display any similarity with RIT elements. Both have counterparts in various bacteria, including some γ -Proteobacteria such as *Shewanella frigidimarina* NCIMB400 and *Vibrio cholerae* O1 biovar *el tor* str. N16961 for the four-

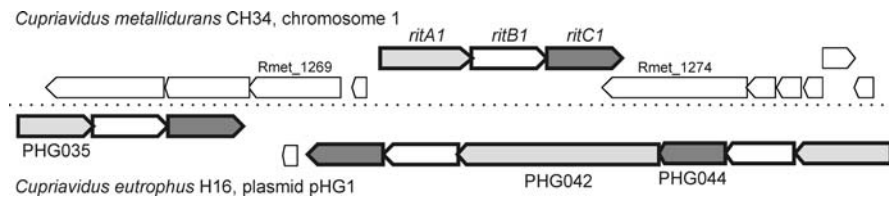


Fig. 6 Synteny between *RitCme1*, located on chromosome 1 of CH34, and plasmid pHG1 of *C.eutrophus* H16. Orthologous genes are similarly highlighted (with a minimum threshold of 35% protein sequence identity on 80% of the length of the smallest protein)

gene structure contrary to RIT elements. It is not known if these counterparts are located in the extremities of mobile genetic elements in other bacteria, as it is the case for CMGI-28b and CMGI-7.

bimA1B1, chromate genes in *pMOL28*
and a chromate island in *B. vietnamiensis*

Throughout this study, we found that the position of *int* genes gave valuable pointers to detect genes from the horizontal gene pool. In plasmid pMOL28, the tyrosine-based site-specific recombinase *bimA* gene (Rmet_6193), and its companion *bimB* (a putative repressor gene) lie close to the mercury resistance transposon, Tn4378. This pair is separated from the *chr* genes by five genes coding for hypothetical proteins with unknown function or phenotypic clues. The *chr* genes, in turn, are followed by the *cnr* genes. Synteny analysis shows that the first two of these five genes, *orf44* (p10049) and *lppY* (p10050), are systematically associated with chromate genes in bacteria as different as *Burkholderia pseudomallei*, *Methylobacterium* sp.4-46, and *Arthrobacter* sp. FB24. Microarray analysis revealed that the next three genes preceding the chromate cluster *chrFE-CAB1* are among those most induced in the presence of chromate. Rmet_6196 encodes a permease of the Major Facilitator Superfamily and will be called *chrP*, while Rmet_6197 and Rmet_6198 encode unknown proteins are denominated *chrN* and *chrO*.

These observations suggest that more previously unnoticed genes must be taken into consideration in detailing the response to or the resistance to chromate in *C. metallidurans* CH34 and other bacteria.

Similarly, the *bimAB* module is present in *B. vietnamiensis* *G4* (O'Sullivan et al. 2007) at four identical copies together with at least 12 genes including *lppY* and the chromate genes *chrNOFE-CAB*. These *bimAB*-chromate structures make two

pairs, one pair wherein one putative Tn3-related element is inserted (located, respectively, on G4 chromosome 1 and plasmid pBVIE01). For the second pair, two putative Tn3-related elements are integrated (located, respectively, on G4 chromosome 3 and plasmid pBVIE03). Again, a more detailed, systematic analysis and *in vivo* experiments are needed to clarify the role of these *bimAB* modules in horizontal gene transfer. Nevertheless, there are two noteworthy observations. First is the putative relation between Tn3-related elements and *bimAB* modules, such as the insertions of Tn3-related elements in the four *bimAB* structures in *B. vietnamiensis* *G4*, the Tn3-family mercury resistance transposons (complete or partial) near *bimA1B1* on pMOL28 and *bimA2B2* on pMOL30 of strain CH34, and the Tn3-related element close to a *bimAB* module in plasmid 2 of *Polaromonas* sp. JS66. Second is the neighboring heavy metal resistance determinants, such as those for chromate in CH34 and *B. vietnamiensis* *G4*, for copper and mercury resistance in *Ralstonia pickettii* 12J (also CH34 if mercury resistance transposons are considered) and for arsenate/arsenite resistance in *Burkholderia multivorans* ATCC17616.

General discussion

Cupriavidus metallidurans strain CH34 belongs to a species mainly defined by multiple metal-resistant (and often hydrogenotrophic) strains isolated from industrial sites mainly linked to mining-, metallurgical-, and chemical-industries. The sites are characterized by high levels of bio-available heavy metals, often above the maximum tolerated concentrations. The sequenced genome of *C. metallidurans* CH34 shows much similarity to the main chromosome (chromosome 1) of the corresponding replicons of *C. eutrophus* H16, *C. taiwanensis* LMG19424, and

C. pinatubonensis JMP134 (Sato et al. 2006). Chromosome 2 is more of a patchwork, as it also was observed in various *Burkholderia* species (Fricke et al. 2009) and contains a variety of heavy metal resistance genes that are not evident in other *Cupriavidus* strains (but many are present in *R. solanacearum*; Mergeay et al. 2003). Niche specialization seems to be linked to the presence of megaplasmids. At first glance, plasmids/megaplasmids seem to gather most of the horizontal gene pool required for niche specialization: this applies to *C. taiwanensis* LMG19424, *C. eutrophus* H16, and likely to *C. pinatubonensis* JMP134.

However, a glance at the genomic maps as displayed by IslandPath (<http://www.pathogenomics.sfu.ca/islandpath/>) shows a striking difference between *C. metallidurans* and the other related genomes due to the much higher abundance of ORFs (complete or fragments) related to transposases or site-specific recombinase genes. For example, in chromosome 1 of *C. metallidurans* CH34 there are 69 ORFs compared with 9 (5 fragments, 3 *bona fide* IS elements and 1 *bona fide int* gene) for *C. taiwanensis* LMG19424 (Amadou et al. 2008).

The differences between *C. taiwanensis* LMG19424 and *C. metallidurans* CH34 may be attributed to their respective environments, namely the symbiotic nodules of legumes, and industrial polluted soils. The nodules might be considered as protected from external microbial influences, or extensively changing physicochemical pressures. From an evolutionary perspective, the industrial sites where the metal resistant *Cupriavidus* strains occur are recent, their advent mainly arose after the industrial revolution a few centuries ago (a few millennia ago if we consider earlier metallurgy). As a species, *C. metallidurans* has undergone part of its constituent evolution quite recently, relying on horizontal gene transfer with recognizable and still active tools. Here, a valuable finding is the presence of mobile genetic elements that are conserved fully between taxonomically distant species. There are examples of full conservation in strain CH34. The full conservation of CMGI-1 between CH34 and *P. aeruginosa* clone C strains, of CMGI-4 between CH34 and *D. acidovorans* SPH-1 as reported here, of ISC_{me3} between CH34 and pHG1 (data not shown), and of IS1071 between CH34 and various environmental bacteria able to degrade aromatic xenobiotic

compounds (Di Gioia et al. 1998; Nakatsu et al. 1991; Peel and Wyndham 1999; Providenti et al. 2006; Wyndham et al. 1994). Possibly, to obtain IS1071 the progenitors of *C. metallidurans* might have coexisted with xenobiotic degraders in polluted soils (many of these strains being β -Proteobacteria).

The rearrangements found in the IS1071 copy on pMOL28 might function by stabilizing the captured genes: this concept may be applicable in the many genetic islands of CH34 (CMGI-1, CMGI-5 to 8, CMGI-28a) and especially the transposon Tn6049 seems to play a major role in this inactivation of genes involved in mobility (mainly, but not exclusively, site-specific recombinase genes). Inactivation of the genes involved in mobility is not only observed by insertion of Tn6049, but also by insertion of IS elements or through accumulation of point mutations or deletions in *int* genes and IS elements. For instance, in pMOL30 deleted tyrosine-based site-specific recombinases flank the *cop* genes (Mergeay et al. 2009; Monchy et al. 2007). All these observed events suggest primary acquisition via a mobile genetic element and successive stabilization or domestication via the rearrangement of the key mobility genes.

A curious feature possibly linked to this process is the clustering of *tnp* and *int* gene fragments in some plasmid regions of pMOL30 and even in chromosomal genomic islands. It was also observed in pHG1 of *C. eutrophus* H16 (Fig. 4) and in pRALTA of *C. taiwanensis* that contains up to 120 ORFs linked to IS elements. All these deleted or rearranged mobility genes may represent scars of events occurring at various periods in the evolutionary history of these strains, and likely encompassing a long time up to rather recent events. An interesting question is if whether these clusters would dissolve completely in time or would they be a kind of storage of deleted or “eroded” mobility genes after usage?

The genome as a source of unknown or poorly noticed mobile genetic elements

The main islands found in the chromosome of *C. metallidurans* CH34 belong to the KLFC2/PAGI-2 family (CMGI-1) (Klockgether et al. 2006; Larbig et al. 2002) or to the Tn4371-family (CMGI-2, CMGI-3, and CMGI-4). Some small islands of interest are the arsenic island (CMGI-7) that expands our view on

heavy metal responses, and CMGI-9 that shares characteristics with the *int* module of CMGI-3. We described new transposons of special interest because they carry genes that are never, or rarely represented in the current transposon collection from clinical- or environmental-studies such as the eight still undefined genes of Tn6048 for which no phenotype is recorded but that are induced by heavy metals, and the unusual but conserved association of Tn6050 genes encoding for an universal stress protein (UspA), a suppressor of DnaK and a permease of sulfate. The latter one is a good example of passenger genes from well disseminated mobile genetic elements for which we do not yet have good phenotypic clues.

Maybe the advantageous power of synteny in annotation studies is the ability to reveal the diversity of mobile genetic elements outside the bias of selection not only in the studied genome but also in a variety of other genomes. We can recall here numerous cases like transposon Tn6053 of *B. xenovorans* LB400 very similar in structure to Tn6049, the four-gene transposon Tn6052 (*pbrR cdfX pbrC tnpA*) similar to IS*Ppu12* (from *P. putida* plasmids) and found in very distant bacteria, a genetic element carrying chromate genes in *B. vietnamiensis* G4 and particularly the Tn4371 family genomic islands of *D. acidovorans* and *B. petrii* carrying various metal responding genes. In this respect, the *D. acidovorans* DAGI-1 genomic island is remarkable as a carrier of metal resistance genes that are gathered in a compact collection: *cop* genes in the accessory module, *czcD*-like in the replication module and *sil*, *pbr*, *mer*, and *ars* genes in the integrase module. Annotation using synteny allows also putting in evidence the great diversity in tyrosine-based site-specific recombinase genes (integrases) going beyond the diversity already known for the bacteriophage and the integron integrases. Among these genes, there are the *int* genes of the genomic islands of the PAGI-2 family and of the integrase modules of the Tn4371 family, the well conserved four-gene elements found in CMGI-7 (*ars*) and CMGI-28b, and the peculiar RIT elements (triads of site-specific recombinase genes with 3 bp overlaps) that are present in CMGI-2 and especially in pHG1. Moreover, the latter are present in many bacteria and especially in Actinobacterial strains belonging to the genera *Frankia*, *Mycobacterium* and *Rhodococcus*.

At this stage, it may be useful to mention that antibiotic resistance genes are practically absent of the mobile genetic elements reported in this study. The diversity of the mobile genetic elements so found is representative of this part of the horizontal genetic pool that was at work when β -Proteobacteria (genera *Cupriavidus*, *Ralstonia*, *Burkholderia*, *Delftia*, *Bordetella*, *Herminiimonas*, and *Janthinobacterium*) began to colonize anthropogenic sites (soils, sediments) characteristic of the modern industrial revolution. The annotation of *C. metallidurans* CH34 allows then a new and refreshing insight on the diversity of bacterial mobile genetic elements.

Acknowledgments This work was supported by the European Space Agency (ESA-PRODEX) and the Belgian Science Policy (Belspo) through the MISSEX project (PRODEX agreements C90254) and by the collaboration agreement between SCK•CEN and Université Libre of Bruxelles. SM was supported by a SCK•CEN Ph.D. grant. Thanks are due to Claudine Médigue, Zoé Rouy, David Vallenet from the MAGE Genoscope platform of annotation for their availability and their kind pedagogy, to the SCK•CEN annotation team and especially Paul Janssen, Pieter Monsieurs, Hugo Moors and Nicolas Morin, to Niels van der Lelie and Safieh Taghavi for 20 years collaboration about the genetics of *C. metallidurans* CH34, to Ariane Toussaint and Jacques Mahillon for comments and suggestions about the mobile genetic elements and the introduction to the ACLAME and ISfinder databases, and to Simon Silver for patience and comprehension.

References

- Amadou C, Pascal G, Mangenot S et al (2008) Genome sequence of the beta-rhizobium *Cupriavidus taiwanensis* and comparative genomics of rhizobia. *Genome Res* 18:1472–1483. doi:10.1101/gr.076448.108
- Anton A, Grosse C, Reissmann J, Pribyl T, Nies DH (1999) CzcD is a heavy metal ion transporter involved in regulation of heavy metal resistance in *Ralstonia* sp. strain CH34. *J Bacteriol* 181:6876–6881
- Audic S, Robert C, Campagna B, Parinello H, Claverie JM, Raoult D, Drancourt M (2007) Genome analysis of *Minibacterium massiliensis* highlights the convergent evolution of water-living bacteria. *PLoS Genet* 3:e138. doi:10.1371/journal.pgen.0030138
- Borremans B, Hobman JL, Provoost A, Brown NL, van Der Lelie D (2001) Cloning and functional analysis of the *pbr* lead resistance determinant of *Ralstonia metallidurans* CH34. *J Bacteriol* 183:5651–5658. doi:10.1128/JB.183.19.5651-5658.2001
- Brim H, Heyndrickx M, de Vos P, Wilmette A, Springael D, Schlegel H, Mergeay M (1999) Amplified rDNA restriction analysis and further genotypic characterisation of metal-resistant soil bacteria and related facultative hydrogenotrophs. *Syst Appl Microbiol* 22:258–268

- Chain PS, Deneff VJ, Konstantinidis KT et al (2006) *Burkholderia xenovorans* LB400 harbors a multi-replicon, 9.73-Mbp genome shaped for versatility. *Proc Natl Acad Sci USA* 103:15280–15287. doi:10.1073/pnas.0606924103
- Chen P, Greenberg B, Taghavi S, Romano C, van der Lelie D, He C (2005) An exceptionally selective lead(II)-regulatory protein from *Ralstonia metallidurans*: development of a fluorescent lead(II) probe. *Angew Chem Int Ed Engl* 44:2715–2719. doi:10.1002/anie.200462443
- Chen PR, Wasinger EC, Zhao J, van der Lelie D, Chen LX, He C (2007) Spectroscopic insights into lead(II) coordination by the selective lead(II)-binding protein PbrR691. *J Am Chem Soc* 129:12350–12351. doi:10.1021/ja0733890
- Collard JM, Provoost A, Taghavi S, Mergeay M (1993) A new type of *Alcaligenes eutrophus* CH34 zinc resistance generated by mutations affecting regulation of the *cnr* cobalt-nickel resistance system. *J Bacteriol* 175:779–784
- Denger K, Weinitschke S, Smits TH, Schleheck D, Cook AM (2008) Bacterial sulfite dehydrogenases in organotrophic metabolism: separation and identification in *Cupriavidus necator* H16 and in *Delftia acidovorans* SPH-1. *Microbiology* 154:256–263. doi:10.1099/mic.0.2007/011650-0
- Di Gioia D, Peel M, Fava F, Wyndham RC (1998) Structures of homologous composite transposons carrying *cbaABC* genes from Europe and North America. *Appl Environ Microbiol* 64:1940–1946
- Diels L, Mergeay M (1990) DNA probe-mediated detection of resistant bacteria from soils highly polluted by heavy metals. *Appl Environ Microbiol* 56:1485–1491
- Dong Q, Sadouk A, van der Lelie D, Taghavi S, Ferhat A, Nuyten JM, Borremans B, Mergeay M, Toussaint A (1992) Cloning and sequencing of IS1086, an *Alcaligenes eutrophus* insertion element related to IS30 and IS4351. *J Bacteriol* 174:8133–8138
- Fricke WF, Kusian B, Bowien B (2009) The genome organization of *Ralstonia eutropha* strain H16 and related species of the *Burkholderiaceae*. *J Mol Microbiol Biotechnol* 16:124–135. doi:10.1159/000142899
- Goris J, De Vos P, Coenye T, Hoste B, Janssens D, Brim H, Diels L, Mergeay M, Kersters K, Vandamme P (2001) Classification of metal-resistant bacteria from industrial biotopes as *Ralstonia campinensis* sp. nov., *Ralstonia metallidurans* sp. nov. and *Ralstonia basilensis*. *Int J Syst Evol Microbiol* 51:1773–1782
- Gross R, Guzman CA, Sebahia M et al (2008) The missing link: *Bordetella petrii* is endowed with both the metabolic versatility of environmental bacteria and virulence traits of pathogenic *Bordetellae*. *BMC Genomics* 9:449. doi:10.1186/1471-2164-9-449
- Hacker J, Hochhut B, Middendorf B, Schneider G, Buchrieser C, Gottschalk G, Dobrindt U (2004) Pathogenomics of mobile genetic elements of toxigenic bacteria. *Int J Med Microbiol* 293:453–461. doi:10.1078/1438-4221-00290
- Haines AS, Jones K, Batt SM, Kosheleva IA, Thomas CM (2007) Sequence of plasmid pBS228 and reconstruction of the IncP-1 α phylogeny. *Plasmid* 58:76–83. doi:10.1016/j.plasmid.2007.01.001
- Hou S, Saw JH, Lee KS et al (2004) Genome sequence of the deep-sea gamma-proteobacterium *Idiomarina loihiensis* reveals amino acid fermentation as a source of carbon and energy. *Proc Natl Acad Sci USA* 101:18036–18041. doi:10.1073/pnas.0407638102
- Hsiao W, Wan I, Jones SJ, Brinkman FS (2003) IslandPath: aiding detection of genomic islands in prokaryotes. *Bioinformatics* 19:418–420. doi:10.1093/bioinformatics/btg004
- Klockgether J, Wurdemann D, Reva O, Wiehlmann L, Tummler B (2006) Diversity of the abundant pKLC102/PAGI-2 family of genomic islands in *Pseudomonas aeruginosa*. *J Bacteriol* 189:2443–2459. doi:10.1128/JB.01688-06
- Klockgether J, Wurdemann D, Wiehlmann L, Tummler B (2008) Transcript profiling of the *Pseudomonas aeruginosa* genomic islands PAGI-2 and pKLC102. *Microbiology* 154:1599–1604. doi:10.1099/mic.0.2007/014340-0
- Larbig KD, Christmann A, Johann A, Klockgether J, Hartsch T, Merkl R, Wiehlmann L, Fritz HJ, Tummler B (2002) Gene islands integrated into tRNA(Gly) genes confer genome diversity on a *Pseudomonas aeruginosa* clone. *J Bacteriol* 184:6665–6680. doi:10.1128/JB.184.23.6665-6680.2002
- Leplae R, Hebrant A, Wodak SJ, Toussaint A (2004) ACLAME: a classification of mobile genetic elements. *Nucleic Acids Res* 32:D45–D49. doi:10.1093/nar/gkh084
- Leplae R, Lima-Mendez G, Toussaint A (2006) A first global analysis of plasmid encoded proteins in the ACLAME database. *FEMS Microbiol Rev* 30:980–994. doi:10.1111/j.1574-6976.2006.00044.x
- Li W, Shi J, Wang X, Han Y, Tong W, Ma L, Liu B, Cai B (2004) Complete nucleotide sequence and organization of the naphthalene catabolic plasmid pND6-1 from *Pseudomonas* sp. strain ND6. *Gene* 336:231–240. doi:10.1016/j.gene.2004.03.027
- Mathee K, Narasimhan G, Valdes C et al (2008) Dynamics of *Pseudomonas aeruginosa* genome evolution. *Proc Natl Acad Sci USA* 105:3100–3105. doi:10.1073/pnas.0711982105
- Mazel D (2006) Integrons: agents of bacterial evolution. *Nat Rev Microbiol* 4:608–620. doi:10.1038/nrmicro1462
- Mela F, Fritsche K, Boersma H, van Elsas JD, Bartels D, Meyer F, de Boer W, van Veen JA, Leveau JH (2008) Comparative genomics of the pIPO2/pSB102 family of environmental plasmids: sequence, evolution, and ecology of pTer331 isolated from *Collimonas fungivorans* Ter331. *FEMS Microbiol Ecol* 66:45–62. doi:10.1111/j.1574-6941.2008.00472.x
- Mergeay M (1991) Towards an understanding of the genetics of bacterial metal resistance. *Trends Biotechnol* 9:17–24. doi:10.1016/0167-7799(91)90007-5
- Mergeay M (2000) Bacteria adapted to industrial biotopes: the metal resistant *Ralstonia*. In: Storz G, Hengge-Aronis R (eds) *Bacterial stress responses*. ASM Press, Washington DC, pp 403–414
- Mergeay M, Houba C, Gerits J (1978) Extrachromosomal inheritance controlling resistance to cadmium, cobalt, copper and zinc ions: evidence from curing in a *Pseudomonas*. *Arch Int Physiol Biochim* 86:440–442
- Mergeay M, Nies D, Schlegel HG, Gerits J, Charles P, Van Gijsegem F (1985) *Alcaligenes eutrophus* CH34 is a facultative chemolithotroph with plasmid-bound resistance to heavy metals. *J Bacteriol* 162:328–334

- Mergeay M, Monchy S, Vallaëys T, Auquier V, Benotmane A, Bertin P, Taghavi S, Dunn J, van der Lelie D, Wattiez R (2003) *Ralstonia metallidurans*, a bacterium specifically adapted to toxic metals: towards a catalogue of metal-responsive genes. *FEMS Microbiol Rev* 27:385–410. doi:10.1016/S0168-6445(03)00045-7
- Mergeay M, Monchy S, Janssen P, Van Houdt R, Leys N (2009) Megaplasmids in *Cupriavidus* genus and metal resistance. In: Schwartz E (ed) *Microbial megaplasmids*. Springer, Berlin, pp 209–238
- Merlin C, Springael D, Mergeay M, Toussaint A (1997) Organisation of the *bph* gene cluster of transposon Tn4371, encoding enzymes for the degradation of biphenyl and 4-chlorobiphenyl compounds. *Mol Gen Genet* 253:499–506. doi:10.1007/s004380050349
- Merlin C, Springael D, Toussaint A (1999) Tn4371: a modular structure encoding a phage-like integrase, a *Pseudomonas*-like catabolic pathway, and RP4/Ti-like transfer functions. *Plasmid* 41:40–54. doi:10.1006/plas.1998.1375
- Monchy S, Vallaëys T, Bossus A, Mergeay M (2006) Metal efflux P1-ATPase genes of *Cupriavidus metallidurans* CH34: a transcriptomic approach. *Int J Environ Anal Chem* 86:677–692. doi:10.1080/03067310600583824
- Monchy S, Benotmane MA, Janssen P, Vallaëys T, Taghavi S, van der Lelie D, Mergeay M (2007) Plasmids pMOL28 and pMOL30 of *Cupriavidus metallidurans* are specialized in the maximal viable response to heavy metals. *J Bacteriol* 189:7417–7425. doi:10.1128/JB.00375-07
- Muller D, Simeonova DD, Riegel P, Mangenot S, Koechler S, Lievreumont D, Bertin PN, Lett MC (2006) *Herminiimonas arsenicoxydans* sp. nov., a metalloresistant bacterium. *Int J Syst Evol Microbiol* 56:1765–1769. doi:10.1099/ijs.0.64308-0
- Muller D, Medigue C, Koechler S et al (2007) A tale of two oxidation states: bacterial colonization of arsenic-rich environments. *PLoS Genet* 3:e53. doi:10.1371/journal.pgen.0030053
- Munkelt D, Grass G, Nies DH (2004) The chromosomally encoded cation diffusion facilitator proteins DmeF and FieF from *Wautersia metallidurans* CH34 are transporters of broad metal specificity. *J Bacteriol* 186:8036–8043. doi:10.1128/JB.186.23.8036-8043.2004
- Nakatsu C, Ng J, Singh R, Straus N, Wyndham C (1991) Chlorobenzoate catabolic transposon Tn5271 is a composite class I element with flanking class II insertion sequences. *Proc Natl Acad Sci USA* 88:8312–8316. doi:10.1073/pnas.88.19.8312
- Nies D, Mergeay M, Friedrich B, Schlegel HG (1987) Cloning of plasmid genes encoding resistance to cadmium, zinc, and cobalt in *Alcaligenes eutrophus* CH34. *J Bacteriol* 169:4865–4868
- Nies A, Nies DH, Silver S (1989) Cloning and expression of plasmid genes encoding resistances to chromate and cobalt in *Alcaligenes eutrophus*. *J Bacteriol* 171:5065–5070
- Nies DH, Rehbein G, Hoffmann T, Baumann C, Grosse C (2006) Paralogs of genes encoding metal resistance proteins in *Cupriavidus metallidurans* strain CH34. *J Mol Microbiol Biotechnol* 11:82–93. doi:10.1159/000092820
- O’Sullivan LA, Weightman AJ, Jones TH, Marchbank AM, Tiedje JM, Mahenthalingam E (2007) Identifying the genetic basis of ecologically and biotechnologically useful functions of the bacterium *Burkholderia vietnamiensis*. *Environ Microbiol* 9:1017–1034. doi:10.1111/j.1462-2920.2006.01228.x
- Peel MC, Wyndham RC (1999) Selection of *clc*, *cba*, and *feb* chlorobenzoate-catabolic genotypes from groundwater and surface waters adjacent to the Hyde park, Niagara Falls, chemical landfill. *Appl Environ Microbiol* 65:1627–1635
- Pohlmann A, Fricke WF, Reinecke F et al (2006) Genome sequence of the bioplastic-producing “Knallgas” bacterium *Ralstonia eutropha* H16. *Nat Biotechnol* 24:1257–1262. doi:10.1038/nbt1244
- Providenti MA, Shaye RE, Lynes KD, McKenna NT, O’Brien JM, Rosolen S, Wyndham RC, Lambert IB (2006) The locus coding for the 3-nitrobenzoate dioxygenase of *Comamonas* sp. strain JS46 is flanked by IS1071 elements and is subject to deletion and inversion events. *Appl Environ Microbiol* 72:2651–2660. doi:10.1128/AEM.72.4.2651-2660.2006
- Roberts AP, Chandler M, Courvalin P et al (2008) Revised nomenclature for transposable genetic elements. *Plasmid* 60:167–173. doi:10.1016/j.plasmid.2008.08.001
- Rosch V, Denger K, Schleheck D, Smits TH, Cook AM (2008) Different bacterial strategies to degrade taurocholate. *Arch Microbiol* 190:11–18. doi:10.1007/s00203-008-0357-7
- Salanoubat M, Genin S, Artiguenave F et al (2002) Genome sequence of the plant pathogen *Ralstonia solanacearum*. *Nature* 415:497–502. doi:10.1038/415497a
- Sato Y, Nishihara H, Yoshida M, Watanabe M, Rondal JD, Concepcion RN, Ohta H (2006) *Cupriavidus pinatubonensis* sp. nov. and *Cupriavidus laharis* sp. nov., novel hydrogen-oxidizing, facultatively chemolithotrophic bacteria isolated from volcanic mudflow deposits from Mt. Pinatubo in the Philippines. *Int J Syst Evol Microbiol* 56:973–978. doi:10.1099/ijs.0.63922-0
- Schleheck D, Knepper TP, Fischer K, Cook AM (2004a) Mineralization of individual congeners of linear alkylbenzenesulfonate by defined pairs of heterotrophic bacteria. *Appl Environ Microbiol* 70:4053–4063. doi:10.1128/AEM.70.7.4053-4063.2004
- Schleheck D, Tindall BJ, Rossello-Mora R, Cook AM (2004b) *Parvibaculum lavamentivorans* gen. nov., sp. nov., a novel heterotroph that initiates catabolism of linear alkylbenzenesulfonate. *Int J Syst Evol Microbiol* 54:1489–1497. doi:10.1099/ijs.0.03020-0
- Schleheck D, Knepper TP, Eichhorn P, Cook AM (2007) *Parvibaculum lavamentivorans* DS-1T degrades centrally substituted congeners of commercial linear alkylbenzenesulfonate to sulfophenyl carboxylates and sulfophenyl dicarboxylates. *Appl Environ Microbiol* 73:4725–4732. doi:10.1128/AEM.00632-07
- Schneiker S, Keller M, Droge M, Lanka E, Puhler A, Selbitschka W (2001) The genetic organization and evolution of the broad host range mercury resistance plasmid pSB102 isolated from a microbial population residing in the rhizosphere of alfalfa. *Nucleic Acids Res* 29:5169–5181. doi:10.1093/nar/29.24.5169
- Schwartz E, Henne A, Cramm R, Eitinger T, Friedrich B, Gottschalk G (2003) Complete nucleotide sequence of

- pHG1: a *Ralstonia eutropha* H16 megaplasmid encoding key enzymes of H₂-based lithoautotrophy and anaerobiosis. *J Mol Biol* 332:369–383. doi:[10.1016/S0022-2836\(03\)00894-5](https://doi.org/10.1016/S0022-2836(03)00894-5)
- Siguier P, Perochon J, Lestrade L, Mahillon J, Chandler M (2006) ISfinder: the reference centre for bacterial insertion sequences. *Nucleic Acids Res* 34:D32–D36. doi:[10.1093/nar/gkj014](https://doi.org/10.1093/nar/gkj014)
- Stokes HW, Elbourne LD, Hall RM (2007) Tn1403, a multiple-antibiotic resistance transposon made up of three distinct transposons. *Antimicrob Agents Chemother* 51:1827–1829. doi:[10.1128/AAC.01279-06](https://doi.org/10.1128/AAC.01279-06)
- Taghavi S (1996) Un mégaplasmide de résistance aux métaux lourds d'*Alcaligenes eutrophus*: analyse génétique et fonctionnelle. Ph.D. Thesis, Université Libre de Bruxelles, Brussels, Belgium
- Taghavi S, Mergeay M, van der Lelie D (1997) Genetic and physical maps of the *Alcaligenes eutrophus* CH34 megaplasmid pMOL28 and its derivative pMOL50 obtained after temperature-induced mutagenesis and mortality. *Plasmid* 37:22–34. doi:[10.1006/plas.1996.1274](https://doi.org/10.1006/plas.1996.1274)
- Tibazarwa C, Wuertz S, Mergeay M, Wyns L, van Der Lelie D (2000) Regulation of the *cnr* cobalt and nickel resistance determinant of *Ralstonia eutropha* (*Alcaligenes eutrophus*) CH34. *J Bacteriol* 182:1399–1409. doi:[10.1128/JB.182.5.1399-1409.2000](https://doi.org/10.1128/JB.182.5.1399-1409.2000)
- Toussaint A, Merlin C, Monchy S, Benotmane MA, Leplae R, Mergeay M, Springael D (2003) The biphenyl- and 4-chlorobiphenyl-catabolic transposon Tn4371, a member of a new family of genomic islands related to IncP and Ti plasmids. *Appl Environ Microbiol* 69:4837–4845. doi:[10.1128/AEM.69.8.4837-4845.2003](https://doi.org/10.1128/AEM.69.8.4837-4845.2003)
- Tseng SP, Hsueh PR, Tsai JC, Teng LJ (2007) Tn6001, a transposon-like element containing the *bla*_{VIM-3}-harboring integron In450. *Antimicrob Agents Chemother* 51:4187–4190. doi:[10.1128/AAC.00542-07](https://doi.org/10.1128/AAC.00542-07)
- Vallenet D, Labarre L, Rouy Z, Barbe V, Bocs S, Cruveiller S, Lajus A, Pascal G, Scarpelli C, Medigue C (2006) MaGe: a microbial genome annotation system supported by synteny results. *Nucleic Acids Res* 34:53–65. doi:[10.1093/nar/gkj406](https://doi.org/10.1093/nar/gkj406)
- Van der Auwera GA, Król JE, Suzuki H, Foster B, Van Houdt R, Brown CJ, Mergeay M, Top EM (2009) Plasmids captured in *C. metallidurans* CH34: defining the PromA family of broad-host-range plasmids. *Antonie van Leeuwenhoek*. doi:[10.1007/s10482-009-9316-9](https://doi.org/10.1007/s10482-009-9316-9)
- Vandamme P, Coenye T (2004) Taxonomy of the genus *Cupriavidus*: a tale of lost and found. *Int J Syst Evol Microbiol* 54:2285–2289. doi:[10.1099/ijs.0.63247-0](https://doi.org/10.1099/ijs.0.63247-0)
- Weightman AJ, Topping AW, Hill KE, Lee LL, Sakai K, Slater JH, Thomas AW (2002) Transposition of DEH, a broad-host-range transposon flanked by IS*Ppu12*, in *Pseudomonas putida* is associated with genomic rearrangements and dehalogenase gene silencing. *J Bacteriol* 184:6581–6591. doi:[10.1128/JB.184.23.6581-6591.2002](https://doi.org/10.1128/JB.184.23.6581-6591.2002)
- Wen A, Fegan M, Hayward C, Chakraborty S, Sly LI (1999) Phylogenetic relationships among members of the *Comamonadaceae*, and description of *Delftia acidovorans* (den Dooren de Jong 1926 and Tamaoka et al. 1987) gen. nov., comb. nov. *Int J Syst Bacteriol* 49:567–576
- Williams PA, Jones RM, Shaw LE (2002) A third transposable element, IS*Ppu12*, from the toluene-xylene catabolic plasmid pWW0 of *Pseudomonas putida* mt-2. *J Bacteriol* 184:6572–6580. doi:[10.1128/JB.184.23.6572-6580.2002](https://doi.org/10.1128/JB.184.23.6572-6580.2002)
- Wyndham RC, Nakatsu C, Peel M, Cashore A, Ng J, Szilagyfi F (1994) Distribution of the catabolic transposon Tn5271 in a groundwater bioremediation system. *Appl Environ Microbiol* 60:86–93
- Yabuuchi E, Kosako Y, Yano I, Hotta H, Nishiuchi Y (1995) Transfer of two *Burkholderia* and an *Alcaligenes* species to *Ralstonia* gen. Nov.: Proposal of *Ralstonia pickettii* (Ralston, Palleroni and Doudoroff 1973) comb. Nov., *Ralstonia solanacearum* (Smith 1896) comb. Nov. and *Ralstonia eutropha* (Davis 1969) comb. Nov. *Microbiol Immunol* 39:897–904
- Zhang YB, Monchy S, Greenberg B, Mergeay M, Gang O, Taghavi S, van der Lelie D (2009) ArsR arsenic-resistance regulatory protein from *Cupriavidus metallidurans* CH34. *Antonie van Leeuwenhoek*. doi:[10.1007/s10482-009-9313-z](https://doi.org/10.1007/s10482-009-9313-z)