Open access • Journal Article • DOI:10.1111/1462-2920.13994

Catabolic task division between two near-isogenic subpopulations co-existing in a herbicide-degrading bacterial consortium: consequences for the interspecies consortium metabolic model. — Source link

Pieter Albers, Cédric Lood, Başak Öztürk, Benjamin Horemans ...+9 more authors

Institutions: Katholieke Universiteit Leuven, Leiden University, IMEC, Universidad Técnica Particular de Loja

Published on: 01 Jan 2018 - Environmental Microbiology (John Wiley & Sons, Ltd)

Topics: Gene cluster

Catabolic task division between two interchangeable near-isogenic subpopulations co-existing in a herbicide-degrading bacterial consortium: consequences for the interspecies consortium metabolic model

Running title: Near isogenic subpopulations in a herbicide-degrading consortium

Authors: Pieter Albers[1], Cédric Lood [2,3], Basak Özturk[1], Benjamin Horemans[1], Rob Lavigne[3], Vera van Noort[2,4], René De Mot[2], Kathleen Marchal[2,5,6], Aminael Sanchez-Rodriguez[2,7] and Dirk Springael[1*].

Affiliation:

[1] Division of Soil and Water Management, Department of Earth and Environmental Sciences, KU Leuven, Belgium

[2] Centre of Microbial and Plant Genetics, Department of Microbial and Molecular Systems, KU Leuven, Belgium

[3] Laboratory of Gene Technology, KU Leuven, Kasteelpark Arenberg 21 box 2462, 3001 Heverlee, Belgium

[4] Institute of Biology, Leiden university, Leiden, The Netherlands

[5]Department of Plant Biotechnology and Bioinformatics, Ghent University

[6]Department of Information Technology, iMinds, Ghent University, Belgium

[7] Departamento de Ciencias Naturales, Universidad Tecnica Particular de Loja, Ecuador

***Corresponding author:**

Dirk Springael

Division of Soil and Water Management, Kasteelpark Arenberg 20 –box 2459, 3001 Leuven,

tel. +32 16 32 16 04,

fax +32 16 32 19 97,

dirk.springael@kuleuven.be

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process which may lead to differences between this version and the Version of Record. Please cite this article as an 'Accepted Article', doi: 10.1111/1462-2920.13994

Originality-Significance Statement

This paper elaborates on the interactions within a tripartite bacterial consortium in which the members synergistically degrade the phenylurea herbicide linuron. The ecological significance of this consortium became recently clear from studies that show that similar networks are active in linuron degradation in agricultural environments treated with linuron. The metabolic interactions in the consortium are however not fully understood since *Variovorax* sp. WDL1 that performs the first step in linuron hydrolysis, contains genes for further metabolism of the hydrolysis product 3,4-dichloroaniline (DCA), while apparently another strain in the consortium, *Comamonas testosteroni* WDL7, grows on DCA. We show here that two WDL1 subpopulations co-occur, one hydrolyzing linuron into DCA and one degrading DCA into chlorocatechol. Both subpopulations are isogenic except for the linuron hydrolysis gene and DCA catabolic genes which are exchanged at the same locus of an extrachromosomal plasmid-like element. This observation results into a new consortium metabolic interactive model that includes the existence of a WDL1 subpopulation producing and leaking DCA for use by other bacteria, including its near-isogenic DCA-degrading counterpart and strain WDL7. It implies further that the crucial WDL1 subpopulation hydrolyzing linuron, a reaction which does not deliver assimilable carbon nor energy, has to gain carbon and energy in another way, either by using metabolites downstream in DCA conversion or from other pathways from the other consortium members or by transiently housing the genes for DCA catabolism through gene transfer. We conclude that the presence of two essential linuron catabolic gene clusters in WDL1 as cargo on two near identical genetic elements with apparent mutual incompatibility is key to the existence of WDL1 in a consortium mode, which is a completely new concept in bacterial interactions.

Summary

Variovorax sp. WDL1 mediates hydrolysis of the herbicide linuron into 3,4-dichloroaniline (DCA) and *N,O*-dimethylhydroxylamine in a tripartite bacterial consortium with *Comamonas testosteroni* WDL7 and *Hyphomicrobium sulfonivorans* WDL6. Although strain WDL1 contains the *dcaQTA1A2B* operon for DCA oxidation, this conversion is mainly performed by WDL7. Phenotypic diversification observed in WDL1 cultures and scrutiny of the WDL1 genome suggest that WDL1 cultures consist of two dedicated subpopulations, i.e., a linuron-hydrolyzing subpopulation (Lin+DCA-) and a DCA-oxidizing subpopulation (Lin-DCA+). Whole genome analysis of strains representing the respective subpopulations revealed that they are identical, aside from the presence of *hylA* (in Lin+DCA- cells) and the *dcaQTA1A2B* gene cluster (in Lin-DCA+ cells), and that these catabolic gene modules replace each other at exactly the same locus on a 1380 kb extra-chromosomal element that shows plasmid features including transferability by conjugation. Both subpopulations proliferate in consortium biofilms fed with linuron, but Lin+DCA- cells compose the main WDL1 subpopulation. Our observations instigated revisiting of the interactions within the consortium and suggest that the physical separation of two essential linuron catabolic gene clusters in WDL1 by mutually exclusive integration in the same mobile genetic element is key to the existence of WDL1 in a consortium mode.

Introduction

Various environmental processes are performed through a joint effort of different microbial species. Together these species form a consortium that achieves a functionality unattainable by the individual species. The mineralization of organic xenobiotics without the formation of harmful metabolites is often due to combined catabolic activities of multiple species that

establish synergistic interactions within a consortium based on the exchange of metabolites (Horemans et al., 2016a). Linuron is a phenylurea herbicide that is metabolized by both bacterial consortia and individual bacterial strains. The first reported linuron-degrading consortium was isolated from an orchard soil with a history of linuron application (Dejonghe *et al*., 2003). Since then several linuron-metabolizing consortia with similar composition were reported (Breugelmans *et al*., 2007). Recent studies suggest the ecological relevance of such consortia. Linuron-metabolizing consortia were shown to be active in environmental settings such as agricultural soil or in on-farm biopurification systems (BPS) that treat linuron-contaminated wastewater (Dealtry et al., 2016; Horemans et al., 2016b).

In the consortium reported by Dejonghe *et al.* (2003), *Variovorax* sp. WDL1 hydrolyses linuron into 3,4-dichloroaniline (DCA) and *N*,*O*-dimethylhydroxylamine (*N,O*-DMHA) (Figure S1). *N,O*-DMHA is excreted into the medium and used by *Hyphomicrobium sulfonivorans* WDL6 as carbon and energy source. It is believed that the DCA conversion rate is low in strain WDL1 creating a bottleneck in linuron mineralization and resulting into leakage of DCA (Dejonghe *et al*., 2003). Excreted DCA sustains the growth of the third consortium member, *Comamonas testosteroni* WDL7, and DCA removal was suggested to increase the rate of linuron hydrolysis in *Variovorax* sp. WDL1 (Dejonghe *et al*., 2003). The fact that during growth on linuron on solid surfaces the consortium forms biofilms consisting of macro-colonies composed of all three strains, supports the suggested metabolic interactions (Breugelmans *et al*., 2008).

Variovorax sp. WDL1 contains all genes necessary for the complete degradation of linuron beyond DCA. The initial hydrolysis of linuron into DCA and *N,O*-DMHA is performed by a periplasmic linuron hydrolase HylA (Bers *et al*., 2013). DCA is likely converted to a

chlorocatechol intermediate by a multicomponent DCA dioxygenase encoded by the *dcaQTA1A2B* gene cluster (Breugelmans *et al*., 2010). The chlorocatechol intermediate is presumably further degraded to β-ketoadipic acid via a modified *ortho*-cleavage pathway encoded by the *ccdCDEF* gene cluster (Bers *et al*., 2013). It is not known yet why strain WDL1 is inefficient in the degradation of DCA and how cross-feeding on the metabolites benefits the consortium as a whole. Strain WDL1 constitutes a substantial fraction of cells in consortium biofilms (Breugelmans *et al*., 2008). Since the hydrolysis of linuron yields no energy or immediate carbon source, strain WDL1 has to acquire sufficient carbon and energy either from the intrinsic metabolism of DCA or from other consortium members that grow on DCA and/or *N,O*-DMHA. Analysis of the linuron catabolic gene clusters in the draft genome sequence of *Variovorax* sp. WDL1 unexpectedly revealed that it contains two genotypic variants of the *dca* genes: one carrying the complete *dcaQTA1A2B* cluster and one consisting of an incomplete *dcaQ* gene interrupted by *hylA*. Scrutiny of WDL1 mutants deficient in either linuron hydrolysis or DCA oxidation showed that these phenotypic variants define two subpopulations of WDL1 of which one contains *hylA* and one harbors the *dcaQTA1A2B* cluster. The growth of both subpopulations in consortium biofilms fed with linuron as sole carbon source was examined to understand community functioning.

Results

Variovorax sp. WDL1 cultures consist of two catabolically distinct subpopulations

The genes for linuron hydrolysis and DCA oxidation were previously located on different contigs of the draft genome sequence of strain WDL1 (Bers *et al*., 2013). As shown in Figure 1, contig 10 (9054 bp) contains the *dcaQTA1A2BR* gene cluster encoding the enzymes for the oxygenation of

DCA to a chlorocatechol intermediate. *dcaQTA1A2BR* shows 99% nucleotide identity with the *tadQTA1A2BR* cluster encoding aniline dioxygenase from *Delftia tsuruhatensis* AD9 (Liang *et al*., 2005). However, in contrast to the *tad* cluster in strain AD9, the WDL1 *dca* cluster displays a truncated homologue of the regulatory gene *tadR* containing only the first 97 nucleotides of the coding region (*dcaR**). On contig 10, *dcaR** is followed immediately by an inverted repeat (IR) of IS*1071*, indicating that truncation of *dcaR* is due to a transposition event involving IS*1071*. An identical IS*1071* IR was found at the other end of contig 10.

Insert figure 1 preferably here

The *hylA* gene*,* required for initial hydrolysis of linuron in strain WDL1, is located on contig 1 (7530 bp). Inspection of the *hylA* upstream region revealed a 43-bp coding region remnant spanning base pairs 650-692 of *dcaQ*. Sequencing of the PCR amplicon generated from primer W1-11 (targeting base pairs 589-602 of *dcaQ*) and primer W1-12 (targeting *hylA*) (Table S1) confirmed that *hylA* is located next to a *dcaQ* remnant (*dcaQ**) that could be allocated to a 4.2-kb long contig 104. The 11.7-kb long contig resulting from the assembly of contig 1 and contig 104 was designated contig 1.2 (Figure 1). Interestingly, the DNA sequence preceding *hylA* on contig 1.2 is 100% identical to the first part of contig 10 including a copy of IS*1071* IR, *orf10.8*, *orf10.7* and the *dcaQ** (*orf 10.6,* comprising only the first 692 nucleotides of the 1479 bp long *dcaQ*). The *dcaQ* gene encodes the glutamine synthetase-like component of the DcaQTA1A2B multicomponent enzyme. As the truncated version almost completely lacks the catalytic domain (Pfam PF00120), *dcaQ** is likely not functional. Similarly to contig 10, the other end of contig 1.2 also carries IS*1071* IR, suggesting that contig 1.2 originated from a gene rearrangement involving contig 10.

We hypothesized that either WDL1 cells contain a copy of both contig 1.2 with *hylA* and contig 10 with *dcaQTA1A2BR**, or that a WDL1 culture consists of two subpopulations, one containing contig 1.2 and one containing contig 10. To examine this, we assessed the linuron and DCA degradation activities of individual colonies of strain WDL1. After plating of a cryostock of WDL1 on R2A supplemented with linuron, 44 colonies were tested for linuron and DCA degradation after one week of growth in R2B medium and they were cryopreserved for further analysis. Three different degradation phenotypes were observed, i.e., 20% of the colonies degraded both linuron and DCA (designated as Lin+DCA+), 57% degraded linuron but not DCA (Lin+DCA-), and 23% degraded 3,4-DCA but not linuron (Lin-DCA+) (Figure 2A).

Insert figure 2 preferably here.

To examine whether colonies originating from the three phenotypic lineages retained their catabolic phenotype, four cultures of all three phenotypes that had been cryopreserved after one week of growth of the original culture in R2B medium were plated again on R2A supplemented with linuron. Per cryoculture, five individual colonies were tested for degradation of linuron and DCA supplemented in R2B (Figure 2B). Only 10% of the colonies (2 out of 20) originating from Lin+DCA+ cultures retained a Lin+DCA+ phenotype indicating that Lin+DCA+ is an unstable phenotype that is readily lost after one week of growth in liquid medium. The remainder of the colonies originating from the Lin+DCA+ culture tested as either Lin+DCA- (40%) or Lin-DCA+ (50%). In contrast, colonies originating from cultures of Lin+DCA- and Lin-DCA+ variants did not show an altered phenotype. Colonies showing a Lin-DCA- phenotype were never observed. These results indicate that the co-occurrence of Lin+ and DCA+ phenotypes is unstable and that both phenotypes tend to be mutually exclusive.

To verify whether we could link the degradation phenotypes to the presence of catabolic sequences corresponding to contig 1.2 and 10, the abundance of the respective sequences in the WDL1 population was determined in cryocultures of all three degradation phenotypes by qPCR using appropriate primer sets (see materials and methods and Table S1). To quantify contig 1.2 DNA, two primer pairs were used, i.e., one targeting the *hylA* and one targeting the *dcaQ*-hylA* intergenic region. Contig 10 DNA was quantified by qPCR using a forward primer targeting a region of *dcaQ* that is in common between contig 1.2 and 10, and a reverse primer targeting a part of *dcaQ* that was unique to contig 10 (Figure 1). The abundances of sequences specific for contig 1.2 and contig 10 relative to the total WDL1 population was calculated as the ratio of target gene copies to copies of the chromosomal gene *omp32.* Relating the abundances of the two genomic regions to the linuron/DCA catabolic phenotypes (Table S2 and Figure 3) showed that cultures displaying a Lin+ phenotype were clearly enriched in sequences corresponding to *hylA*-carrying contig 1.2 whereas cultures with a DCA+ phenotype contained primarily sequences matching contig 10 with the *dcaQTA1A2B* operon. Moreover, cultures displaying Lin- or DCA- phenotypes essentially lacked respectively *hylA* (contig 1.2) or the *dcaQTA1A2B* cluster (contig 10). The sum of the abundances of both contig-specific sequences was not significantly different across the different linuron degradation variants (one way ANOVA, $p = 0.09$ for the sum of *dcaQ*-hylA* intergenic region and *dcaQTA1A2B*), suggesting that the presence of genomic regions with either the *hylA* gene or the *dcaQTA1A2B* operon in WDL1 cultures is complementary (Table S2). Further evidence for this was provided by PCRs targeting *dcaT* and *dcaA1* (primer sets in Table S1). *dcaT* and *dcaA1* were only present in cultures with a DCA+ phenotype (Figure S2). These results taken, together with the finding that a WDL1 culture

showing a Lin+DCA+ phenotype segregates into variants with either a Lin+DCA- or Lin-DCA+ phenotype, suggest that (i) genomic regions corresponding to contig 1.2 and contig 10 are not stably maintained together in an individual cell, (ii) a Lin+DCA+ culture is actually a mix of two stable subpopulations, i.e. a linuron-hydrolyzing variant carrying *hylA* and a DCA-oxidizing variant carrying a complete *dcaQTA1A2B* cluster and that (iii) the contig 1.2 gene configuration may have arisen by replacing the *dcaQTA1A2B* gene cluster in the same strain and, hence, that the two subpopulations are variants of the same strain and not two different strains.

Insert figure 3 preferably here.

The two *Variovorax* sp. WDL1 subpopulations have arisen from exchange at the same locus of the *hylA* gene and *dcaQTA1A2BR* gene cluster.

Illumina *de novo* sequencing and nanopore sequencing technology was combined to compare the genomic sequences of a strain representing the *hylA*-containing Lin+DCA- WDL1 subpopulation (strain B2) and a strain representing the *dcaQTA1A2BR*-containing Lin-DCA+ WDL1 subpopulation (strain B4). As such we examined whether these subpopulations are indeed variants of the same strain and analyzed the genetic context of the *hylA* and *dcaQTA1A2BR* in the subpopulations. Nineteen and 16 contigs were recovered in strain B2 and B4, respectively. The contigs could be allocated to replicons using the Bandage software (Figure S3). The genomes of both strains consist of six replicons including the main chromosome of around 5400 kb and five smaller replicons of 1380 kb, 1240 kb, 200 kb, 24 kb and 20 kb (Figure S3). The genetic organization of contigs 1.2 and 10 was confirmed in respectively strain B2 and strain B4. In both strains, the two catabolic gene clusters are located on the extrachromosomal

element of around 1380 kb, and in both cases are flanked at both ends by complete IS*1071* elements constituting a composite transposon. The organization of a large region of this element depicting the most relevant gene functions in the respective strains is provided in Figure 4. The organization of the replicon is exactly the same in both strains with the exception of the *hylA* and *dcaQTA1A2BR** region contained between the two IS*1071* elements. The replicon shows different features of plasmids including genes for partitioning, maintenance and conjugation (Table S3). Its average GC content of 62.9 % is substantially different from the GC content of 67.2% found on the main chromosome (Figure S4 and Table S4). Moreover, the element appears to be a reservoir for adaptive accessory traits including metal resistance (in particular multiple resistance-nodulation-cell division (RND)-driven trans-envelope exporter systems) and catabolism of aromatic compounds different from linuron. In both strains, the element also carries the genes that are likely responsible for chlorocatechol degradation (*ccdEDFC genes*), located around 18 kb distant from the *hylA*/*dcaQTA1A2BR** region. The element is further characterized by a high number of integrase/transposase gene functions and IS elements including 6 of the 7 copies of IS*1071* present in WDL1. Genes that show clear homology with core genes were absent. In contrast to the replicon carrying *hylA*/*dcaQTA1A2BR**, the 1240 kb replicon does not encode known plasmid functions and its GC content is identical to this of the main chromosome. The smaller replicons of 200, 24 and 20 kb are indisputably plasmids showing plasmid replication functions. The 24 kb plasmid carries the seventh IS*1071* copy while the 200 kb plasmid contains other aromatic catabolic genes and a large number of IS elements (data not shown). Whole genome analysis showed that the B2 and B4 strains have the same genomic organization (Figure S3) as well as identical nucleotide

sequences except for the regions carrying the *hylA*/*dcaQTA1A2BR* genes. Indeed, the average nucleotide identity (as analyzed by ANI) between both genomes is 100%, and dot plots of the different genomic components further confirm the sequence identity (Figure S5). We conclude that the two subpopulations are two variants of the same strain only differing in *hylA* and *dcaQTA1A2BR** genes and that the subpopulations represented by strain B2 probably developed from strain B4 by replacing the *dcaQTA1A2BR** gene cluster with *hylA*.

Insert figure 4 preferably here.

Contributions of *Variovorax* sp. WDL1 subpopulations to biomass in a linuron-metabolizing consortium biofilm

The identification of two WDL1 subpopulations has important implications for the metabolic interactions within the consortium since it implies that the WDL1 population, that composes 31% of the total community in linuron-metabolizing consortium biofilms, actually consists of two variants of which the Lin-DCA+ variant competes with *C. testosteroni* WDL7 for DCA produced by WDL1 Lin+DCA- variants. Since the hydrolysis of linuron in the Lin+DCA- variant yields no readily assimilable carbon nor energy, the abundant WDL1 population in consortium biofilms might mainly consist of Lin-DCA+ cells that are able to acquire carbon for growth from linuron unless the Lin+DCA- variant in its turn would acquire further downstream metabolites from other consortium members, including the Lin-DCA+ WDL1 variant (Breugelmans *et al.,* 2008). To study this, we grew the consortium as biofilms in flow chambers continuously fed with linuron as sole carbon source in exactly the same way as done previously by Breugelmans et al. (2008), starting from a cryoculture of the WDL1 strain containing both subpopulations. After reaching steady-state activity regarding linuron degradation for at least one week, the relative

contribution of both WDL1 variants to the total WDL1 biomass in the biofilm was estimated by determining the abundance of sequences corresponding to contig 1.2 (by means of qPCR targeting either *hylA* or the *dcaQ*-hylA* junction) and contig 10 (by means of qPCR targeting the complete *dcaQ* gene). Two independent biofilm experiments with identical setup were performed using two different WDL1 precultures that were both initiated using a streak of WDL1 cells taken from R2A plates supplemented with linuron as inoculum. Both experiments included three replicates.

In the first experiment, the consortium reached steady-state linuron degradation after 28 days with the linuron effluent concentration stagnating around 6 mg L^{-1} (Figure 5A). The maximal transiently accumulated 3,4-DCA concentration was 1.3 mg L^{-1}, in accordance with previous observations (Breugelmans *et al*., 2008). At harvest, relative abundances of the WDL1 and WDL7 populations in the consortium biofilms, estimated by measuring the ratio of the number of 16S rRNA gene copies of WDL1 and WDL7 to the total number of bacterial 16S rRNA gene copies, were respectively 32% ± 8% and 55% ± 18% for WDL1 and WDL7, similar to previous estimates (Breugelmans *et al*., 2008). The abundance of the contig 1.2 markers for the Lin+DCA- subpopulation relative to the total WDL1 population was around two-fold (Table S5; Figure 5B). In contrast, the relative abundance of the contig 10 marker for the Lin-DCA+ subpopulation was one tenth. The second biofilm experiment showed similar linuron and DCA degradation kinetics (Figure S6) and the relative abundance of Lin+DCA- subpopulation was about 1, while the relative abundance of the Lin-DCA+ subpopulation was on average more than 1000 times lower than in the first biofilm experiment (Table S5).

Insert figure 5 preferably here.

Discussion

The linuron catabolic genes in *Variovorax* sp. WDL1 are located on putative composite transposons

Our results show that strain WDL1 consist of two near-isogenic subpopulations that only differ in a catabolic region that carry crucial genes for linuron metabolism. One subpopulation carries the linuron hydrolysis gene *hylA* and the other subpopulation the DCA catabolic genes for conversion of DCA into chlorocatechol. Both genomic regions are flanked by IS*1071* elements. Such organization represents a composite transposon structure and in few cases actual transposition of such a structure was shown (Larrain-Linton *et al*., 2006; Ng and Wyndham, 2011). The association of both linuron catabolic gene clusters with IS*1071* supports the previous appraisal that the linuron catabolic pathway in WDL1 has resulted from patchwork assembly of catabolic gene modules allowing strain WDL1 to use linuron as the sole source of carbon and energy (Bers *et al*., 2013). Interestingly, the linuron hydrolysis gene and DCA catabolic genes in the two subpopulations are exchanged at the same locus within the putative composite transposon and seem to have replaced one another. Moreover, the acquisition of the linuron hydrolase-encoding gene *hylA* by strain WDL1 appears to have caused truncation of *dcaQ* and the loss of the remainder of the *dca* genes. This suggests that the composite transposon carrying the *hylA* gene originated from the transposon carrying the *dca* cluster and that rather than recruitment of *hylA* by a single transpositional insertion event, a double homologous

recombination event between a DNA element carrying *hylA* and the DNA element carrying the *dca*-cluster took place resulting in the deletion of most of the *dca*-cluster.

Plasmid-based mutual exclusion of linuron degradation genes is responsible for the existence of catabolic subpopulations of *Variovorax* sp. WDL1

IS*1071*-based composite catabolic transposons are often located on plasmids (Sota *et al*., 2006). Our results show that also *hylA*/*dcaQTA1A2BR** in WDL1 are located on an extrachromosomal replicon of 1.38 Mb. The replicon encodes plasmid replication and partitioning systems and the majority of their genes confer accessory functions. Furthermore, the replicon shows features of a conjugative nature which indicates that this element can be transferred between cells. The conjugative nature of the element is further supported by the presence of a high number of adaptive trait gene functions, the large abundance of IS elements and integrases that are used to recruit other genes. Its average GC content is substantially lower than the GC content on the main chromosome (4.3% difference). Together with the absence of genuine housekeeping genes, this indicates that the 1.38-Mb replicon represents a megaplasmid rather than a chromid, whereas the replicon of 1.24 Mb not associated with linuron degradation would qualify as a chromid. According to Harrison *et al.* (2010), chromids have a nucleotide composition close to that of the chromosome and carry at least one gene required for cell viability. Putative chromids have been proposed to occur in *Variovorax paradoxus* strains B4 (1.35 Mb) and S110 (1.13 Mb) (diCenzo and Finan, 2017). Replicons of similar size as the WDL1 megaplasmid have been experimentally shown not to be essential in the α-proteobacterial plant symbiont *Ensifer (Sinorhizobium) meliloti* (diCenzo et al., 2014) and in the β-proteobacterial

human pathogen *Burkholderia lata* (megaplasmid pC3 of 1.4 Mb). No previous report exists of the *hylA*, *dcaQTA1A1B* and *ccdCDEF* genes being present on such a large replicon. Homologues of the *dcaQTA1A1B* and *ccdCDEF* genes have been associated with IncP1 plasmids. In contrast, not much information exists about the genetic context of the *hylA* gene, but recently IncP1 plasmids that apparently can transfer the *hylA* gene were exogenously isolated from the microbial community of an on-farm biopurification system enriched for linuron degrading populations (Dealtry et al., 2016). As such, we can conclude that *hylA* and d*ca* gene cluster in the two WDL1 subpopulations are carried by plasmid variants that only differ in accessory load composition. The encoding of both genotypes on the same plasmid precludes stable simultaneous inheritance of *hylA* and *dcaQTA1A2BR** in WDL1 cells due to plasmid incompatibility when both plasmids are present in the same cell, explaining the observed segregation in either a Lin+DCA- or Lin-DCA+ phenotype upon cell division and the mutual exclusion of the two catabolic markers. The (only temporary) occurrence of both plasmids in the same cell in Lin+DCA+ colonies after plating of the WDL1 culture is then explained by the exchange of the two plasmid variants between both WDL1 subpopulations in the culture. Conjugation of a particular plasmid into cells carrying an incompatible plasmid has been shown before for plasmid curing purposes, for instance for IncP1 plasmids (Springael *et al*., 1993).

Existence of two *Variovorax* sp. WDL1 subpopulations: linuron degradation by consortium biofilms revisited

It is currently assumed that in the tripartite consortium, *C. testosteroni* WDL7 is responsible for the further degradation of most of the DCA formed during linuron hydrolysis by WDL1 (Dejonghe *et al*., 2003). As in previous studies (Breugelmans *et al*. 2008), this was reflected in

the high abundance of strain WDL7 constituting about 55% of the total community (compared to about 32% for the total WDL1 population) and the abundance of the DCA-oxidizing WDL1 subpopulation in consortium biofilms, being on average much lower than WDL7. These observations confirm that strain WDL7 is the main DCA oxidizer when WDL1 is grown in a consortium. However, the competition between strain WDL7 and the DCA-oxidizing WDL1 subpopulation might have been affected by the initial ratio of DCA-oxidizing WDL1 cells to WDL7 cells in the inoculum. We decided to perform the biofilm experiments in exactly the same way as we previously did in order to be able to make comparisons with previous experiment examining the consortium composition in biofilms. Hence, we started from the same WDL1 cryoculture containing both subpopulations. The abundance of the Lin-DCA+ WDL1 subpopulation in the inoculum was not measured but we determined the abundance of Lin-DCA+ WDL1 cells relative to the total WDL1 population in the cryoculture that was used for generating the inoculum in the biofilm experiments, showing a fairly low abundance of 0.6 % and therefore we cannot exclude that the Lin-DCA+ WDL1 subpopulations will perform better in case initially present at higher numbers. On the other hand, Dejonghe et al. (2003) showed that DCA degradation rate in WDL1 were substantially lower than in strain WDL7. In the consortium biofilms, the Lin+DCA- WDL1 subpopulation represented the main fraction of the total WDL1 population. This suggests that the Lin+DCA- WDL1 subpopulation proliferates in the biofilm system and acquires carbon and energy, despite the fact that linuron hydrolysis yields no energy nor immediate carbon source for the Lin+DCA- subpopulation carrying the HylA encoding megaplasmid. Two explanations for this observation can be put forward. First, linuron hydrolyzers might depend for growth on metabolites provided by other consortium members

(i.e. WDL6, WDL7 and the Lin-DCA+ WDL1 variant) that in return rely for growth on DCA or *N,O*-DMHA released from linuron-hydrolyzing cells. These metabolites might be compounds generated further downstream in the catabolic pathway like chlorocatechol produced by strain WDL7 or by the Lin-DCA+ WDL1 variant. Otherwise, they might be products from anabolic pathways active in either WDL7, WDL6 or the Lin-DCA+ WDL1 variant. This implies that interaction between WDL1 and WDL6/WDL7 is not going one way but that WDL1 and WDL6/WDL7 mutually benefit from each other. Alternatively, the continuous exchange of the large replicon carrying the *hylA* gene or *dcaQTA1A1B* cluster between the WDL1 subpopulations will create cells that temporally harbor copies of both plasmid variants and during that time are able to produce both HylA and DcaQTA1A2B enzymes.

Mutual exclusion of the *hylA* and *dcaQTA1A2B* genes in *Variovorax* sp. WDL1: living in a consortium for a reason

The presence of two functionally interconnected catabolic modules on otherwise identical genetic elements and their apparent mutual exclusion might answer questions about the lifestyle of *Variovorax* sp. WDL1 and its co-habitation with *C. testosteroni* WDL7 that uses DCA from strain WDL1 to grow. The incompatibility of both catabolic genotypes means that strain WDL1 will always hold a Lin+DCA- subpopulation that only converts linuron to DCA without performing the further enzymatic steps and hence releases DCA. The release of DCA from the Lin+DCA- WDL1 subpopulation will lead to opportunities for bacteria other than the Lin-DCA+ WDL1 subpopulation to use DCA. Organisms that more efficiently take up and metabolize DCA from the environment than Lin-DCA+ WDL1 cells will thus benefit from the activity of the Lin+DCA- WDL1 subpopulation. In the original consortium described by (Dejonghe et al., 2003),

apart from WDL7, other DCA degraders were identified showing that the metabolic interactions between WDL1 and WDL7 are not unique and that also other strains can profit from the cross-feeding of DCA (Dejonghe *et al*., 2003). However, one can wonder why the IS*1071* composite transposons harboring *dcaQTA1A2BR** and *hylA* do not insert in other parts of the genome. For instance, genomic integration of either *dcaQTA1A2BR** or *hylA* would result into a WDL1 variant that would stably inherit both catabolic pathways and to directly use DCA. This apparently does not occur either because site-specific integration of IS*1071* composite transposons in IncP1 sequences is required or, alternatively, because the separation of both pathways is evolutionary advantageous for WDL1. In the latter scenario, the presence of the two pathways on mutually exclusive genetic elements provides a vehicle to keep this advantageous constellation. *Variovorax* strains carrying a complete pathway for linuron degradation in one cell exist, but instead of *hylA* they contain a different linuron hydrolase gene, *libA* (Bers *et al*., 2011b). As linuron-degrading consortia similar in taxonomic composition to the WDL1/WDL6/WDL7 consortium exist in the environment (Breugelmans et al., 2007; Dealtry et al., 2016; Horemans et al., 2016b), it would be of interest to examine whether WDL1 counterparts in other linuron-assimilating bacterial consortia also consist of subpopulations that have a similar genetic context of the *hylA* gene and *dca* gene cluster, which would argue for an evolutionary driver for such constellation.

Experimental Procedures

Chemicals

Pestanal analytical standards (99.9 %) of linuron (3-(3,4-dichlorophenyl)-1-methoxy-1-methylurea) and 3,4-dichloroaniline were purchased from Sigma-Aldrich (Belgium).

Bacterial strains, media and growth conditions

Strains *Variovorax* sp. WDL1 (LMG 27260), *C. testosteroni* WDL7 (LMG 27261) and *H. sulfonivorans* WDL6 (LMG 27262) were described before (Dejonghe *et al*., 2003) and routinely grown as described by Horemans *et al.* (2013). Briefly, *Variovorax* sp. WDL1 was grown in liquid R2B medium (R2A medium without agar) (Reasoner and Geldreich, 1985) supplemented with 20 mg L^{-1} linuron. *C. testosteroni* WDL7 was grown in Luria-Bertani (LB) medium (Sambrook and Russell, 2001). *H. sulfonivorans* WDL6 was grown in MMO minimal medium (Boon *et al*., 2001) supplemented with 1% (vol/vol) methanol as carbon source.

Draft genome reconstruction and sequence analysis of *Variovorax* sp. WDL1

A cryoculture of WDL1 was plated on R2A supplemented with linuron, a smear of grown cells was inoculated in R2B supplemented with linuron, and grown until exponential phase (OD_{600} 0.15). To assure the Lin+ and DCA+ phenotype of the culture, depletion of linuron and DCA was verified by determining their concentrations on a regular basis by High Performance Liquid Chromatography (HPLC) analysis as reported by Horemans *et al.* (2014). Genomic DNA was extracted from 1-mL culture using the Puregene Core kit A (Qiagen, Belgium) following the manufacturer's instructions. Three hundred MB of 75 cycle pair-ended reads of genomic DNA of WDL1 was obtained by Illumina Hi-Seq sequencing performed by Baseclear (The Netherlands).

Nucleotides with a PHRED quality score <20 were trimmed from the end of reads and trimmed reads with a minimal length <10 were discarded using the FASTX-Toolkit-0.0.12 software (http://hannonlab.cshl.edu/fastx_toolkit). The draft genome sequence was composed by assembling the trimmed pair-ended reads to contigs using Velvet software version 1.2.01 with an optimized k-mer length of 41 and a minimal contig length of 100 bp (Zerbino and Birney, 2008). The n50 of the assembly was 90473 bp and the resulting 8,170,112 bp draft genome sequence consists of 493 contigs with an average coverage depth of 31x. Gap closing with primer pair W1-11, W1-12 (Table S1) was used to merge contig_104 and contig_1 (Genbank accession file numbers LMTS01000339 and KC146403 respectively), forming contig 1.2. The draft genome was uploaded onto the web-based RAST server for automatic annotation using the Rapid Annotation using Subsystem Technology (Aziz *et al*., 2008). Linuron catabolic gene regions were manually curated by recalling the genes of interest in the draft genome by localBLAST using BLASTX to identify proteins in the non-redundant protein database with significant sequence similarity to the recalled gene products (Altschul *et al*., 1990).

Draft genome sequences of WDL1 subpopulations and genome analysis

Strains B2 and B4 representing the two WDL1 subpopulations were grown in R2B medium supplemented with 20 mg L^{-1} linuron or 14 mg L^{-1} DCA, respectively. The phenotype of each subpopulation was confirmed by measuring the depletion of linuron or DCA. Linuron and DCA concentrations were determined by HPLC analysis. Cultures were grown to an OD_{600} of 1 and 10 ml of cell culture was used to extract DNA with a cetyl trimethylammonium bromide (CTAB)-lysozyme method (Larsen *et al*., 2007). Short read sequencing of the two genomes was performed using Illumina MiSeq technology with a 2 x 150 bp, paired-end approach. Quality of

the reads was assessed using FastQC (Andrews S., 2010), and reads were processed using BBDuk for both quality (trimming of PHRED scores < 28), length (minimum 50 bp), and adapters contamination. Long reads sequencing of the two genomes was performed by nanopore sequencing using the MinION sequencer (Oxford Nanopore Technology). The genomic DNAs of the two strains were processed separately and mechanically sheared to an average fragment length of 8 kb, barcoded and subsequently pooled. The library preparation was done using the 1D ligation approach and the result was sequenced on a SpotION flowcell (R9.4). Genome assembly was performed using Unicycler (Wick et al., 2017), which uses a hybrid approach that combined the short reads from Illumina with the long reads from Nanopore. The genomic organization and subsequent assignment of the contigs to their respective components (replicons) was done by inspecting the assembly graph produced by the assembler using Bandage (Wick et al., 2015). The two genomes were functionally annotated using Prokka (Seemann, 2014). To assess the genome sequence identity of both subpopulations, the average nucleotide identity (ANI) was calculated using the ANI Calculator (Rodriguez-R and Konstantinidis, 2016) and further analysis was performed with the dotplot tool Gepard (Krumsiek et al., 2007).

Suspension batch linuron degradation assay

A cryoculture of WDL1 was plated on R2A supplemented with linuron. Individual colonies were suspended in 20 µL 10 mM $MgSO_4$ and subsequently transferred to 5 mL R2B medium supplemented with 20 mg L^{-1} linuron or 14 mg L^{-1} DCA. R2B was used as background medium to provide an additional source of carbon to acquire sufficient biomass for activity. After one week,

1-mL samples were taken, centrifuged at 10,000 x*g* for 5 min, and the supernatant stored at -20°C until HPLC for determining linuron and DCA concentration (Horemans *et al.* (2014).

Biofilm experiments

Consortium biofilms were cultivated at 25 °C in continuously fed flow chamber systems (©BioCentrum DTU, Denmark) as described previously (Breugelmans *et al*., 2008). Briefly, a cryoculture of WDL1 was plated on R2A supplemented with linuron and a smear of colonies was inoculated in liquid 50 mL R2B medium supplemented with linuron. Cultures of WDL6 and WDL7 were grown in 20 mL MMO with 1% (vol/vol) methanol and in LB medium, respectively. All cultures were harvested at late exponential phase and washed two times with 10 mM $MgSO_4$ solution. Cell densities were adjusted to OD_{600} 0.25 for WDL6 and WDL7 and OD_{600} 0.5 for WDL1. Equal volumes of the three cultures were mixed and 200 µL of the mixed inoculum was injected in the flow chambers using a syringe (Myjector U−100 Insulin needle, Terumo, Japan). The cells were left to settle on the glass substratum for 1 hour after which the flow chambers were irrigated with MMO containing 20 ± 2 mg L^{-1} linuron at 3.5 ml h^{-1}. At regular time intervals, linuron and DCA concentrations in the effluent were determined by HPLC. One-ml effluent samples were taken, centrifuged at 10,000 x*g* for 5 min and the supernatant stored at -20°C until measurement. Harvesting of biofilm cells for DNA extraction was performed as follows. The cells were collected from the flow chambers by injecting and pipetting up and down (10 times) 1 mL of 10 mM $MgSO_4$ in the channel. The suspension was centrifuged (4,000 x *g* for 15 min), the DNA extracted from the pellet using the CTAB-lysozyme method (Larsen *et al*., 2007) and the DNA was dissolved in deioinized water.

Real-time quantitative PCR

Real-time quantitative PCR (qPCR) was performed on either DNA extracted from biofilm cultures or in case of cryocultures, on cell lysates obtained by taking a loop of cryoculture cells suspended in 20 µL deionized water and boiling the suspension for 10 min at 100°C. qPCR targeting the bacterial 16S rRNA gene was performed as described previously (Haest *et al*., 2011). The copy number of the 16S rRNA gene of WDL1 was quantified by performing a qPCR targeting the *Variovorax* 16S rRNA gene as described previously (Bers *et al*., 2011a). qPCR targeting *hylA* was performed as described by Horemans *et al.* (2016). For all other qPCR assays the reaction mixture contained 7.5 µL of Absolute QPCR SYBRs Green mix (Thermo Scientific, USA), 0.03 µL of both forward and reverse primer (100 µM), 5.94 µL of water and 1.5 µL of extracted biofilm DNA or cryosuspension. Primers, listed in Table S1, were designed using Primer3-v.0.4.0 software (Untergasser *et al*., 2012) and tested for the desired specificity using genomic DNA of strains WDL1, WDL7 and WDL6 as template. All primers were purchased from Integrated DNA Technologies (Belgium). qPCR was performed in duplicate in a Rotor Gene 6000 real-time cycler apparatus (Qiagen, Germany). For qPCR assays targeting *dcaQ*, the *dcaQ*-hylA* intergenic region and *omp32*, the reaction consisted of an initial denaturation step of 15 min at 94°C, followed by 40 cycles of 20 s at 94°C, 20 s at 60°C and 20 s at 72°C. The qPCR assay targeting the 16S rRNA gene of WDL7 consisted of an initial denaturation step of 15 min at 94°C, followed by 40 cycles of 15 s at 94°C, 15 s at 50°C and 15 s at 72°C. To generate a standard curve, a 10-fold dilution series of 10^8 copies μL^{-1} of amplicons of appropriate gene fragments generated by conventional PCR on genomic DNA of WDL1 or WDL7, was included in every qPCR assay. The amplicons were purified from agarose gels using the QIAquick Gel Extraction Kit

(Qiagen, Germany) and DNA concentrations of the purified DNA fragments were determined with the NanoDrop 1000 Spectrophotometer (Thermo Scientific, USA). The detection limit of all qPCR assays was 10 copies μL^{-1}. The copy number of the chromosomal gene *omp32* (encoding APY03_1563, putative outer membrane porin 32 precursor) was used to normalize the concentration of target gene copies to the number of WDL1 cells in each sample. The abundances of the different target genes relative to the total WDL1 population, was expressed as ratio of target gene copies to gene copies of the chromosomal gene *omp32*. In case the target gene copy number was below the detection limit, the detection limit was used to calculate the abundances. Significant differences in abundance between samples were examined by one way analysis of variance (ANOVA).

Conventional PCR

Amplification of standard template for real time quantitative PCR assays and amplification of catabolic genes *dcaT* and *dcaA1* was performed using the PCR primer sets listed in Table S1. Primers were designed using Primer3-v.0.4.0 software (Untergasser *et al*., 2012). PCR was performed in a total volume of 25 µL containing 2.5 µL 10x Dream Taq Green buffer (Thermo Scientific, USA), 2 µL 2.5 mM dNTPs, 0.125 µL of both primers, 0.125 µL DreamTaq DNA Polymerase (5 U/µL) (Thermo Scientific, USA). As a template, 1 µL of DNA was used for standard template amplification and 1 µL of cryoculture lysate was used for the amplification of the catabolic genes. PCR was performed in a Biometra Thermocycler (AnalytikJena, Germany) and consisted of an initial denaturation step of 5 min at 94°C, followed by 35 cycles of 30 sec at 94°C, 20 sec at 60°C and 20 sec at 72°C and a final elongation step at 72°C for 5 min. Amplicons were

visualized by agarose gel electrophoresis (1% agarose, 60 min, 90V) using Gelred (Biotum, USA) as nucleic acid stain.

Nucleotide sequence accession numbers

The draft genome sequence of *Variovorax* sp. WDL1 and contig 1.2 have been deposited at DDBJ/EMBL/GenBank under the accession numbers LMTS00000000 and KU973601, respectively. The draft genomes of *Variovorax* sp. WDL1 subpopulations B2 and B4 can be accessed under the accession numbers NNBR00000000 and NNBQ00000000 respectively.

Acknowledgments

This work was funded by IWT (Innovation by Science and Technology) project SB/101628, the KU Leuven projects OT/10/030 and STRT/13/004, the FWO post-doctoral fellow grant n° 12Q0215N to B.H., FWO project G.0371.06 and by the Inter-University Attraction Pole (IUAP) "μ-manager" of the Belgian Science Policy (BELSPO, P7/25). The authors declare no competing interests.

References

Agnoli, K., Schwager, S., Uehlinger, S., Vergunst, A., Viteri, D.F., Nguyen, D.T. et al. (2012) Exposing the third chromosome of *Burkholderia cepacia* complex strains as a virulence plasmid. *Mol Microbiol* **83**: 362-378.

Altschul, S.F., Gish, W., Miller, W., Myers, E.W., and Lipman, D.J. (1990) Basic local alignment search tool. *J Mol Biol* **215**: 403-410.

Andrews, S. (2010) FastQC: a quality control tool for high throughput sequence data. [WWW document]. URL http://www.bioinformatics.babraham.ac.uk/projects/fastqc.

Aziz, R.K., Bartels, D., Best, A.A., DeJongh, M., Disz, T., Edwards, R.A. et al. (2008) The RAST Server: rapid annotations using subsystems technology. *BMC Genomics* **9**: 75.

Bers, K., Batisson, I., Proost, P., Wattiez, R., De Mot, R., and Springael, D. (2013) HylA, an alternative hydrolase for initiation of catabolism of the phenylurea herbicide linuron in *Variovorax* sp. strains. *Appl Environ Microbiol* **79**: 5258-5263.

Bers, K., Sniegowski, K., Albers, P., Breugelmans, P., Hendrickx, L., De Mot, R., and Springael, D. (2011a) A molecular toolbox to estimate the number and diversity of *Variovorax* in the environment: application in soils treated with the phenylurea herbicide linuron. *FEMS Microbiol Ecol* **76**: 14-25.
Bers, K., Leroy, B., Breugelmans, P., Albers, P., Lavigne, R., Sorensen, S.R. et al. (2011b) A novel hydrolase identified by genomic-proteomic analysis of phenylurea herbicide mineralization by *Variovorax* sp. strain SRS16. *Appl Environ Microbiol* **77**: 8754-8764.
Boon, N., Goris, J., De Vos, P., Verstraete, W., and Top, E.M. (2001) Genetic diversity among 3-chloroaniline- and aniline-degrading strains of the C*omamonadaceae*. *Appl Environ Microbiol* **67**: 1107-1115.
Breugelmans, P., D'Huys, P.J., De Mot, R., and Springael, D. (2007) Characterization of novel linuron-mineralizing bacterial consortia enriched from long-term linuron-treated agricultural soils. *FEMS Microbiol Ecol* **62**: 374-385.
Breugelmans, P., Horemans, B., Hofkens, J., and Springael, D. (2010) Response to mixed substrate feeds of the structure and activity of a linuron-degrading triple-species biofilm. *Res Microbiol* **161**: 660-666.
Breugelmans, P., Barken, K.B., Tolker-Nielsen, T., Hofkens, J., Dejonghe, W., and Springael, D. (2008) Architecture and spatial organization in a triple-species bacterial biofilm synergistically degrading the phenylurea herbicide linuron. *FEMS Microbiol Ecol* **64**: 271-282.
Dealtry, S., Nour, E.H., Holmsgaard, P.N., Ding, G.C., Weichelt, V., Dunon, V. et al. (2016) Exploring the complex response to linuron of bacterial communities from biopurification systems by means of cultivation-independent methods. *FEMS Microbiol Ecol* **92**.
Dejonghe, W., Berteloot, E., Goris, J., Boon, N., Crul, K., Maertens, S. et al. (2003) Synergistic degradation of linuron by a bacterial consortium and isolation of a single linuron-degrading *Variovorax* strain. *Appl Environ Microbiol* **69**: 1532-1541.
diCenzo, G.C., and Finan, T.M. (2017) The Divided Bacterial Genome: Structure, Function, and Evolution. *Microbiology and Molecular Biology Reviews* **81**.
diCenzo, G.C., MacLean, A.M., Milunovic, B., Golding, G.B., and Finan, T.M. (2014) Examination of prokaryotic multipartite genome evolution through experimental genome reduction. *PLoS Genet* **10**: e1004742.
Haest, P.J., Philips, J., Springael, D., and Smolders, E. (2011) The reactive transport of trichloroethene is influenced by residence time and microbial numbers. *J Contam Hydrol* **119**: 89-98.
Harrison, P.W., Lower, R.P., Kim, N.K., and Young, J.P. (2010) Introducing the bacterial 'chromid': not a chromosome, not a plasmid. *Trends Microbiol* **18**: 141-148.
Horemans, B., Smolders, E., and Springael, D. (2013) Carbon source utilization profiles suggest additional metabolic interactions in a synergistic linuron-degrading bacterial consortium. *FEMS Microbiol Ecol* **84**: 24-34.
Horemans, B., Albers, P., and Springael, D. (2016a) The Biofilm concept from a bioremediation perspective. In *Biofilms in Bioremediation*. Lear, G. (ed). Poole, UK: Caister Academic Press.
Horemans, B., Hofkens, J., Smolders, E., and Springael, D. (2014) Biofilm formation of a bacterial consortium on linuron at micropollutant concentrations in continuous flow chambers and the impact of dissolved organic matter. *FEMS Microbiol Ecol* **88**: 184-194.
Horemans, B., Bers, K., Ruiz-Romero, E., Pose-Juan, E., Dunon, V., De Mot, R., and Springael, D. (2016b) Functional redundancy of linuron degradation in microbial communities of agricultural soil and biopurification systems. *Appl Environ Microbiol*: accepted.
Jacques, R.J., Okeke, B.C., Bento, F.M., Teixeira, A.S., Peralba, M.C., and Camargo, F.A. (2008) Microbial consortium bioaugmentation of a polycyclic aromatic hydrocarbons contaminated soil. *Bioresour Technol* **99**: 2637-2643.
Krumsiek, J., Arnold, R., and Rattei, T. (2007) Gepard: a rapid and sensitive tool for creating dotplots on genome scale. *Bioinformatics* **23**: 1026-1028.

Larrain-Linton, J., De la Iglesia, R., Melo, F., and Gonzalez, B. (2006) Molecular and population analyses of a recombination event in the catabolic plasmid pJP4. *J Bacteriol* **188**: 6793-6801.
Larsen, M.H., Biermann, K., Tandberg, S., Hsu, T., and Jacobs, W.R., Jr. (2007) Genetic Manipulation of *Mycobacterium tuberculosis*. *Curr Protoc Microbiol* **Chapter 10**: Unit 10A.2.
Liang, Q., Takeo, M., Chen, M., Zhang, W., Xu, Y., and Lin, M. (2005) Chromosome-encoded gene cluster for the metabolic pathway that converts aniline to TCA-cycle intermediates in *Delftia tsuruhatensis* AD9. *Microbiology* **151**: 3435-3446.
Ng, J., and Wyndham, R.C. (2011) IS1071-mediated recombinational equilibrium in *Alcaligenes* sp. BR60 carrying the 3-chlorobenzoate catabolic transposon Tn5271. *Canadian Journal of Microbiology* **39**: 92-100.
Reasoner, D.J., and Geldreich, E.E. (1985) A new medium for the enumeration and subculture of bacteria from potable water. *Appl Environ Microbiol* **49**: 1-7.
Rodriguez-R, L.M., and Konstantinidis, K.T. (2016) The enveomics collection: a toolbox for specialized analyses of microbial genomes and metagenomes. *PeerJ Preprints* **4**: e1900v1901.
Sambrook, J., and Russell, D.W. (2001) *Molecular cloning : a laboratory manual*. Cold Spring Harbor, N.Y.: Cold Spring Harbor Laboratory Press.
Seemann, T. (2014) Prokka: rapid prokaryotic genome annotation. *Bioinformatics* **30**: 2068-2069.
Sota, M., Yano, H., Nagata, Y., Ohtsubo, Y., Genka, H., Anbutsu, H. et al. (2006) Functional analysis of unique class II insertion sequence *IS1071*. *Appl Environ Microbiol* **72**: 291-297.
Springael, D., Kreps, S., and Mergeay, M. (1993) Identification of a catabolic transposon, *Tn4371*, carrying biphenyl and 4-chlorobiphenyl degradation genes in *Alcaligenes eutrophus* A5. *J Bacteriol* **175**: 1674-1681.
Untergasser, A., Cutcutache, I., Koressaar, T., Ye, J., Faircloth, B.C., Remm, M., and Rozen, S.G. (2012) Primer3--new capabilities and interfaces. *Nucleic Acids Res* **40**: e115.
Wick, R.R., Judd, L.M., Gorrie, C.L., and Holt, K.E. (2017) Unicycler: Resolving bacterial genome assemblies from short and long sequencing reads. *PLoS Comput Biol* **13**: e1005595.
Zerbino, D.R., and Birney, E. (2008) Velvet: algorithms for de novo short read assembly using de Bruijn graphs. *Genome Res* **18**: 821-829.

Figure legends

Fig. 1. Organization of the genomic regions of *Variovorax* sp. WDL1 associated with linuron hydrolysis and DCA oxidation. Gene cluster *dcaQTA1A2BR** coding for DCA oxidation is located on contig 10 (Bers *et al*., 2013). The originally described contig 1 containing the linuron hydrolase gene *hylA* (Bers *et al*., 2013) was extended, creating contig 1.2 (sequence deposited in GenBank under accession number KU973601). Black and grey boxes below the contigs represent

respectively qPCR and conventional PCR target regions using primer sets listed in supplemental Table S1.

Fig. 2. Phenotypic diversification of *Variovorax* sp. WDL1 during growth on liquid medium (R2B) containing linuron. (A) Frequencies of occurrence (as percentages) of the three different linuron/DCA degradation phenotypes observed by testing 44 colonies grown from the original *Variovorax* sp. WDL1 cryostock after plating on R2A medium supplemented with 20 mg L^{-1} linuron. (B) Frequencies of occurrence (as percentages) of the degradation phenotypes observed for 20 colonies each of the three catabolic variants after plating on linuron-supplemented R2A medium.

Fig. 3. Abundances of the *Variovorax* sp. WDL1 genomic fragments carrying *dcaQTA1A2BR** or *dcaQ*-hylA junction* in different linuron/DCA degradation variants, expressed as log-ratio of marker gene copies to copies of the chromosomal gene *omp32*. The position of the quantified amplicons is shown in Figure 1. The data points and error bars represent the means and standard deviations based on data from at least four biological replicates.

Fig. 4. Schematic representation of a 800-1000 kb region of the 1380 kb extrachromosomal replicon carrying the *hylA* gene in strain B2 (representing the Lin+ DCA- subpopulation of WDL1) and the *dcaQTA1A1BR** gene cluster in strain B4 (representing the Lin- DCA+ WDL1 subpopulation). Alignment of the contigs in B2 and B4 reveals identical sequence makeup except for the region of interest that is highlighted in red. Features of interest found in the annotation are highlighted such as defense systems, metal resistance genes, catabolic genes different from the *hylA/dcaQTA1A2BR* genes and IS*1071* elements. The region in purple at the

right-hand side contains catabolic genes extraneous to the *hylA/dcaQTA1A2BR** genes, including the *ccdCDEF* genes responsible for chlorocatechol degradation. Regions that show homology to plasmid maintenance and transfer functions are indicated as P1 to P4 regions with details provided in Table S3.

Fig. 5. Degradation of linuron (open symbols) and accumulation of DCA (filled symbols) in three consortium biofilms composed of strains WDL1, WDL6 and WDL7 from the first biofilm experiment (A) and abundances of gene markers for *hylA*, *dcaQ*-hylA* junction and *dcaQ* in DNA extracted from the consortium biofilms at harvest after 39 days of irrigation (B). Abundances are expressed as log-ratio of marker gene copies to copies of chromosomal gene *omp32.* In all panels, data points and error bars represent the mean and standard deviations for consortium biofilms (n=3).

Supporting information

Fig. S1. Postulated pathway of linuron degradation by the triple species consortium.

Fig. S2. Agarose gel electrophoresis of PCR amplicons of *dcaT* (A) and *dcaA1* (B) in cryocultures of the different linuron degradation variants. Primers are listed in Table S1. Lanes corresponding to the different linuron degradation variants are indicated above the gel, Lane L: Generuler 100bp plus DNA ladder (Thermo Scientific, USA), Lane +: PCR positive control.

Fig. S3. Replicons in strains B2 and B4. Genomic organization and subsequent assignment of the contigs to their respective components (replicons) was done by inspecting the assembly graph produced by the assembler using Bandage (Wick et al., 2015). Bandage allows to visualize assembly graph and displays potential connections between contigs. The thin lines in the replicon pictures indicate positions where connection between contigs has multiple possibilities. The assembly graphs of both subpopulations show identical genomic architecture consisting of 6 replicons and for both strains all replicons included the same sequences. The numbers indicated on the different components of the replicons correspond to the contig number that can be found in the respective assemblies. Details of each component are found in table S4.

Fig. S4. Agarose gel electrophoresis of PCR amplicons of *dcaT* (A) and *dcaA1* (B) in cryocultures of the different linuron degradation variants. Primers are listed in Table S1. Lanes corresponding to the different linuron degradation variants are indicated above the gel, Lane L: Generuler 100bp plus DNA ladder (Thermo Scientific, USA), Lane +: PCR positive control.

Fig. S5. Dot plots between the contigs of the different components (2 chromosomes and 4 plasmids) of the genomes of strains B2 and B4, displaying sequence identity. Inversions and discontinuity in the dot plots are a consequence of the de novo assembly process.

Fig. S6. Degradation of linuron (open symbols) and transient accumulation of DCA (filled symbols) by consortium biofilms composed of strains WDL1, WDL6 and WDL7 in the second biofilm experiment. Data points and error bars represent the mean concentration values and standard deviations of triplicate measurements.

Table S1. Primers used in this study. Target regions for *hylA*, *dcaQ*-hylA* junction, *dcaQ*, *dcaT* and *dcaA1* are indicated in Figure 1.

Table S2. Abundances ± SD of contigs 1.2 and 10 as determined by qPCR targeting the indicated catabolic gene markers in different degradation variants of WDL1. Abundances are expressed as log-ratio of marker gene copies to copies of reference gene *omp32* in the same sample. "Contig 1.2 + contig 10" represents the log-ratio of the sum of copies of the *dcaQ*-hylA* junction and *dcaQ* to number of copies of *omp32*. For each gene marker, different letters in bold indicate significant differences in abundance between the degradation phenotypes (one way ANOVA, p<0,05).

Table S3. Details of the regions P1 to P4 containing plasmid related genes as indicated in Figure 4. For each entry, the locus tag and the coordinates of the gene on both subpopulation contigs is given, together with the annotation, and the protein sequence.

Table S4. GC content and length of the contigs composing the different replicons recognized in strains B2 and B4 representing the Lin+DCA- and Lin-DCA+ subpopulations of WDL1, respectively as well as average GC content and length of the replicons.

Table S5. Abundances ± SD of the catabolic gene markers in the WDL1 population in consortium biofilms. Abundances are expressed as log-ratio of marker gene copies to copies of reference gene *omp32* in the same sample. Contig 1.2 + contig 10 represents the log-ratio of the sum of copies of the *dcaQ*-hylA* junction and *dcaQ* gene to the number of copies of *omp32*.

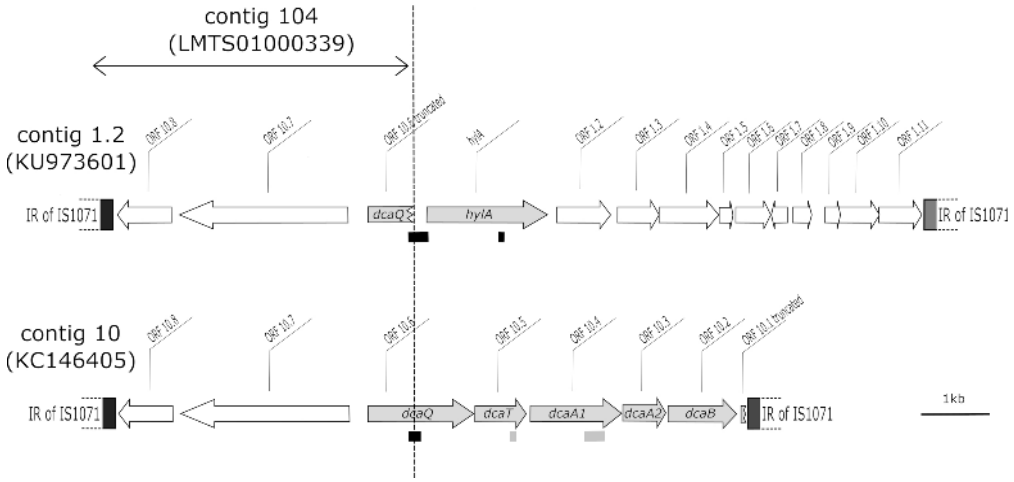

Fig. 1. Organization of the genomic regions of Variovorax sp. WDL1 associated with linuron hydrolysis and DCA oxidation. Gene cluster dcaQTA1A2BR* coding for DCA oxidation is located on contig 10 (Bers et al., 2013). The originally described contig 1 containing the linuron hydrolase gene hylA (Bers et al., 2013) was extended, creating contig 1.2 (sequence deposited in GenBank under accession number KU973601). Black and grey boxes below the contigs represent respectively qPCR and conventional PCR target regions using primer sets listed in supplemental Table S1.

200x96mm (300 x 300 DPI)

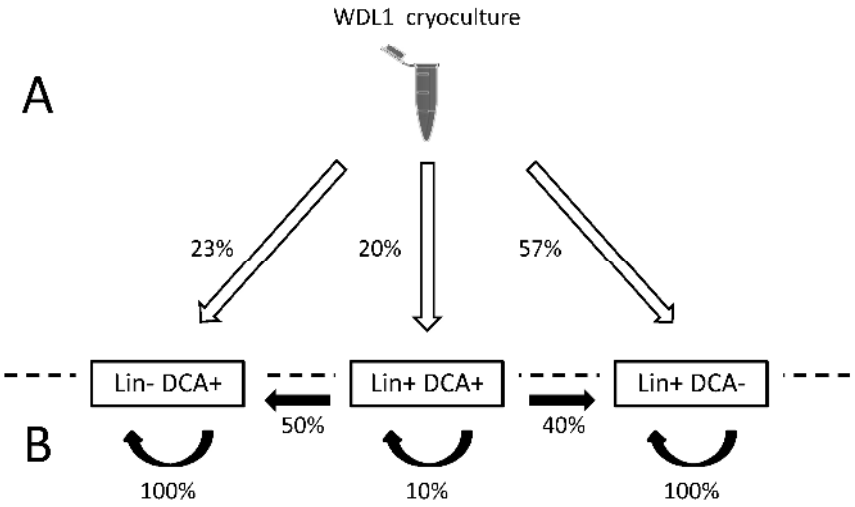

Fig. 2. Phenotypic diversification of Variovorax sp. WDL1 during growth on liquid medium (R2B) containing linuron. (A) Frequencies of occurrence (as percentages) of the three different linuron/DCA degradation phenotypes observed by testing 44 colonies grown from the original Variovorax sp. WDL1 cryostock after plating on R2A medium supplemented with 20 mg L-1 linuron. (B) Frequencies of occurrence (as percentages) of the degradation phenotypes observed for 20 colonies each of the three catabolic variants after plating on linuron-supplemented R2A medium.

254x190mm (300 x 300 DPI)

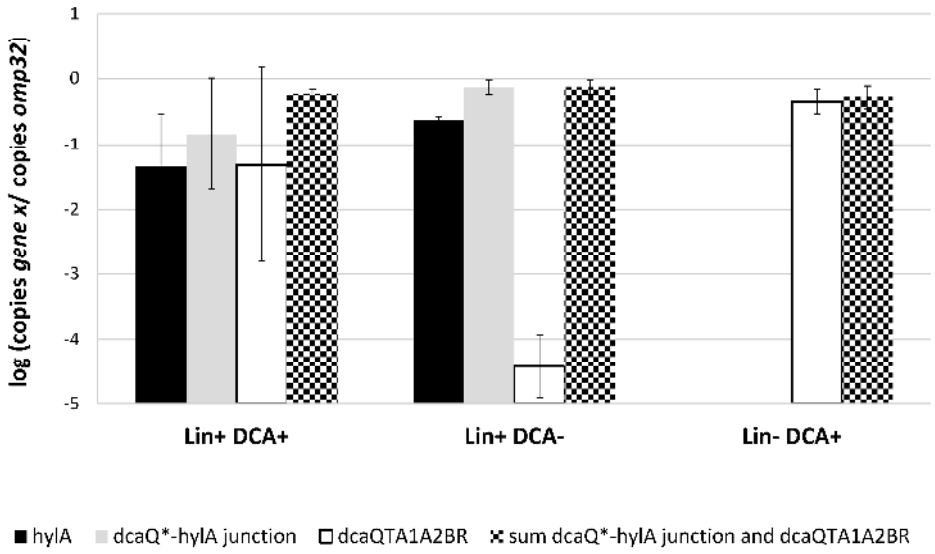

Fig. 3. Abundances of the Variovorax sp. WDL1 genomic fragments carrying dcaQTA1A2BR* or dcaQ*-hylA junction in different linuron/DCA degradation variants, expressed as log-ratio of marker gene copies to copies of the chromosomal gene omp32. The position of the quantified amplicons is shown in Figure 1. The data points and error bars represent the means and standard deviations based on data from at least four biological replicates.

254x190mm (300 x 300 DPI)

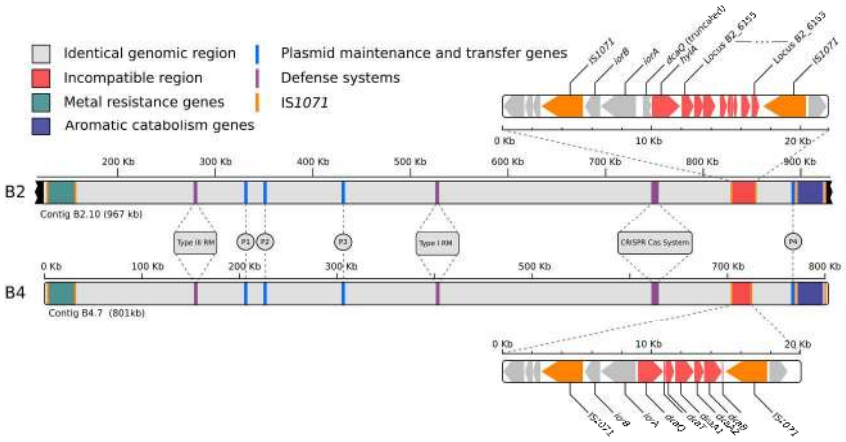

Fig. 4. Schematic representation of a 800-1000 kb region of the 1380 kb extrachromosomal replicon carrying the hylA gene in strain B2 (representing the Lin+ DCA- subpopulation of WDL1) and the dcaQTA1A1BR* gene cluster in strain B4 (representing the Lin- DCA+ WDL1 subpopulation). Alignment of the contigs in B2 and B4 reveals identical sequence makeup except for the region of interest that is highlighted in red. Features of interest found in the annotation are highlighted such as defense systems, metal resistance genes, catabolic genes different from the hylA/dcaQTA1A2BR genes and IS1071 elements. The region in purple at the right-hand side contains catabolic genes extraneous to the hylA/dcaQTA1A2BR* genes, including the ccdCDEF genes responsible for chlorocatechol degradation. Regions that show homology to plasmid maintenance and transfer functions are indicated as P1 to P4 regions with details provided in Table S3.

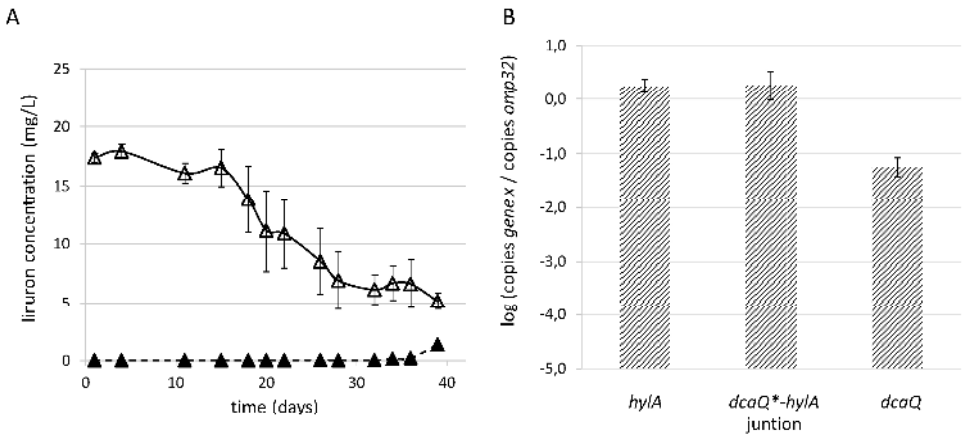

Fig. 5. Degradation of linuron (open symbols) and accumulation of DCA (filled symbols) in three consortium biofilms composed of strains WDL1, WDL6 and WDL7 from the first biofilm experiment (A) and abundances of gene markers for hylA, dcaQ*-hylA junction and dcaQ in DNA extracted from the consortium biofilms at harvest after 39 days of irrigation (B). Abundances are expressed as log-ratio of marker gene copies to copies of chromosomal gene omp32. In all panels, data points and error bars represent the mean and standard deviations for consortium biofilms (n=3).

254x190mm (300 x 300 DPI)