Ding, C., Rogers, M.J., Yang, K.-L., He, J. (2017):
Loss of the *ssrA* genome island led to partial debromination in the PBDE respiring *Dehalococcoides mccartyi* strain GY50
Environ. Microbiol. **19** (7), 2906 – 2915

Loss of the *ssrA* genome island led to partial debromination in the PBDE respiring *Dehalococcoides mccartyi* strain GY50

Chang Ding [1†], Matthew J. Rogers [1], Kun-Lin Yang [2], Jianzhong He [1*]

[1] Department of Civil and Environmental Engineering,

[2] Department of Chemical and Biomolecular Engineering,

National University of Singapore, Singapore, 117576

[†] Present address: Department of Isotope Biogeochemistry,

Helmholtz-Centre for Environmental Research - UFZ, Leipzig, Germany, 04318

* Corresponding author:

Department of Civil and Environmental Engineering

National University of Singapore

Block E2-02-13

1 Engineering Drive 3, Singapore 117576

Phone: (65) 6516-3385; Fax: (65) 6774-4202

Email: jianzhong.he@nus.edu.sg

Running Title: Strain GY50 respires on PBDEs

Keywords: reductive dehalogenation, *Dehalococcoides*, polybrominated diphenyl ether, bioremediation, *ssrA*-specific genome island

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: 10.1111/1462-2920.13817

Summary

Polybrominated diphenyl ethers (PBDEs), chemicals commonly used as flame-retardants in consumer products, are emerging persistent organic pollutants that are ubiquitous in the environment. In this study, we report the PBDE-respiring isolate - *Dehalococcoides mccartyi* strain GY50, which debrominates the most toxic tetra- and penta-BDE congeners (~1.4 µM) to diphenyl ether within 12 days with hydrogen as the electron donor. The complete genome sequence revealed 26 reductive dehalogenase homologous genes (*rdhA*s), among which three genes (*pbrA1*, *pbrA2*, and *pbrA3*) were highly expressed during PBDE debromination. After ten transfers of GY50 with trichloroethene or 2,4,6-trichlorophenol as the electron acceptor instead of PBDEs, the *ssrA*-specific genome island (*ssrA*-GI) containing *pbrA1* and *pbrA2* was deleted from the genome of strain GY50, leading to two variants (strain GY52 with trichloroethene, strain GY55 with 2,4,6-trichlorophenol) with identically impaired debromination capabilities (debromination of penta-/tetra-BDEs ceased at di-BDE 15). Through analysis of Illumina paired-end sequencing data, we identified read pairs that probably came from variants that contain *ssrA*-GI deletions, indicating their possible presence in the original strain GY50 culture. The two variant strains provide real-time examples on rapid evolution of organohalide-respiring organisms. As PBDE-respiring organisms, GY50-like strains may serve as key players in detoxifying PBDEs in contaminated environments.

Introduction

The widespread usage of polybrominated diphenyl ethers (PBDEs) as flame retarding materials since the 1960s has led to their ubiquitous presence in various biotas (Norén and Meironyté, 2000; Hites, 2004; Tanabe, 2008). Due to concerns over their toxicity and bioaccumulating behavior, tetra- to hepta-BDEs were recently classified as "Chemicals of Concern" by the United States Environmental Protection Agency and as Persistent Organic Pollutants (POPs) by the United Nations (UNEP, 2009; USEPA, 2009). Since the first observation of the toxic effects of PBDEs two decades ago (Fowles *et al.*, 1994), people have been looking for efficient ways to detoxify these recalcitrant compounds. Bioremediation via reductive debromination by anaerobic bacteria could be a viable solution for PBDE contamination because it takes place in the subsurface anoxic environment and has few side effects to the ecological system (Iwamoto and Nasu, 2001). Several microcosms and cultures have been found to debrominate PBDEs (He *et al.*, 2006; Robrock *et al.*, 2008; Tokarz *et al.*, 2008; Lee and He, 2010; Lee *et al.*, 2011; Ding *et al.*, 2013; Pan *et al.*, 2017). However, most of them debrominate PBDEs slowly and incompletely to lower brominated congeners that are even more toxic. The only exceptional case is a co-culture GY2 consisting of *Dehalococcoides mccartyi* and *Desulfovibrio*, which is able to debrominate tetra- and penta-BDEs completely to diphenyl ether (Lee *et al.*, 2011).

So far, no reductive dehalogenase homologous genes (*rdhAs*) have been discovered for PBDE debromination. PBDEs were usually dosed at very low aqueous concentrations (<1 μM) in reported debrominating cultures due to their high hydrophobicity. Therefore, the effects of PBDEs on the transcriptome/proteome of the debrominators in mixed cultures could hardly be observed, leading to difficulties in

identifying genes/enzymes responsible for debromination. The discovery of the co-culture GY2 offered an opportunity to identify these genes, because the *Dehalococcoides* cells in the co-culture GY2 are responsible for the debromination activity, and their growth are closely coupled to the debromination process (Lee *et al.*, 2011). The strict dependence on PBDEs for cell growth would greatly facilitate the identification of debromination genes in the *Dehalococcoides* species.

In this study, *D. mccartyi* strain GY50 was isolated from the co-culture GY2, and retains the same debromination activities as its parent culture – debrominating tetra-BDE 47 and penta-BDEs 99 & 100 to non-brominated diphenyl ether, via tri-BDE 27, di-BDE 15, and mono-BDE 4 as the intermediates. The complete genomes of strain GY50 and its two variant strains were sequenced and revealed an interesting genome island deletion event that leads to the two variants. Subsequent transcriptomic and proteomic analyses were performed to identify novel *rdhA* genes that potentially catalyze debromination reactions.

Results

Isolation of *D. mccartyi* strain GY50 and emergence of two variant strains

Isolation of the *Dehalococcoides* strain in co-culture GY2 by serial dilution was not feasible under normal cultivation conditions because the cell density of *Dehalococcoides* was an order of magnitude lower than that of *Desulfovibrio* (1.02×10^6 cells/mL versus 4.83×10^7 cells/mL) (Lee *et al.*, 2011). Therefore, we subcultured GY2 repeatedly in 1 g/L ampicillin in order to inhibit *Desulfovibrio*. The 16S-DGGE showed that *Dehalococcoides* became the major population in the ampicillin treated GY2 culture (Figure 1). After three serial dilutions in ampicillin, a pure culture of

Dehalococcoides was obtained, as confirmed by both microscopic observation and DGGE (Fig. S1). This isolate was designated *D. mccartyi* strain GY50 belonging to the Victoria subgroup based on the 1,484/1,485 bp identity in the 16S rRNA gene sequence with *D. mccartyi* strain VS. Strain GY50 grows at pH 7.2-7.4 and 30°C. It debrominates tetra-BDE 47 and penta-BDEs 99 and 100 to diphenyl ether via tri-BDE 28, di-BDE 15, and mono-BDE 4 as the intermediates, in a defined mineral salt medium (with vitamins) amended with acetate and hydrogen as the carbon source and the electron donor, respectively.

Besides PBDEs, GY50 also dechlorinates TCE and 2,4,6-TCP, with *cis*-dichloroethene (*cis*-DCE) and 2,4-dichlorophenol (2,4-DCP) as the final products, respectively. Interestingly, after ten transfers of strain GY50 fed with TCE or TCP instead of PBDEs, the resultant strains partially or completely lost debromination capabilities. Transfers of GY50 with TCE generated a variant strain GY52 that can only debrominate tetra- and penta-BDEs to di-BDE 15, while transfers with TCP yielded strain GY55 with no detectable debromination activity after inoculating 3% GY55 grown on TCP (100 μM) to tetra- and penta-BDEs spiked medium. The purity of strains GY52 and GY55 were further confirmed by picking colonies from agar shake tubes amended with TCE and TCP, respectively.

Dehalogenation profiles of *D. mccartyi* strains GY50, GY52, and GY55

Strain GY50 completely debrominated tetra-BDE 47 and penta-BDEs 99 & 100 (~1.4 μM in sum) to diphenyl ether in 12 days (Figure 2A). The cell number of strain GY50 increased 28 fold to 5.2×10^6 cells/mL on day 7, and then slowly drops to 2.9×10^6 cells/mL on day 12. The cell yield from PBDE debromination was approximately $(9.20 \pm 0.06) \times 10^{14}$ cell mole of bromine released. Strain GY50 also dechlorinated 100

µM of 2,4,6-TCP to 2,4-DCP within 11 days (Fig. S1), and 714 µM of TCE to predominantly *cis*-DCE within 15 days (Fig. S2). The cell yields from TCP and TCE were (4.08±0.09) × 10^{14} and (3.45±0.12) × 10^{14} cell mole of chlorine released, respectively. Strain GY52 debrominated tetra-BDE 47 and penta-BDEs 99 & 100 (~1.4 µM in sum) to di-BDE 15 as the only debromination product within 11 days (Figure 2B). The corresponding cell yield was (9.24±0.40) × 10^{14} cell mole of bromine released, or (3.63±0.16) g cell / mole of bromine released.

Surprisingly, strain GY55 did not show debromination activity on tetra-BDE 47 and penta-BDEs 99 & 100 as its parent strain GY50 does. Later, it was found that the residual DCP (~3-4 µM) in the medium, which was carried over from the inoculum culture grown on 100 µM of TCP, inhibited strain GY55's debromination activity when fed with tetra- and penta-BDEs. In the presence of 4 µM of DCP, PBDE debromination by strain GY50 or GY52 was also completely inhibited (Fig. S3). When strain GY55 grown in TCP was serially diluted and transferred to PBDE-spiked medium again to minimize the effect of residual DCP, strain GY55 did show the same partial debromination activity as strain GY52. Apart from the difference in PBDE debromination, these two strains perform similarly as strain GY50 does in dechlorinating TCP and TCE.

Complete genome sequence of strain GY50

The complete genome sequence of strain GY50 (NCBI accession number CP006730) contains 1,407,418 base pairs encoding 1,519 predicted proteins (Table S1). The genome of strain GY50 has the closest relationship with the genome of strain CG1 based on average nucleotide identity (ANI) (Konstantinidis and Tiedje, 2005; Goris *et*

al., 2007) and *in silico* DNA-DNA Hybridization (DDH) values (Meier-Kolthoff *et al.*, 2013) (Table 1, Table S2).

A total of 26 reductive-*dehalogenase*-homologs (*rdhA*s) were identified in the genome of strain GY50 (Table S3). Similar to other *Dehalococcoides* strains (Kube *et al.*, 2005; McMurdie *et al.*, 2009), strain GY50 harbors two high plasticity regions (HPRs) at both sides of the replication origin (HPR1: 57,513-233,636 bp, HPR2: 1,198,936-1,361,570 bp), where 25 of the 26 *rdhAs* reside. Six putative mobile genetic elements (MGEs) were identified in strain GY50 (Table S4) with three MGEs located in the HPRs. The largest mobile element – MGE V has the typical characteristics of the *ssrA*-specific genome island (*ssrA*-GI) as described previously (McMurdie *et al.*, 2011), including its position in the genome (near the transporter-messenger RNA gene *ssrA*), its size, as well as its genetic components. The second largest mobile element – MGE II spans from 628.5 kbp to 665.3 kbp, and splits a large piece of conserved region which is continuous in most other *Dehalococcoides* genomes (Fig. S4). Both MGEs show abnormal dinucleotide and trinucleotide signatures, indicating their foreign origins (Fig. S5).

Deletion of the *ssrA*-specific genome island in strains GY52 and GY55

The full genomes of strains GY52 and GY55 were also sequenced and compared with the genome of strain GY50. Interestingly, the MGE V (*ssrA*-GI) (McMurdie *et al.*, 2011) was deleted in both genomes (27,166 bp/32 genes in strain GY52, 27,136 bp/32 genes in strain GY55) (Figure 3, Table S5). The deleted regions in the GY52 and GY55 genomes are 2.2 kbp and 5.2 kbp downstream of the *ssrA* gene, respectively. The *ssrA*-GIs are MGEs commonly observed in *Dehalococcoides* strains, which have boundaries with unique characteristics, i.e., having the *ssrA* gene at the 5' terminal and a 20-bp

direct repeat of *ssrA*'s 3' end at the 3' terminal (*ssrA* repeat) (McMurdie *et al.*, 2011). There are three such *ssrA* repeats in strain GY50 at 17.6 kbp, 27.1 kbp, and 47.4 kbp downstream of the *ssrA* gene, which are probably remnants of multiple *ssrA*-specific integration events. Notably, the deleted *ssrA*-GIs in strains GY52 and GY55 are surrounded by different sets of repeated sequences (repeated sequences 1 and 2 as shown in Figure 3). Therefore, in contrast to the strict site-specific mechanism of *ssrA*-GI integration (Williams, 2003), the deletions of the *ssrA*-GIs are less site-specific, and are not associated with the 20-bp *ssrA* repeats.

Most importantly, the deleted *ssrA*-GI regions of strains GY52 and GY55 contain two *rdhA* genes – GY50_1350 (designated *pbrA1*) and GY50_1352 (designated *pbrA2*), while other deleted genes include those annotated as DNA modification proteins, a CbiZ protein (for vitamin B_{12} salvaging) (Woodson and Escalante-Semerena, 2004), and hypothetical proteins (Table S5). Apart from the deletions at *ssrA*-GI, there are only a few single nucleotide polymorphisms (SNPs) in the genomes of strains GY52 (7 SNPs) and GY55 (2 SNPs), compared to the genome of strain GY50 (Table S6). This indicates that at least one of the deleted genes *pbrA1* and *pbrA2* catalyzes debromination from di-BDE to diphenyl ether.

Evidence for existence of variants with *ssrA*-GI deletions in strain GY50

The fact that variants with *ssrA*-GI deletions became dominant in only ten transfers raised the question whether the variants existed already in the original culture of strain GY50. Because the excision points of deletions are within highly similar sequences (repeated sequences 1 and 2, Figure 3), it is difficult to detect presence of variants that contain *ssrA*-GI deletions by conventional PCR, due to probable formation of PCR chimera. Instead, we looked into the Illumina paired-end sequencing data of

strain GY50 (pair end library, 473 bp average insert size, 90 bp read length, ~400x sequencing depth) and attempted to obtain read pairs that indicate signs of deletions.

The genome region of 1224191-1259367 that covers repeated sequences 1 and 2 as well as 1000 bp upstream and downstream (total length: 35,177 bp) was chosen for this analysis. Through two rounds of BLASTN search, a total of 76,640 correctly oriented read pairs were identified that match to this region better than any other regions of the GY50 genome (Supplementary method). Among these 76,640 read pairs, six read pairs have apparent insert sizes >27 kbp. Of the six read pairs, three (read pairs No. 2, 3, and 5) are probably resulted from *ssrA*-GI deletions (Table S7), based on the criteria that 1) two reads within one read pair match to repeated sequence 1a and 1b (or 2a and 2b) respectively (a and b refer to the two copies of the repeated sequences 1 and 2) and therefore exhibit an apparent insert size of > 27 kbp, and that 2) for the two reads that match criterion 1, the best and second best BLASTN hits in the genome of GY50 (in repeated sequence 1a and 1b (or 2a and 2b)) have at least three mismatches (Table S8).

As shown in Table S7, read pair 2 should be from mutants where excision points are located within repeated sequence 1, while read pairs 3 and 5 are from the mutants where excision points are within repeated sequence 2. None of the three read pairs have the known GY52/GY55 excision points in between, and therefore they are new mutants other than strains GY52/GY55.

Identification of RdhAs that are responsible for PBDE debromination

In order to identify the RdhAs that are responsible for PBDE debromination, the transcription levels of all 26 *rdhA* genes were monitored during PBDE debromination

by strain GY50 and strain GY52. Among the *rdhA* genes, *pbrA1* and *pbrA2* were highly transcribed in strain GY50, while GY50_1427 (designated *pbrA3*) was highly transcribed in both strains GY50 and GY52 (Figure 4). Transcript numbers per cell of the three genes reached their maximum points (27.5 ±0.2 for *pbrA1* in GY50, 14.6 ±0.8 for *pbrA2* in GY50, 28.5 ±1.0 for *pbrA3* in GY50, and 66.6 ±6.7 for *pbrA3* in GY52) three days after the addition of PBDEs. The transcription levels of the remaining *rdhA* genes were all below 2 transcripts per cell in both strain GY50 and strain GY52. To confirm the protein translation of the transcribed *rdhA* genes, whole cell proteins of strain GY50 were harvested and subjected to SDS-PAGE (Fig. S6). Three most prominent bands were excised and sent for MALDI-TOF-MS analysis using a customized protein digestion database based on the GY50 genome (Table S9). In total, four proteins were identified in these three bands: a BNR/Asp-box repeat-containing protein (GY50_1244) which was believed to be an S-layer component (Morris *et al.*, 2006), a membrane-bound Ni/Fe hydrogenase (HupL, GY50_0142), a chaperonin GroEL (GY50_1266), and the *rdhA*s PbrA1. PbrA2 and PbrA3 were not detected in these three bands. To achieve higher sensitivity, nano-liquid chromatography (LC)-MS-MS was performed to capture the proteomes of strains GY50 and GY52. In total, 178 and 154 proteins were detected in strain GY50 and strain GY52, respectively (confidence level > 95%). PbrA1, PbrA2, and PbrA3 were the only three RdhAs detected in the proteome of strain GY50 (Table 2), while PbrA3 was the most abundant RdhA in the proteome of strain GY52 (Table 3).

Discussion

In this study, the PBDE-respiring microbe, *D. mccartyi* strain GY50 was isolated to be able to rapidly and completely debrominate tetra- and penta- BDE

congeners 47, 99, and 100. The three congeners as substrates are the most environmentally abundant, toxic, and bioaccumulative PBDE congeners among all (de Wit, 2002; Hites, 2004; Burreau *et al.*, 2006), making strain GY50 environmentally relevant. In most previous studies, microbial debromination of PBDEs was slow and incomplete, partially because most of the debrominating cultures co-metabolize PBDEs rather than utilizing them as energy sources (He *et al.*, 2006; Robrock *et al.*, 2008; Tokarz *et al.*, 2008). The discovery of the strain GY50 that respires on PBDEs leads to rapid detoxification of these recalcitrant halogenated compounds.

Three *rdhA* genes (*pbrA1*, *pbrA2*, and *pbrA3*) in strain GY50 were highly expressed as shown by transcriptional and proteomic analysis, and are probably involved in debromination of PBDEs. The comparative genomic analysis of strain GY50 and its variant strains GY52 and GY55 suggests that the absence of the genes in the *ssrA*-GI (most likely *pbrA1* and *pbrA2*) caused strain GY52/GY55 to lose its capability to debrominate di-BDE to diphenyl ether (*para*-debromination). In other words, *pbrA1* and/or *pbrA2* are probably involved in the last *para*-debromination of di-BDE to diphenyl ether, while the other highly transcribed *rdhA* gene - *pbrA3* gene could be responsible for debromination from tetra-/penta-BDEs to di-BDE (*ortho*- and *meta*-debromination). Due to the extremely low biomass of strain GY50 grown on PBDEs, we could not verify the *in vitro* activity of these gene products on native protein gels.

The amino acid sequences of PbrA1 and PbrA2 show high similarity to the RdhAs X792_07240 (encoding PcbA1, similar to PbrA1), ASJ33_07610 (similar to PbrA2), and X792_07250 (similar to PbrA2) in the PCB-dechlorinating strains CG1 and CG3 (Wang *et al.*, 2014), respectively (>99% amino acid (aa) identity), and share very low similarity with all other RdhAs (PbrA1, 43% aa identity with DET0311 of strain

195; PbrA2, 38% aa identity with cbdb_A80 of strain CBDB1). PbrA3 is highly similar to DhcVS_1421 in strain VS and X792_07710 in strain CG1 with only one amino acid difference (474/475 aa identity). PbrA3 also shows similarity to RdhAs in strain CBDB1 (cbdb_A1618, 454/475 aa identity), strain GT (DehalGT_1338, 454/475 aa identity), and strain BTF08 (btf_1481, 452/475 aa identity), but only shows 43% similarity to DET0876 in strain 195 which was reported to debrominate PBDEs co-metabolically (He *et al.*, 2006). Therefore, from a genetic view point, strains CG1, CG3, VS, CBDB1, GT, and BTF08 may also possess debromination capabilities. Indeed, strain CBDB1 was reported to debrominate brominated benzenes and two brominated phenolic compounds (tetrabromobisphenol A and bromophenol blue) (Wagner *et al.*, 2012; Yang *et al.*, 2015). Since PBDEs have not been reported to occur naturally and their anthropogenic production was initiated only decades ago, it is not known whether strain GY50 already carried the *pbrA1, pbrA2, and pbrA3* genes before the introduction of PBDEs to the environment, or evolved and developed debromination capabilities over a short period of time. No matter which case it is, the *rdhA* genes in *Dehalococcoides* strains may have greater dehalogenation capabilities that we currently know.

Although *ssrA* is widely regarded as the frequent target for MGE integration, it was not known how stable the integrated sequences are in the host genomes. There have been hints of *ssrA*-GI deletions in prokaryotes. By comparing 770 *Dichelobacter* strains isolated from different geographical sites, it was observed that 10 strains (1.3%) exhibited *ssrA*-GI deletions (Billington *et al.*, 1999). However, since these strains were all isolated independently, the real-time deletion events were not captured. In this study, the mutation of strain GY50 to strains GY52 and GY55 was observed after repeated

transfers of the same strain in TCE and TCP, which provides a real-time example of *ssrA*-GI deletion events. By sequencing the complete genomes of both variants, it is confirmed that the *ssrA*-GI were deleted in both genomes in merely ten transfers. The exact locations where the excision of the genome islands took place were also pinpointed. The deletion was probably initiated by recombinases that recognize repeated sequences, since both excision points are located in the repeated regions (Figure 3). We believe that deletion of the *ssrA*-GI is constantly taking place in strain GY50 when growing in PBDEs, leading to trace amounts of *ssrA*-GI-deficient cells in the total population. When TCP or TCE replaced PBDEs as the electron acceptor, PBDE-related *rdhA* genes and their associated RNAs/proteins became a metabolic burden to cells, and *ssrA*-GI-deficient cells may possess an advantage over wild type cells (Koskiniemi *et al.*, 2012). It is possible that these *ssrA*-GI-deficient cells could gradually out-competed the normal cells and eventually became dominant.

Attempts have been made to look for *ssrA*-GI-deficient cells in strain GY50 by looking at the paired-end Illumina sequencing data. Indeed, we found three read pairs that probably come from *ssrA*-GI-deficient cells and none of them share the same excision points as strains GY52/GY55. This suggests that the deletion events are relatively random along the repeated sequences. In order to further confirm the presence of *ssrA*-GI deletions in strain GY50, amplification-free genome sequencing is needed, but it is impractical given the low biomass of strain GY50 grown in PBDEs.

In summary, *D. mccartyi* strain GY50 was isolated in this study, which exhibited PBDE respiring capabilities that led to complete debromination of tetra- and penta-BDEs. Moreover, we successfully identified three *rdhA* genes (*pbrA1*, *pbrA2*, and *pbrA3*) in strain GY50, which can be used as biomarkers for *in situ* PBDE

debromination. The discovery of *D. mccartyi* strain GY50 sheds light on microbial detoxification of recalcitrant PBDEs in contaminated environments.

Experimental Procedures

Culture cultivation and sample analysis. Culture medium for strains GY50, GY52, and GY55 was prepared as described for co-culture GY2 (Lee *et al.*, 2011) with acetate and hydrogen as an carbon source and electron donor, respectively. Medium was reduced by 0.2 mM of $_{L}$-cysteine, 0.2 mM of $Na_2S{\cdot}9H_2O$ and 0.5 mM of $_{DL}$-dithiothreitol. A vitamin solution including 0.05 $mg{\cdot}L^{-1}$ vitamin B_{12} was also added to the medium (He *et al.*, 2007). PBDE congeners dissolved in isooctane (10 µL isooctane into 100 mL medium) were added to a total final nominal concentration of ~1.4 µM (727 ± 29 nM of BDE 47, 390 ± 21 nM of BDE 99, and 380 ± 15 nM of BDE 100). TCP dissolved in hexane (100 µL hexane into 100 mL medium) was added to a final nominal concentration of 100 µM. Neat TCE was added to a final nominal concentration of 714 µM in the medium. Culture bottles were incubated at 30 ºC in the dark without shaking. Autoclaved controls and un-inoculated controls were prepared to monitor possible abiotic dehalogenation. All kinetics studies were performed in triplicates. Isolation of GY52 and GY55 was achieved by picking colonies grown in serially diluted agar shake tubes amended with TCE and TCP (same concentrations as above), respectively. Agar shakes were prepared using the same medium recipe as liquid cultures except that 0.6% of low melting temperature agarose (SeaPlaque, Lonza) were added as the solidifying agent. Colonies appeared in dilutions of up to 10^{-7} after 2-3 months. Halogenated compound concentrations were determined by gas chromatography as described previously (Mun *et al.*, 2008; Cheng *et al.*, 2010; Lee *et*

al., 2011). Cell numbers were quantified by quantitative real-time PCR using primers GY50_1408F and GY50_1408R as mentioned below in the transcriptional analyses method section.

Genome sequencing and annotation. Due to the extremely low biomass of strain GY50 grown in PBDEs, genomic DNA from strain GY50 was amplified using REPLI-g mini kit (QIAGEN) with GY50 cells as the template, and purified using QIAamp DNA Mini kit (QIAGEN) according to the manufacturer's instructions. Genomic DNA from GY52 and GY55 was extracted using Genomic-Tip 100/G (QIAGEN) from cells grown in TCE and TCP, respectively. The genome of strain GY55 was sequenced by DOE Joint Genome Institute (CA, USA) using pair-end Illumina sequencing with an average insert size of 181 bp, while genomes of strain GY50 and GY52 were sequenced by BGI (Shenzhen, China) using pair-end Illumina sequencing with average insert sizes of 471 bp and 482 bp, respectively. Genomes were assembled by using SOAPdenovo package with optimized parameters (Li *et al.*, 2010). Gaps in the assembled genome of GY50 were closed by Sanger sequencing using primers specific to the beginning and end of each scaffold (Tettelin *et al.*, 1999). Sequencing errors were manually screened and corrected using BamView in the Artemis genome browser (Carver *et al.*, 2013). Open reading frames (ORFs) were identified using the online version of GLIMMER v3.02 (Delcher *et al.*, 1999), and compared with available *Dehalococcoides* annotations for perfect matches. The un-matched ORFs were then translated to protein sequences and searched with BLASTP against NCBI non-redundant protein database for annotations (Altschul *et al.*, 1997). Manual inspection of gaps between identified ORFs in the Artemis genome browser was done to screen for ORFs that were potentially missed. Annotations of protein

coding sequences were performed by searching against Pfam 27.0 (Punta *et al.*, 2012), NCBI CDD (Marchler-Bauer *et al.*, 2013), and KEGG database (Moriya *et al.*, 2007), and by referring to BLASTP results and annotations of known *Dehalococcoides* genomes. Ribosomal RNA, transport RNA, and transport-messenger RNA were identified using RNAmmer v1.2 (Lagesen *et al.*, 2007), tRNAscan-SE v1.3.1 (Lowe and Eddy, 1997), and ARAGORN v1.2.36 (Laslett and Canback, 2004). Protein molecular weight and theoretical pI were calculated using an online tool on web.expasy.org. Di- and tri-nucleotide signatures (Karlin, 1998) were calculated using a script written in Perl with a window size of 2000 bp and a step size of 1000 bp. Visualized genome comparison among fourteen *Dehalococcoides* genomes was done using Mauve v2.3.1 with Progressive Mauve algorithm and a seed weight of 15 (Darling *et al.*, 2004; Darling *et al.*, 2010).

Transcriptional analyses. At day 0, PBDEs were amended to cultures that had been starved of electron acceptors for two weeks. At each sampling point, 1.5 mL and 1 mL of culture (duplicates) were harvested for RNA and DNA extraction, respectively, by pelleting at 10,000 g, 4 ºC for 10 min. DNA extraction was done by using QIAGEN DNeasy Blood and Tissue kit as described previously (Cheng and He, 2009). RNA extraction and reverse transcription were done as described previously (Ding *et al.*, 2014), using the Trizol reagent (Invitrogen), the QIAGEN RNeasy mini column, and the QIAGEN Sensiscript RT kit. Real-time PCR was performed in 20-µL reactions using Bioline SensiFAST SYBR Lo-ROX Kit on an ABI 7500 FAST Real-Time PCR system. Cell numbers were calculated as the copy numbers of a single copy *rdhA* gene GY50_1408 in the genomes of the three *Dehalococcoides* strains, which were quantified using primers GY50_1408F and GY50_1408R. Luciferase cDNA copies

were quantified using primers LucF and LucR. Transcripts of individual *rdhA* genes were quantified using primers specifically designed based on their gene sequences. Primers used in this study are listed in Table S10. Numbers of gene transcripts per cell were obtained by normalizing transcript copies based on luciferase mRNA losses, and then divided by cell numbers at each time point.

Proteomic analyses. The culture of strain GY50 (800 mL) was harvested at the late exponential phase by pelleting at 14,000 g, 4 ºC for 15 min. The barely visible cell pellet was then re-suspended in 0.5 mL of lysis buffer (12% w/v of SDS, 0.6 M of DTT, 47 % v/v of glycerol, 60 mM of Tris at pH 6.8, and 0.06% w/v of bromophenol blue). Cells were disrupted for 5 min (working time) on ice with 5s/15s working/cooling pulse using a VCX 130 ultrasonicator (Sonics & Materials, USA). The resulted whole cell extracts were separated on a 10% SDS PAGE gel using Bio-Rad Mini-PROTEAN Tetra cell with 25 µL of extracts loaded in each lane in a total of 10 lanes. The gel was silver stained using PlusOne Silver Staining kit (GE Healthcare), and prominent bands were excised from all 10 lanes. Bands at the same position of the lanes were combined and sent for trypsin digestion and MALDI-TOF-MS analysis at Protein and Proteomics Centre in National University of Singapore. The acquired data were then searched against the strain GY50 genome using Mascot (Matrixscience) for significant hits. For nano-liquid chromatography (LC)-MS-MS analysis, late exponential phase cultures of strains GY50 and GY52 (1.8 L for each strain) were filtered through 0.2 µm pore size polycarbonate membrane filters (GE Water & Process Technologies). Cells were re-suspended in a buffer containing 10 mM Tris (pH 8.0) and 1 mM DTT, and then ultrasonicated for 5 min on ice (5s/15s working/cooling pulse). Resulted cell extracts were concentrated to 0.1 mL for each strain using Vivaspin molecular filter (5 kDa

cutoff) before trypsin digestion and analysis on an AB SCIEX TripleTOF 5600 system at Protein and Proteomics Centre in National University of Singapore. Results were then searched against the strain GY50 genome using ProteinPilot v4.5 beta for significant hits.

References

Altschul, S.F., Madden, T.L., Schaffer, A.A., Zhang, J.H., Zhang, Z., Miller, W., and Lipman, D.J. (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res* **25**: 3389-3402.

Billington, S.J., Huggins, A.S., Johanesen, P.A., Crellin, P.K., Cheung, J.K., Katz, M.E. *et al.* (1999) Complete nucleotide sequence of the 27-kilobase virulence related locus (*vrl*) of *Dichelobacter nodosus*: Evidence for extrachromosomal origin. *Infect Immun* **67**: 1277-1286.

Burreau, S., Zebuhr, Y., Broman, D., and Ishaq, R. (2006) Biomagnification of PBDEs and PCBs in food webs from the Baltic Sea and the northern Atlantic ocean. *Sci Total Environ* **366**: 659-672.

Carver, T., Harris, S.R., Otto, T.D., Berriman, M., Parkhill, J., and McQuillan, J.A. (2013) BamView: visualizing and interpretation of next-generation sequencing read alignments. *Briefings in Bioinformatics* **14**: 203-212.

Cheng, D., and He, J. (2009) Isolation and characterization of "*Dehalococcoides*" sp. strain MB, which dechlorinates tetrachloroethene to *trans*-1,2-dichloroethene. *Appl Environ Microbiol* **75**: 5910-5918.

Cheng, D., Chow, W.L., and He, J. (2010) A *Dehalococcoides*-containing co-culture that dechlorinates tetrachloroethene to *trans*-1,2-dichloroethene. *The ISME Journal* **4**: 88-97.

Darling, A.C.E., Mau, B., and Perna, N.T. (2010) progressiveMauve: multiple genome alignment with gene gain, loss and rearrangement. *PLoS One* **5**: e11147.

Darling, A.C.E., Mau, B., Blattner, F.R., and Perna, N.T. (2004) Mauve: Multiple alignment of conserved genomic sequence with rearrangements. *Genome Res* **14**: 1394-1403.

de Wit, C.A. (2002) An overview of brominated flame retardants in the environment. *Chemosphere* **46**: 583-624.

Delcher, A.L., Harmon, D., Kasif, S., White, O., and Salzberg, S.L. (1999) Improved microbial gene identification with GLIMMER. *Nucleic Acids Res* **27**: 4636-4641.

Ding, C., Chow, W.L., and He, J. (2013) Isolation of *Acetobacterium* sp. strain AG that reductively debrominates octa- and penta- brominated diphenyl ether technical mixtures. *Appl Environ Microbiol* **79**: 1110-1117.

Ding, C., Zhao, S., and He, J. (2014) A *Desulfitobacterium* sp. strain PR reductively dechlorinates both 1,1,1-trichloroethane and chloroform. *Environ Microbiol* **16**: 3387-3397.

Fowles, J.R., Fairbrother, A., Baecher-Steppan, L., and Kerkvliet, N.I. (1994) Immunologic and endocrine effects of the flame-retardant pentabromodiphenyl ether (DE-71) in C57BL/6J mice. *Toxicology* **86**: 49-61.

Goris, J., Konstantinidis, K.T., Klappenbach, J.A., Coenye, T., Vandamme, P., and Tiedje, J.M. (2007) DNA–DNA hybridization values and their relationship to whole-genome sequence similarities. *Int J Syst Evol Microbiol* **57**: 81-91.

He, J., Robrock, K.R., and Alvarez-Cohen, L. (2006) Microbial reductive debromination of polybrominated diphenyl ethers (PBDEs). *Environ Sci Technol* **40**: 4429-4434.

He, J., Holmes, V.F., Lee, P.K.H., and Alvarez-Cohen, L. (2007) Influence of vitamin B_{12} and cocultures on the growth of *Dehalococcoides* isolates in defined medium. *Appl Environ Microbiol* **73**: 2847-2853.

Hites, R.A. (2004) Polybrominated diphenyl ethers in the environment and in people: a meta-analysis of concentrations. *Environ Sci Technol* **38**: 945-956.

Iwamoto, T., and Nasu, M. (2001) Current bioremediation practice and perspective. *J Biosci Bioeng* **92**: 1-8.

Karlin, S. (1998) Global dinucleotide signatures and analysis of genomic heterogeneity. *Curr Opin Microbiol* **1**: 598-610.

Konstantinidis, K.T., and Tiedje, J.M. (2005) Genomic insights that advance the species definition for prokaryotes. *Proceedings of the National Academy of Sciences of the United States of America* **102**: 2567-2572.

Koskiniemi, S., Sun, S., Berg, O.G., and Andersson, D.I. (2012) Selection-driven gene loss in bacteria. *PLoS Genet* **8**: e1002787.

Kube, M., Beck, A., Zinder, S.H., Kuhl, H., Reinhardt, R., and Adrian, L. (2005) Genome sequence of the chlorinated compound respiring bacterium *Dehalococcoides* species strain CBDB1. *Nat Biotechnol* **23**: 1269-1273.

Lagesen, K., Hallin, P., Rødland, E.A., Stærfeldt, H.-H., Rognes, T., and Ussery, D.W. (2007) RNAmmer: consistent and rapid annotation of ribosomal RNA genes. *Nucleic Acids Res* **35**: 3100-3108.

Laslett, D., and Canback, B. (2004) ARAGORN, a program to detect tRNA genes and tmRNA genes in nucleotide sequences. *Nucleic Acids Res* **32**: 11-16.

Lee, L.K., and He, J. (2010) Reductive debromination of polybrominated diphenyl ethers by anaerobic bacteria from soils and sediments. *Appl Environ Microbiol* **76**: 794-802.

Lee, L.K., Ding, C., Yang, K.L., and He, J. (2011) Complete debromination of tetra- and penta- brominated diphenyl ethers by a coculture consisting of *Dehalococcoides* and *Desulfovibrio* species. *Environ Sci Technol* **45**: 8475-8482.

Li, R., Zhu, H., Ruan, J., Qian, W., Fang, X., Shi, Z. *et al.* (2010) De novo assembly of human genomes with massively parallel short read sequencing. *Genome Res* **20**: 265-272.

Lowe, T.M., and Eddy, S.R. (1997) tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. *Nucleic Acids Res* **25**: 955-964.

Marchler-Bauer, A., Zheng, C., Chitsaz, F., Derbyshire, M.K., Geer, L.Y., Geer, R.C. *et al.* (2013) CDD: conserved domains and protein three-dimensional structure. *Nucleic Acids Res* **41**: D348-D352.

McMurdie, P.J., Hug, L.A., Edwards, E.A., Holmes, S., and Spormann, A.M. (2011) Site-specific mobilization of vinyl chloride respiration islands by a mechanism common in *Dehalococcoides*. *BMC Genomics* **12**: 287.

McMurdie, P.J., Behrens, S.F., Muller, J.A., Goke, J., Ritalahti, K.M., Wagner, R. *et al.* (2009) Localized plasticity in the streamlined genomes of vinyl chloride respiring *Dehalococcoides*. *PLoS Genet* **5**: e1000714.

Meier-Kolthoff, J., Auch, A., Klenk, H.-P., and Goker, M. (2013) Genome sequence-based species delimitation with confidence intervals and improved distance functions. *BMC Bioinformatics* **14**: 60.

Moriya, Y., Itoh, M., Okuda, S., Yoshizawa, A.C., and Kanehisa, M. (2007) KAAS: an automatic genome annotation and pathway reconstruction server. *Nucleic Acids Res* **35**: W182-W185.

Morris, R.M., Sowell, S., Barofsky, D., Zinder, S., and Richardson, R. (2006) Transcription and mass-spectroscopic proteomic studies of electron transport oxidoreductases in *Dehalococcoides ethenogenes*. *Environ Microbiol* **8**: 1499-1509.

Mun, C.H., He, J., and Ng, W.J. (2008) Pentachlorophenol dechlorination by an acidogenic sludge. *Water Res* **42**: 3789-3798.

Norén, K., and Meironyté, D. (2000) Certain organochlorine and organobromine contaminants in Swedish human milk in perspective of past 20-30 years. *Chemosphere* **40**: 1111-1123.

Pan, Y., Chen, J., Zhou, H., Farzana, S., and Tam, N.F.Y. (2017) Vertical distribution of dehalogenating bacteria in mangrove sediment and their potential to remove polybrominated diphenyl ether contamination. *Mar Pollut Bull*: DOI: 10.1016/j.marpolbul.2016.1012.1030.

Punta, M., Coggill, P.C., Eberhardt, R.Y., Mistry, J., Tate, J., Boursnell, C. *et al.* (2012) The Pfam protein families database. *Nucleic Acids Res* **40**: D290-D301.

Robrock, K.R., Korytar, P., and Alvarez-Cohen, L. (2008) Pathways for the anaerobic microbial debromination of polybrominated diphenyl ethers. *Environ Sci Technol* **42**: 2845-2852.

Tanabe, S. (2008) Temporal trends of brominated flame retardants in coastal waters of Japan and South China: Retrospective monitoring study using archived samples from *es*-Bank, Ehime University, Japan. *Mar Pollut Bull* **57**: 267-274.

Tettelin, H., Radune, D., Kasif, S., Khouri, H., and Salzberg, S.L. (1999) Optimized multiplex PCR: efficiently closing a whole-genome shotgun sequencing project. *Genomics* **62**: 500-507.

Tokarz, J.A., Ahn, M.Y., Leng, J., Filley, T.R., and Nies, L. (2008) Reductive debromination of polybrominated diphenyl ethers in anaerobic sediment and a biomimetic system. *Environ Sci Technol* **42**: 1157-1164.

UNEP (2009) Depositary notification. Report Number: C.N.524.2009.TREATIES-4 (Depositary Notification).

USEPA (2009) Polybrominated diphenyl ethers (PBDEs) action plan. URL: epa.gov/sites/production/files/2015-09/documents/pbdes_ap_2009_1230_final.pdf

Wagner, A., Cooper, M., Ferdi, S., Seifert, J., and Adrian, L. (2012) Growth of *Dehalococcoides mccartyi* strain CBDB1 by reductive dehalogenation of brominated benzenes to benzene. *Environ Sci Technol* **46**: 8960-8968.

Wang, S., Chng, K.R., Wilm, A., Zhao, S., Yang, K.-L., Nagarajan, N., and He, J. (2014) Genomic characterization of three unique *Dehalococcoides* that respire on persistent polychlorinated biphenyls. *Proceedings of the National Academy of Sciences of the United States of America* **111**: 12103-12108.

Williams, K.P. (2003) Traffic at the tmRNA gene. *J Bacteriol* **185**: 1059-1070.

Woodson, J.D., and Escalante-Semerena, J.C. (2004) CbiZ, an amidohydrolase enzyme required for salvaging the coenzyme B_{12} precursor cobinamide in archaea.

Proceedings of the National Academy of Sciences of the United States of America **101**: 3591-3596.

Yang, C., Kublik, A., Weidauer, C., Seiwert, B., and Adrian, L. (2015) Reductive dehalogenation of oligocyclic phenolic bromoaromatics by *Dehalococcoides mccartyi* strain CBDB1. *Environ Sci Technol* **49**: 8497-8505.

Acknowledgements

This work was jointly supported by the Singapore Agency for Science, Technology and Research (A*STAR) of the Science and Engineering Research Council under Project No. 102 101 0025 and the Singapore National Research Foundation Competitive Research Programme from the Prime Minister's office under Project No.: NRF-CRP5-2009-05. Genome sequencing data of strain GY55 were produced by the US Department of Energy Joint Genome Institute http://www.jgi.doe.gov/ in collaboration with the user community (CSP 2010).

Competing Financial Interests The authors declare no competing financial interests.

Corresponding Author Correspondence and requests for materials should be addressed to Jianzhong He at (jianzhong.he@nus.edu.sg).

The genome sequence of strain GY50 was deposited in GenBank with accession number CP006730.

Figure Legends

Figure 1. DGGE profile suggests culture purity of *D. mccartyi* strain GY50. Lane 1: gDNA of *D. mccartyi* strain GY50. **Lane 2**: gDNA of the *Desulfovibrio* strain isolated from co-culture GY2. **Lane 3**: gDNA of co-culture GY2 treated with 1g/L ampicillin. DGGE running conditions: 120 V for 12 hours at 60 ºC, denaturing gradient: 35%~75%.

Figure 2. Debromination of tetra- and penta-BDE congeners by *D. mccartyi* strains GY50 and GY52. A, Debromination of tetra- and penta-BDE congeners to diphenyl ether by strain GY50. **B**, Debromination of tetra- and penta-BDE congeners to di-BDE 15 by strain GY52. Reported values are averages of triplicates with error bars indicating standard deviations. No debromination products were detected in un-inoculated and autoclaved controls.

Figure 3. Schematic illustration of *ssrA*-GI deletion in *Dehalococcoides* strains GY52 and GY55. Upper panel. The genome of strain GY50 showing the position of the deleted *ssrA*-GI in strains GY52 and GY55. **Lower panel.** Detailed genetic components in the deleted regions. Coordinates shown are absolute base pair positions in the genome of strain GY50. *ssrA* repeat 1 and 2: the first two 20-bp *ssrA* repeats (the third direct repeat is beyond the figure area).

Figure 4. Transcripts of *pbrA1*, *pbrA2*, and *pbrA3* during PBDE debromination by *D. mccartyi* strains GY50 and GY52. Numbers of gene transcripts per cell were obtained by normalizing transcript copies based on luciferase mRNA losses, and then divided by cell numbers at each time point. Data points were averaged from duplicate samples with error bars indicating standard deviations.

Tables

Table 1 Average nucleotide identity (ANI) among representative *Dehalococcoides* genomes

Subgroup	Victoria					Cornell		Pinellas					
Strain	GY50	CG1	CG3	VS	UCH007	195	CG4	BAV1	BTF08	CBDB1	CG5	DCMB5	GT
GY50		99.1	99.1	97.6	98.6	89.5	89.6	86.5	86.5	86.3	86.3	86.5	86.5
CG1	98.9		99.4	97.2	98.2	89.4	89.4	86.4	86.4	86.4	86.3	86.4	86.4
CG3	98.9	99.3		97.1	98.2	89.3	89.4	86.5	86.6	86.4	86.6	86.1	86.5
VS	97.6	97.3	97.4		97.2	89.3	89.4	86.4	86.6	86.2	86.2	86.3	86.6
UCH007	98.5	98.2	98.2	97.1		89.4	89.3	86.4	86.8	86.2	86.4	86.3	86.5
195	89.7	89.7	89.5	89.6	89.7		98.0	86.2	86.4	86.1	86.2	86.0	86.4
CG4	89.5	89.5	89.5	89.5	89.5	97.9		85.7	85.8	85.6	85.7	85.5	85.9
BAV1	86.5	86.4	86.5	86.4	86.4	86.2	85.8		98.8	99.1	98.8	98.8	98.7
BTF08	86.4	86.4	86.6	86.6	86.8	86.3	85.8	98.7		99.0	98.9	98.8	98.9
CBDB1	86.3	86.4	86.4	86.3	86.2	85.9	85.6	99.1	99.1		99.5	99.3	99.4
CG5	86.3	86.3	86.6	86.2	86.5	86.0	85.8	98.8	99.0	99.5		98.9	99.1
DCMB5	86.4	86.3	86.2	86.3	86.2	85.9	85.4	98.7	98.8	99.4	99.0		98.9
GT	86.5	86.4	86.6	86.5	86.6	86.3	85.8	98.5	98.8	99.3	98.9	98.7	

Note: The values were calculated by running the python module pyani (github.com/widdowquinn/pyani) with the ANIb method (Goris *et al.*, 2007). ANIb uses BLASTN (BLAST+ version 2.6.0) to align 1020 bp fragments of the input sequences. BLASTN options are "-xdrop_gap_final 150 -dust no -evalue 1e-15 -max_target_seqs 1". The values were presented in a color scale for clarity. A higher value indicates a shorter distance between two genomes.

Table 2 Top protein hits and all reductive dehalogenase hits detected in strain GY50's proteome by nano-LC-MS-MS

Rank	Score	Name	Size (amino acid)	Coverage (%)	Amino acid hit
1	39.2	BNR/Asp-box repeat-containing protein	1016	27.0	274
2	32.0	reductive dehalogenase PbrA1	504	36.1	182
3	29.5	Ni/Fe hydrogenase, large subunit	526	35.9	189
4	28.0	hypothetical protein GY50_0639	843	23.6	199
5	28.0	hypothetical protein GY50_0640	468	21.8	102
6	26.1	molybdopterin oxidoreductase, formate dehydrogenase subunit alpha	993	16.2	161
7	24.0	DNA-directed RNA polymerase subunit beta	1272	10.6	135
8	24.0	chaperonin GroEL (HSP60)	537	19.0	102
9	21.8	F-type H+-transporting ATPase subunit beta	464	27.4	127
10	19.1	F-type H+-transporting ATPase subunit alpha	503	25.5	128
11	18.1	reductive dehalogenase PbrA2	513	24.8	127
12	17.1	DNA-directed RNA polymerase subunit beta'	1295	11.4	148

13	16.8	membrane-bound proton-translocating pyrophosphatase	708	12.4	88
14	16.0	iron-sulfur cluster-binding protein	600	16.3	98
15	16.0	oligopeptide transport system substrate-binding protein	558	16.5	92
16	16.0	molecular chaperone DnaK family	636	18.5	118
17	12.7	Ni/Fe hydrogenase, iron-sulfur cluster-binding subunit	267	36.0	96
18	12.1	DNA-directed RNA polymerase subunit alpha	330	19.7	65
19	12.0	S-adenosylmethionine synthetase	406	21.4	87
20	10.3	glyceraldehyde-3-phosphate dehydrogenase, type I	334	25.5	85
99	4.0	reductive dehalogenase PbrA3	475	5.3	25

Note: scores >2 correspond to a confidence level of >99%; scores >1.3 correspond to a confidence level of >95%

Table 3 Top protein hits and all reductive dehalogenase hits detected in strain GY52's proteome by nano-LC-MS-MS

Rank	Score	Name	Size (amino acid)	Coverage (%)	Amino acid hit
1	35.4	BNR/Asp-box repeat-containing protein	1016	20.0	203
2	33.6	hypothetical protein GY50_0640	468	47.2	221
3	32.8	chaperonin GroEL (HSP60)	537	27.4	147
4	26.8	molybdopterin oxidoreductase, formate dehydrogenase subunit alpha	993	17.8	177
5	22.0	hypothetical protein GY50_0639	843	22.1	186
6	20.7	Ni/Fe hydrogenase, large subunit	526	21.3	112
7	19.9	membrane-bound proton-translocating pyrophosphatase	708	11.6	82
8	18.0	oligopeptide transport system substrate-binding protein	558	18.6	104
9	16.0	iron-sulfur cluster-binding protein	600	16.2	97
10	14.8	polyribonucleotide nucleotidyltransferase	720	12.8	92
11	14.3	type IV pilus assembly protein PilM	494	16.4	81
12	13.4	F-type H+-transporting ATPase subunit beta	464	13.4	62
13	12.8	F-type H+-transporting ATPase subunit alpha	503	16.3	82
14	12.2	acetyl coenzyme A synthetase (ADP forming), alpha domain-containing protein	891	12.2	109
15	12.0	glyceraldehyde-3-phosphate dehydrogenase, type I	334	21.6	72
16	12.0	reductive dehalogenase PbrA3	475	18.3	87
17	11.0	DNA-directed RNA polymerase subunit beta	1272	7.1	90
18	10.2	LemA family protein	188	27.7	52
19	10.1	DNA-directed RNA polymerase subunit beta'	1295	7.1	92
20	9.9	cell division protease	608	11.8	72
28	7.2	reductive dehalogenase GY50_1443	500	9.8	49
97	2.5	reductive dehalogenase GY50_1397	518	6.8	35
144	2.0	reductive dehalogenase GY50_1381	448	7.1	32

Note: scores >2 correspond to a confidence level of >99%; scores >1.3 correspond to a confidence level of >95%