2018

Divergence in Gene Regulation Contributes to Sympatric Speciation of *Shewanella baltica* Strains

Jie Deng
East China Normal University

Jennifer M. Auchtung
Baylor College of Medicine, jauchtung2@unl.edu

Konstantinos T. Konstantinidis
Georgia Institute of Technology

Alejandro Caro-Quintero
Georgia Institute of Technology

Ingrid Brettar
Helmholtz Centre for Infection Research

See next page for additional authors

AMERICAN SOCIETY FOR MICROBIOLOGY | Applied and Environmental Microbiology®

EVOLUTIONARY AND GENOMIC MICROBIOLOGY

Divergence in Gene Regulation Contributes to Sympatric Speciation of *Shewanella baltica* Strains

Jie Deng,[a,b,c] Jennifer M. Auchtung,[d] Konstantinos T. Konstantinidis,[e,f] Alejandro Caro-Quintero,[f] Ingrid Brettar,[g] Manfred Höfle,[g] James M. Tiedje[c]

[a]Shanghai Key Lab for Urban Ecological Processes and Eco-Restorations, School of Ecological and Environmental Sciences, East China Normal University, Shanghai, China

[b]Shanghai Institute of Eco-Chongming, Shanghai, China

[c]Center for Microbial Ecology, Michigan State University, East Lansing, Michigan, USA

[d]Molecular Virology and Microbiology, Baylor College of Medicine, Houston, Texas, USA

[e]School of Civil and Environmental Engineering, Georgia Institute of Technology, Atlanta, Georgia, USA

[f]School of Biology, Georgia Institute of Technology, Atlanta, Georgia, USA

[g]Helmholtz Centre for Infection Research, Braunschweig, Germany

ABSTRACT Niche partitioning and sequence evolution drive genomic and phenotypic divergence, which ultimately leads to bacterial diversification. This study investigated the genomic composition of two *Shewanella baltica* clades previously identified through multilocus sequencing typing and recovered from the redox transition zone in the central Baltic Sea. Comparative genomic analysis revealed significantly higher interclade than intraclade genomic dissimilarity and that a subset of genes present in clade A were associated with potential adaptation to respiration of sulfur compounds present in the redox transition zone. The transcriptomic divergence between two representative strains of clades A and D, OS185 and OS195, was also characterized and revealed marked regulatory differences. We found that both the transcriptional divergence of shared genes and expression of strain-specific genes led to differences in regulatory patterns between strains that correlate with environmental redox niches. For instance, under anoxic conditions of respiratory nitrate ammonification, OS185—the strain isolated from a nitrate-rich environment—upregulated nearly twice the number of shared genes upregulated by OS195—the strain isolated from an H_2S-containing anoxic environment. Conversely, OS195 showed stronger induction of strain-specific genes, especially those associated with sulfur compound respiration, under thiosulfate-reducing conditions. A positive association between the level of transcriptional divergence and the level of sequence divergence for shared genes was also noted. Our results provide further support for the hypothesis that genomic changes impacting transcriptional regulation play an important role in the diversification of ecologically distinct populations.

IMPORTANCE This study examined potential mechanisms through which co-occurring *Shewanella baltica* strains diversified to form ecologically distinct populations. At the time of isolation, the strains studied composed the major fraction of culturable nitrate-reducing communities in the Baltica Sea. Analysis of genomic content of 13 *S. baltica* strains from two clades representing different ecotypes demonstrated that one clade specifically possesses a number of genes that could favor successful adaptation to respire sulfur compounds in the portion of the water column from which these strains were isolated. In addition, transcriptional profiling of fully sequenced strains representative of these two clades, OS185 and OS195, under oxygen-, nitrate-, and thiosulfate-respiring conditions demonstrated that the strains exhibit relatively similar transcriptional responses during aerobic growth but more-distinct transcriptional responses under nitrate- and thiosulfate-respiring conditions. Results from this study provide insights into

Received 16 September 2017 **Accepted** 30 November 2017

Accepted manuscript posted online 8 December 2017

Citation Deng J, Auchtung JM, Konstantinidis KT, Caro-Quintero A, Brettar I, Höfle M, Tiedje JM. 2018. Divergence in gene regulation contributes to sympatric speciation of *Shewanella baltica* strains. Appl Environ Microbiol 84:e02015-17. https://doi.org/10.1128/AEM.02015-17.

Editor Frank E. Löffler, University of Tennessee and Oak Ridge National Laboratory

Address correspondence to James M. Tiedje, tiedjej@msu.edu.

J.D. and J.M.A. contributed equally to this article.

how genomic and gene regulatory diversification together impacted the redox specialization of the *S. baltica* strains.

KEYWORDS *Shewanella baltica*, comparative transcriptomics, comparative genomics, adaptation

Characterizing the potential mechanisms that bacterial strains have used to diverge into distinct, functionally coherent populations has been greatly facilitated by improved tools for comparative genomics, which can provide a more comprehensive assessment of evolutionary history. Several important applications developed include estimation of evolutionary relatedness using whole-genome sequences (e.g., average nucleotide identity [ANI; 1]), profiling gene content differences among populations favoring distinct environments (2), and assessing the extent of genetic material exchange (i.e., recombination frequencies) to infer dynamics of genome evolution (3, 4). These tools have generated important insights into trajectories and mechanisms regarding bacterial genome evolution (5–7). In the meantime, the relative importance of transcriptional regulation in bacterial diversification and adaptation has also been reported. Insights are emerging on the breadth of intraclade-specific transcriptional divergence that may develop among closely related strains (8–10). Increasing evidence has shown that differences in gene expression patterns may contribute to niche specialization (11–13). Nonetheless, further exploration of this field is required, and combined genomic, transcriptomic, and phenotypic data will provide a more holistic understanding of the mechanisms underlying the bacterial speciation process.

The aim of this study was to investigate the contribution of gene content and gene expression differences to diversification of *Shewanella baltica* populations. The strains used in this study were part of a strain collection obtained during 1986 to 1987 from the Gotland Deep station in the Baltic Sea. The Gotland Deep is the largest basin in the Baltic Sea. The presence of a steep halocline limits vertical mixing and promotes stable stratification of the water column where oxic and anoxic waters meet. The oxic/anoxic transition zone is an energy-rich layer with changing concentrations of oxygen, nitrate, and hydrogen sulfide. Periodic disruptions, such as the inflow of salt water from the North Sea that occurred in 1993, can promote renewal of the deep water and temporary disruption of stratified zones in the water column, allowing bacterial populations previously separated in stratified layers to mix and coexist until the water column restratifies. At the time that strains were collected in 1986, the water column had been stably stratified for approximately 9 years (14).

All strains were assigned to the species named *S. baltica* based on conventional genetic and phenotypic similarity standards. Nonetheless, further investigations based on finer-scale genotyping and metabolic profiling revealed substantial diversity at both the genomic and phenotypic levels (4). Specialization of the *S. baltica* populations corresponded to nutrient availability, redox conditions, particle association, and temporal distribution. Among the phylogenetic clades delineated through multilocus sequence typing (MLST), two clades, MLST-A and MLST-D, were proposed to represent two putative *S. baltica* ecotypes. MLST-A (corresponding to previously described clade RAPD-A [15]) was the largest *S. baltica* clade obtained during 1986 and 1987. Clade A strains were enriched in the oxic water-anoxic water transition zone where H_2S was present and were favorably obtained from anaerobic medium with supplementary thiosulfate. This population is also metabolically versatile with regard to carbon source utilization and respiratory capabilities, and its members were recovered at greater frequencies as well as depths in the latter year. Compared to MLST-A, strains of MLST-D (corresponding to the larger RAPD-C clade) share similar patterns with regard to particle association and favored culturing conditions and yet were most frequently obtained from the narrow zone where nitrate concentration and denitrification activity peaked (14, 15). Therefore, this genotype is considered to be potentially important in N cycling in the Baltic Sea. In a previous report, MLST-D and MLST-A strains were shown to have extensively recombined in the recent past and to have exchanged over 20% of

their genomes (16). Nonetheless, it is unclear what forces contributed to maintenance of these strains as two genetically and phenotypically separated populations regardless of the high frequency of recombination and their commonality in physiological characteristics or whether such forces were ecologically relevant.

Here, we used comparative genomic hybridization (CGH) to fingerprint the genomic contents of 10 MLST-A strains and three MLST-D strains and used comparative transcriptional hybridization (CTH) to assess the differences in gene expression of two representative strains, OS195 and OS185, from MLST-A and MLST-D, respectively. The strains were grown in medium with minimal nutrient input under oxic conditions or were grown anoxically with supplementation of nitrate or thiosulfate as the sole electron acceptor, mimicking simplified *in situ* conditions in the Baltic Sea. The results yielded insights into how genomic and transcriptomic divergence contributes to specialization of *S. baltica* populations and into whether there are trends associated with ecological adaptation by these populations.

RESULTS

Comparative genomics revealed intraclade genomic homogeneity and ecologically relevant clade-specific genes. Hierarchical clustering grouped the targeted strains into two clades, clade A and clade D (see Fig. S1 in the supplemental material), consistent with the grouping by the MLST approach (4). The intraclade similarities for strains of clades A and D were 99.06% and 99.40%, respectively, contrasting with the average interclade similarity of 91.75% (Sorenson similarity based on gene presence/absence profiles). Over 300 genes were unique to each clade; the majority encode hypothetical proteins or were poorly characterized (Fig. S1). (Clade D unique genes were defined as those genes present in all three strains of clade D and not found among the clade A unique genes; clade A unique genes were defined as those genes present in ≥9 of 10 clade A genomes and not found among the clade D unique genes.) Nonetheless, a number of clade A-specific genes that are predicted to play a role in respiration of sulfur compounds and that could be ecologically important for redox adaptation were identified. Examples include the type I anaerobic dimethyl sulfoxide (DMSO) reductase (*S. baltica* 195_2225-2230 [Sbal195_2225-2230]) and the sulfite dehydrogenase (Sbal195_0006-0009) gene operons, both of which were unique to and present in all clade A genomes. Transcriptomic analysis (see below) confirmed expression of these genes under anoxic conditions. In summary, comparative genomic analyses revealed intraclade homogeneity and that gene content differences were potentially associated with redox specialization of clade A strains.

In order to assess whether the genes unique to clade A or D resulted from lateral gene acquisition or gene loss, we mapped their presence/absence among a greater number of *S. baltica* strains (4, 16). These include strains of clade B, which shares a recent common ancestor with clade A; strains of clade C, which shares a recent common ancestor with clade D; and 18 additional *S. baltica* strains. Figure S2 depicts the phylogenetic relationship among these strains. (Table S1 provides gene-level resolution of gene loss/acquisition analysis.) Briefly, we consider genes/gene cassettes to be laterally acquired when they appear in only one of these clades, whereas absence of genes from a branch more likely associates with the process of gene loss. Here, most clade A- or D-specific genes appeared to result from gene loss. Examples include loss of sulfite dehydrogenases (Sbal195_0006-0009) in clade D and loss of a type III secretion pathway-associated gene(s) (Sbal195_2246-2281) in clades C and D. Nevertheless, we acknowledge that this method was based on simplified scenarios of the dynamics of gene flow. For instance, the DMSO reductase genes were categorized as representing gene loss by the ancestor of clades C and D based on their presence not only in clade A strains but also in four additional strains. However, the gene cassette was also adjacent to a phage integrase (Sbal195_2224), implying a possibility of being laterally acquired and spread through horizontal gene transfer. Accordingly, the frequency of gene loss may be overestimated through this analysis.

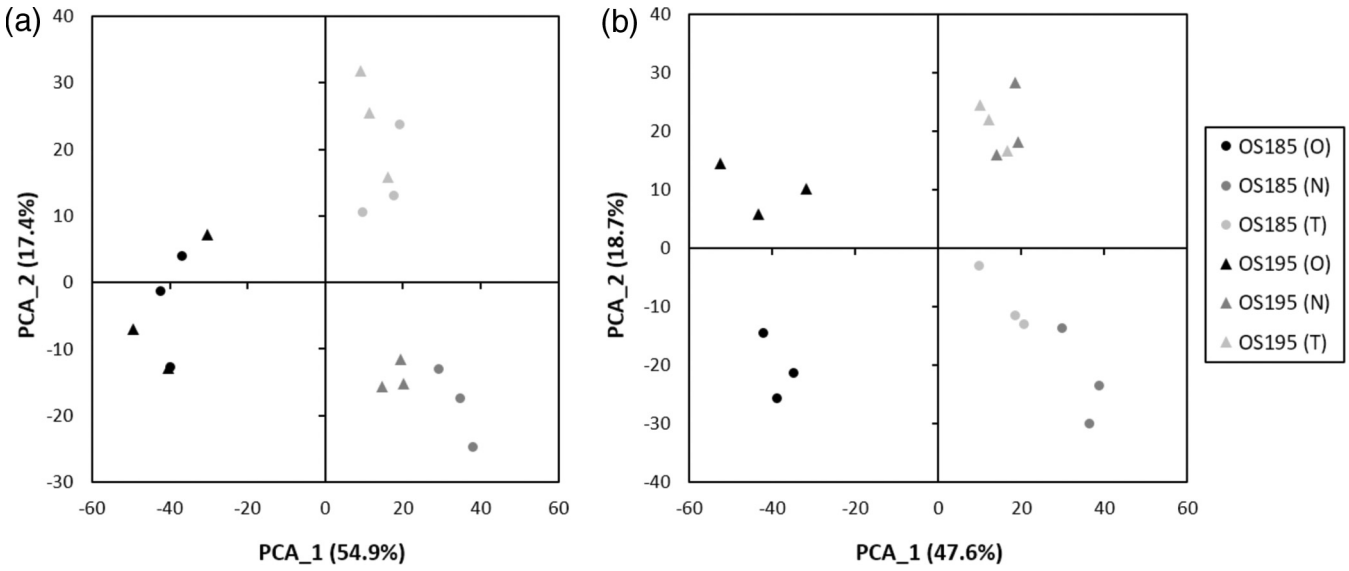

FIG 1 Principal-component analysis (PCA) of genes differentially expressed by OS185 and OS195 grown under oxygen (O), thiosulfate (T) and nitrate (N)-respiring conditions. Analysis was done on (a) genes shared by OS185 and OS195 and (b) all differentially expressed genes.

On the other hand, gene acquisition was predicted for a number of genes and at various stages. For instance, polysaccharide biosynthesis genes (Sbal195_3017-3035) appeared to be acquired by the ancestor of clades A and B but were subsequently lost by OS187 in clade B and preserved among clade A strains. Cytochromes corresponding to Sbal195_0256-0258 might have been gained by the common ancestor of clades A, B, C, and D and yet were lost in clade D. For the sulfite reductase *sirABIGCDJKLM* genes (Shew185_3866-3877 or Sbal195_3991-4002), a part (Shew185_3866-3870 and Sbal195_3991-3994) was shared among clades A, B, and D, with over 94% sequence identity between OS185 and OS195. However, the other part of the operon, i.e., Shew185_3870-3877 in OS185, had lower similarity with its counterparts in OS195 (lower than 70% for some genes) and yet shared over 93% sequence identity with CEQ32_13660-13695 in *Shewanella* sp. strain FDAARGOS_354. Our results imply a possible scenario regarding the evolutionary trajectory of this gene operon. The genes might have been initially acquired by the common ancestor of clades A, B, C, and D and yet were subsequently lost in clade C. Then, part of the gene cassette in OS185 underwent gene exchange, possibly through homologous recombination with their counterparts in *Shewanella* sp. strain FDAARGOS_354, therefore resulting in a "mosaic" pattern in the origin of genes within the same gene cassette. Taken together, our analyses provide additional information on possible evolutionary trajectories of a subset of the clade A- and D-specific genes. Nonetheless, more-accurate prediction of gene gain and loss will require whole-genome sequencing of a greater number of *S. baltica* strains.

Global transcriptional patterns of OS185 and OS195. Transcriptomic profiling revealed a total of 1,209 genes in OS185 and OS195 that were differentially expressed under oxic or anoxic nitrate- or thiosulfate-respiring growth conditions. These differentially expressed genes included those showing similar responses to different electron acceptors between strains as well as those with divergent responses. Among the approximately 300 genes that were unique to each strain, 49 (OS185) or 97 (OS195) genes were differentially expressed under one of the conditions tested. We used principal-component analysis (PCA) to visualize the global transcriptional patterns in response to the availability of different electron acceptors based on genes shared by both the OS185 and OS195 genomes (Fig. 1a) or on all differentially expressed genes (including clade-specific genes) (Fig. 1b). In both cases, the differences between aerobic and anaerobic respiration dominated and drove the ordination along PCA_1, which alone accounted for 43.9% and 51.6% of the total variance, respectively. Considering

TABLE 1 Dissimilarity analyses[a] of OS185 versus OS195 on shared genes

	Euclidean distance	
Condition	Adonis *P*	Adonis R^{2b}
N	0.097	0.645
O	0.206	0.294
T	0.090	0.468

[a]The dissimilarity test was performed with gene expression profiles of strains growing in the presence of the indicated electron acceptors: N, nitrate; O, oxygen; T, thiosulfate.
[b]R^2, partition of the sum of squares.

only genes shared by OS185 and OS195, the transcriptional responses to nitrate or thiosulfate respiratory conditions were more similar between strains, with OS185 and OS195 samples clustering together and being separated by anoxic electron acceptors along PCA_2 (Fig. 1a). When clade-specific genes were also included in the PCA, the separation by anoxic electron acceptors was lost and the samples clustered by strain along PCA_2 (Fig. 1b). Therefore, despite much smaller numbers of strain-specific genes that were differentially regulated when grown in the presence of different terminal electron acceptors, expression of these genes significantly contributed to overall transcriptomic differences between strains OS185 and OS195, especially under anoxic conditions. In addition, the differential results with respect to expression levels of strains of shared genes were more significant under nitrate and thiosulfate-respiratory conditions than under oxygen-respiratory conditions (Table 1). Therefore, our results suggest that growth conditions not only determine the overall expression patterns of shared genes but also impact the extent of divergence in gene expression levels between strains.

The extent of transcriptional divergence decreases with sequence identity. To identify potential mechanisms underlying differences in the transcriptional responses of genes shared between the two strains (OS185 and OS195) under different respiratory conditions, we determined the fold change differences for each shared gene under each respiratory condition and plotted the results in relation to the percentage of nucleotide identity of that gene. This analysis revealed that fold change differences in gene expression decreased with increasing sequence identity between OS185 and OS195 ($P < 0.0001$ in all cases) (Fig. 2), suggesting a positive correlation between divergence in transcriptional regulation and divergence in genomic sequences. This pattern was more pronounced under nitrate-reducing conditions than under thiosulfate-reducing conditions.

A previous analysis (16) demonstrated that large sections of the genomes in OS185 and OS195 had undergone recent recombination and typically show greater than

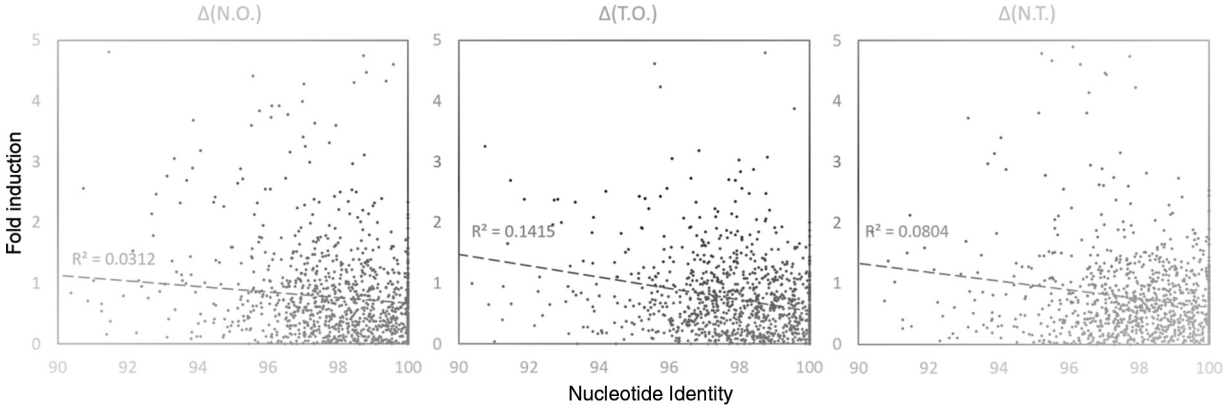

FIG 2 Differences in transcriptional responses of genes shared between OS185 and OS195 under different respiratory conditions [Δ(N.O.), nitrate versus oxygen; Δ(T.O.), thiosulfate versus oxygen; Δ(N.T.), nitrate versus thiosulfate] decrease with sequence identity. Differences between OS185 and OS195 in the magnitude of gene expression changes (Log_2) under the indicated respiratory conditions were calculated for each gene (*y* axis) and plotted relative to the percentage of nucleotide identity of that gene (*x* axis).

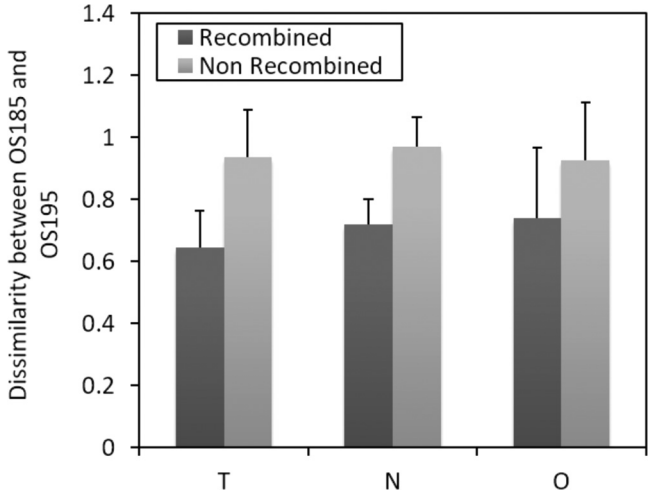

FIG 3 Recombined genes share lower dissimilarity in expression patterns between OS185 and OS195. Dissimilarities reported represent Euclidean distances calculated from gene expression profiles.

99.7% nucleotide identity. We reasoned that these recently recombined genes, which may be undergoing convergent evolution, would be more likely to share regulatory regions and have similar transcriptional profiles. To test this, we separated differentially expressed genes into populations of recently recombined genes and nonrecombined genes and determined overall dissimilarity scores between the two strains grown in the presence of different terminal electron acceptors. Our results show that recombined genes shared more similar overall expression patterns than nonrecombined genes (Fig. 3), which further confirmed that genes that share greater sequence identity between strains have lower divergence in gene expression patterns.

Characteristic gene expression under oxygen, nitrate, and thiosulfate respiratory conditions. In the presence of oxygen as the terminal electron acceptor, there was increased expression of genes related to fast cell growth in both strains (Table S3 and Fig. S3), e.g., upregulation of genes involved in membrane synthesis, nucleotide metabolism, and translation processes. Aerobic growth also induced the nitrogen starvation response, evidenced by upregulation of carbon cycling and amino acid metabolism genes coupled with those encoding P-II and NtrC/NtrB regulatory proteins. Upregulation of specific energy metabolism-related genes, including those encoding a number of *c*-type cytochromes (Fig. S4), and of many others that have been previously confirmed to be involved in aerobic respiration by *Shewanella oneidensis* MR-1 (17–19) was also observed.

Under nitrate- and thiosulfate-respiring conditions, expression of genes required for anaerobic respiration increased (Table S4 and Fig. S5). Growth in the presence of nitrate caused highly induced expression of both nitrate reductase subsystems in both strains, i.e., *napBAD* of subsystem I (Shew185_0823-0825) and *napBC* of subsystem II (*napEDABC*; Shew185_1934-1938), as well as the *nrfA* gene encoding nitrite reductase (Shew185_3711) (Fig. S6). However, the other nitrite reductase subsystem, *nrfDCBA* (Shew185_4151-4154), was induced only under nitrate-reducing conditions in OS195 but was induced under thiosulfate-reducing conditions in both strains. In particular, this operon was highly (6.65-fold) upregulated in OS195 under conditions of thiosulfate growth, suggesting that it may have functions different from those indicated by bioinformatic annotation. Indeed, one of the genes in this operon (Shew185_4151) has been alternatively assigned functions as a polysulfide reductase (Pfam) or a nitrate reductase (KEGG Orthology), indicating the need for further resolution based upon functional characterization.

Growth in the presence of thiosulfate increased the expression in both strains of genes involved in inorganic ion transport and metabolism (Fig. S3; category T). Characteristic induction included a number of those genes related to sulfur respi-

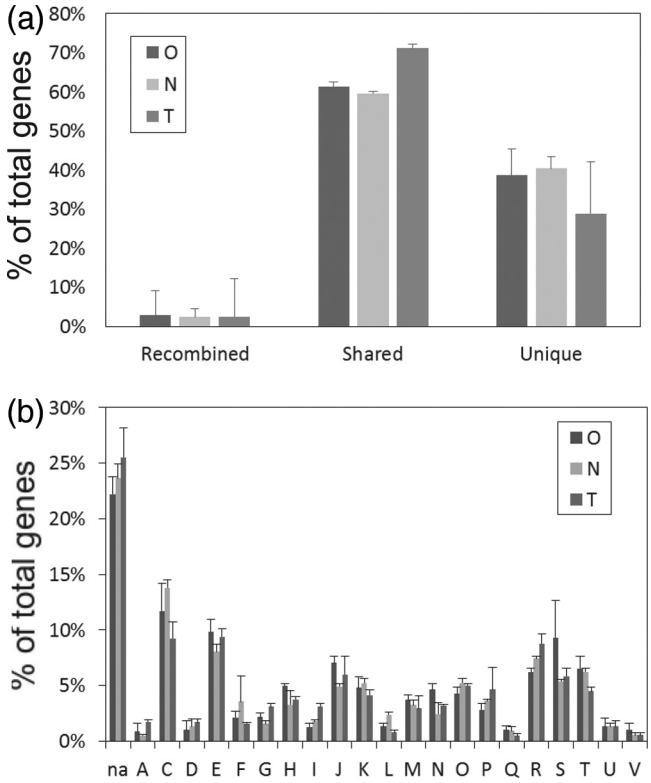

FIG 4 (a) Percentages of total genes from different gene pools (genes recently recombined between OS185 and OS195, genes shared between OS185 and OS195, or genes unique to OS185 or OS195) contributing to overall transcriptomic differences. (b) Clusters of orthologous genes (COG) category distribution of genes shared between OS185 and OS195 that exhibit dissimilar responses to oxygen (O), nitrate (N), or thiosulfate (T) respiratory conditions between strains. See Table S2 for code description for COG categories.

ration (Fig. S5), e.g., the genes encoding sulfur reductase (*phsABC* Shew185_0531-0533), sulfite/thiosulfate reductase (*sirABIGCDJKLM* Sbal195_3991-4002 in OS195, Shew185_3866-3877 in OS185) (20), and a sulfate transporter-associated periplasmic monoheme cytochrome *c* (Shew185_0460) (21). Most of the genes mentioned above were also induced under nitrate respiratory conditions, suggesting that similar transcriptional regulators were activated during anaerobic respiration of nitrate and thiosulfate.

Regulatory differences in shared genes. Among the shared genes that were differentially expressed under at least one set of conditions (1,063 genes), fewer than 34% had consistent expression patterns under all conditions (i.e., were induced under the same condition[s]) in both strains and 39% and 23% (309 and 241 genes) were differentially expressed exclusively in OS185 and OS195, respectively. Differential expression of shared genes contributed to over 60% of the variation between OS195 and OS185 (Fig. 4a). Under all three growth conditions, the greatest fractions of regulatory differences in shared genes were attributed to genes with unknown function, followed by those associated with energy production and conservation, amino acid metabolism, and membrane synthesis. Genes involved in cell cycle control, nucleotide metabolism, gene replication, and translation showed more similarity in expression patterns between strains (Fig. 4b). Genes associated with coenzyme transport, metabolism, and cell motility contributed relatively greater fractions of transcriptional differences between OS185 and OS195 under oxic conditions. Under nitrate-reducing conditions, genes related to energy production and conversion, as well as those involved in replication, recombination, and repair, showed relatively higher contributions to regulatory differences. Thiosulfate-reducing conditions led to higher regulatory divergence of genes related to carbohydrate and lipid transport and metabolism.

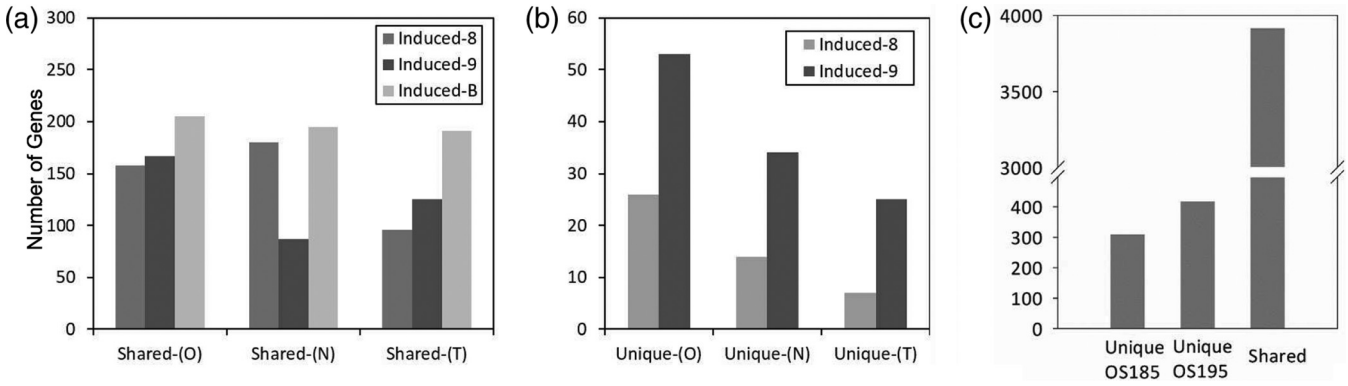

FIG 5 (a and b) The numbers of (a) upregulated and shared and (b) unique genes corresponding to strain OS185 (Induced-8) or strain OS195 (Induced-9) or both strains (Induced-B) under oxygen (O)-, nitrate (N)-, or thiosulfate (T)-respiring conditions. (c) The total numbers of unique and shared genes in the genomes of OS185 and OS195.

In particular, our results revealed that the number of shared genes specifically upregulated in OS185 markedly outnumbered the number specifically upregulated in OS195 under nitrate-reducing conditions (Fig. 5). Specifically, 28 of these genes encode transcriptional regulators, signal transduction components, and membrane sensors (COG K and T categories), among which eight genes were upregulated by over 3-fold. The predicted transcriptional regulators include those belonging to the TtrR, TetR, LysR, GntR, HlyU, and HexR families, as well as other transmembrane and winged helix transcriptional regulators. In contrast, only seven genes associated with the categories listed above were found among the shared genes specifically upregulated by OS195, none of which were highly induced. These observations potentially imply distinct regulatory mechanisms exhibited by OS185 under nitrate-reducing conditions.

The pattern indicating that OS185 upregulated more shared genes in the presence of nitrate as an electron acceptor also applies to highly induced genes (Table S5 and Fig. S6 and S7), especially with regard to redox-related genes. For instance, the induction of nitrate reductase subsystem I (*napBAD*; Shew185_0823-0825) relative to its expression level under aerobic conditions was 3.5-fold higher in OS185 than in OS195. Similarly, the induction of the *mtrABC* operon was 3-fold higher in OS185 than in OS195, that of *frdBAC* (Shew185_0614-0616) was 2.4-fold higher, and those of nitrite reductase *nrfA* (Shew185_3711), formate dehydrogenase (Shew185_4100-4112), and the NiFe hydrogenase (Shew185_1904-1915) were 1.3-fold higher. As a result, the level of induction of shared genes was significantly higher in OS185 under nitrate-reducing conditions (Fig. 6). Our results imply that these unique regulatory patterns of OS185 may potentially be related to adaptation of clade D strains under nitrate-rich, oxygen-limiting conditions.

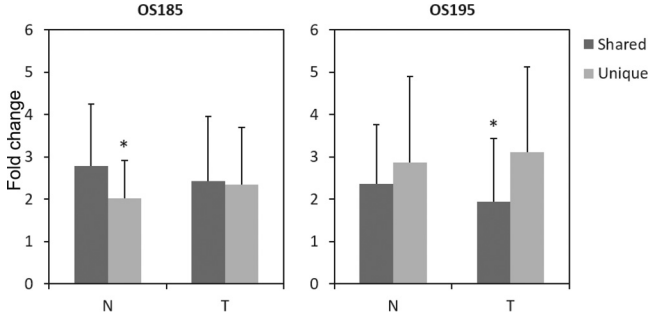

FIG 6 Mean levels of induction (fold change) in shared and unique genes. Asterisks (*) indicate significant differences in the levels of induction.

Regulatory patterns of strain-specific genes. Under all three growth conditions, greater numbers of strain-specific genes were upregulated in OS195 than in OS185 (Fig. 5), which also applies to highly induced genes. For both strains, most of the differentially expressed unique genes were of unknown functions. In OS185, the majority were upregulated under oxygen and nitrate respiratory conditions, and many were distributed in a few putative genomic islands. These include a number of polysaccharide biosynthesis genes (Shew185_2888-2903) and prophage components (Shew185_2093-2095) that were aerobically induced and a partially anaerobically upregulated genomic region (Shew185_2534-2551). We did not identify OS185-specific genes whose induction could be clearly associated with clade D strains thriving under nitrate-respiring conditions.

With regard to differentially expressed OS195-specific genes, nearly half of aerobically induced OS195-specific genes were involved in phage activity or encoded hypothetical proteins, with another 14% associated with membrane and polysaccharide biosynthesis, whereas those induced under anoxic conditions involved a number of oxidoreductases likely important for anoxic adaptation. Specifically, 14 OS195-specific genes were upregulated under both nitrate and thiosulfate respiratory conditions (Fig. S5), including the type I anaerobic DMSO reductase (Sbal195_2225-2230) and part of the sulfite/thiosulfate reductase operon (*sirABIGCDJKLM*; Sbal195_3991-4002 in OS195). (This operon, though homologous to Shew185_3866-3877 in OS185, had sufficient sequence divergence to be represented by different probes.) The sulfite dehydrogenase (Sbal195_0006-0009) operon, also specific to the OS195 genome, was upregulated under thiosulfate-respiring conditions by 4.8-fold relative to its expression under oxygen-respiring conditions. As a result, the average fold induction of OS195-specific genes was significantly higher than that of the shared genes under thiosulfate-reducing conditions (Fig. 6). Hence, expression of strain-specific genes, especially those associated with metabolism of sulfur compounds, may contribute to adaptation of OS195 in anoxic water, especially where sulfur compounds are present as potential alternative electron acceptors.

DISCUSSION

By comparing transcriptional responses to different environmentally relevant electron acceptors between two fully sequenced members from closely related clades of *S. baltica* strains, we demonstrated that both transcriptional divergence of shared gene content and expression of clade-specific genes may contribute to niche adaptation. These results highlight the importance of transcriptional characterization for studying bacterial diversification into ecologically distinct populations. Compared to the number of studies that have investigated diversification in genome content among closely related microorganisms, far fewer studies have assessed ecological implications of transcriptomic diversification. Parkinson et al. revealed that among coral-associated *Symbiodinium* strains, species-specific expression differences were enriched in photosynthesis-related genes, whereas strain-specific expression differences may contribute to host-symbiont genotype parings (12). Carreto et al. reported that, compared to environmental, clinical, and laboratorial *Saccharomyces cerevisiae* strains, the winemaking strains had transcriptional patterns associated with better adaptation, especially during fermentation under nitrogen-limiting and stress-imposing conditions (13). Through investigating gene expression differences between two *Burkholderia cenocepacia* strains, one isolated as a soil dweller from an agricultural field and the other as a pathogen from a cystic fibrosis (CF) patient, under conditions mimicking those two environments, Yoder-Himes et al. showed that the soil strain exhibited a stronger global transcriptional response to its environment, which may provide competitive advantages under nutrient-fluctuating soil conditions (22). Here, our report presents an additional example of transcriptional divergence among closely related strains, i.e., *S. baltica* OS185 and OS195, and provides evidence indicating that differences in gene expression were associated with the presence of terminal electron acceptors rele-

vant to their *in situ* conditions. Our results indicate that transcriptional diversification, in addition to gene gain/loss, serves as an important mechanism underlying bacterial adaptation and diversification.

Expression divergence of shared genes is related to sequence dissimilarity. Differential expression of shared genes played a more significant role in overall transcriptional diversification between OS185 and OS195 than differences in expression levels of unique genes. This observation is consistent with a few previous reports (9, 23) and suggests that patterns of gene regulation may evolve faster than genome content. However, there has been limited evidence about the degree of transcriptional divergence in relation to evolutionary distances. Vital et al. (9) reported that among ecologically distinct *Escherichia coli* strains, highly differentially expressed genes shared significantly lower sequence identity. Here, the similarity in genome content between OS185 and OS195, with an ANI of 96.6%, is higher than the similarity in those genomes investigated by Vital et al. (mean pairwise ANI of 93.1%) (9), suggesting that such patterns hold across even shorter evolutionary distances. Greater similarity in expression profiles of recombined genes further confirmed the trend. However, there was also evidence of rapid transcriptional divergence after recombination, with several recently recombined genes exhibiting nearly or over 2-fold differences in induction between OS185 and OS195, including genes encoding a methyl-accepting chemotaxis sensory transducer protein (Shew185_3397), a redox-active disulfide protein (Shew185_3831), and a few proteins with unknown functions. A possible scenario is that growth conditions may differentially impact gene expression patterns of *S. baltica* strains by imposing different levels of selective pressure, which further implies that the extent of niche partitioning may impact the level of transcriptional divergence.

Niche overlap between OS185 and OS195. Although clades A and D were categorized as representing two individual ecotypes, recombination analysis revealed exchange of a significant part of their genomes. A former habitat association analysis also suggested that strains in the two clades shared similar patterns regarding isolation medium and particle association (15). These observations may suggest that strains in clades A and D occupy partially "overlapping" ecological niches. We show that OS185 and OS195 also share overlapping regulatory patterns of shared genes, with gene expression profiles influenced more by respiratory electron acceptor than by strain boundary (Fig. 1a). In particular, expression patterns of shared genes showed greater similarity under oxic growth conditions than under anoxic conditions (Fig. 1 and Table 1). These more similar patterns for aerobic respiration and adaptation to specific anoxic niches may reflect the strains' adaptation to the ecological cycling of their environment. During periods of stable stratification of the water column, strains diverge to adapt to specific electron acceptor niches within the water column. During periods of water column renewal, strains coexist in the same, oxic environment until the water column returns to its previous stratified state. However, how niche overlap among closely related bacteria translates to overlap in gene expression patterns or genomic contents is rarely known. This report provides some insights into this topic, and yet more controlled experiments are required to quantify the extent of the associations. For instance, investigating overlap as well as divergence in gene expression patterns among experimentally evolved strains with known niche associations may lead to a more complete understanding.

Specialization of OS185 and OS195 through differential regulation of shared genes and niche-adaptive strain-specific genes. Despite the comparable fractions of shared and strain-specific genes that were differentially expressed in OS185 and OS195 (23.8% and 23.9% for OS185; 22.1% and 22.9% for OS195), the patterns of overall induction suggest specialization of the two strains occurring likely through two different strategies: specialization through acquisition or preservation of niche-adaptive genes (OS195) and specialization through changes in transcriptional regulation (OS185). For OS195, strain-specific genes showed significantly higher levels of induction than shared genes under thiosulfate-reducing conditions. In particular, these included the gene encoding DMSO reductase and a number of other genes associated with

respiration of sulfur compounds which were upregulated by OS195 under thiosulfate- or nitrate-reducing conditions. These observations support the hypothesis that processes of gene loss/acquisition contribute to the specialization of OS195 as well as of other clade A strains in the anoxic water zone of the Baltic Sea.

In contrast, we did not find evidence of OS185-specific genes directly related to nitrate respiration and yet instead observed marked regulatory differences among shared genes. OS185 showed both higher levels of induction and a greater number of shared genes upregulated than OS195 under nitrate-respiring conditions, including a large number of putative transcriptional regulators that exhibited increased expression under nitrate-respiring conditions. This indicates that the transcriptional response of OS185 under nitrate-respiring conditions is more comprehensive than simply increasing expression of genes associated with nitrate respiration. On the basis of these results, we propose that adaptation of OS185 occurred primarily through mutations that altered gene expression rather than through acquisition of niche-adaptive genes through horizontal gene transfer.

The ultimate root of the transcriptional differences that occur during evolution is the change of sequences in genes or other DNA regions related to transcriptional regulation. In an extreme case, during a long-term evolution experiment using *E. coli*, mutations in the sequence of one regulator gene resulted in changes of gene expression at the global scale (23, 24). Although it is not known whether differential expression of transcriptional regulators was directly responsible for the performance of OS185 under anoxic nitrate-respiring conditions, our report provides a reference for more-informed selection of genes associated with *S. baltica* adaptation for further experimental investigation.

Implications for transcriptional divergence with ecotype distinction. The level of regulatory differences that translates to a new ecotype has rarely been assessed. We found that OS185 and OS195 share 65% overall regulatory similarity and 75% to 77% regulatory similarity under any specific respiratory condition (based on Gower dissimilarity index data). This provides a rough estimate of the divergence in the levels of gene expression that may develop between strains sharing comparable levels of genetic relatedness. We acknowledge that the level of gene expression divergence may vary with the growth conditions tested and with the investigation platform. Analysis of expression patterns in a greater number of strains from each clade will no doubt improve understanding of the extent of regulatory consistency within a specific genotype. Further, studying gene expression in strains spanning an evolutionary gradient will allow a more confident assessment of the relatedness between gene regulation and evolutionary distance and may provide more understanding of how genetic similarity impacts similarities in gene expression. This report offers some insights into this issue and indicates future directions for investigating the gene expression divergence that takes place as speciation occurs, as well as for describing species from the angles of comparative transcriptomics and evolution of bacterial regulatory networks.

MATERIALS AND METHODS

Array design. Microarrays were made by MYcroarrays, LLC (Ann Arbor, MI, USA), using a proprietary light-directed oligonucleotide synthesis method to synthesize 44-to-48-nucleotide-long probes on glass slides. Probes were designed from the genomic sequences of four *S. baltica* strains initially sequenced. For the purpose of probe design, those genes that shared 90% sequence identity over 90% of their length between sequenced genomes were considered homologous and were treated as a single gene, even when represented by more than one sequence. As homologous gene sequences diverged, probes were designed using the following sequence preference: OS185 → OS195 → OS223 → OS155. This strategy resulted in generation of 30,000 probes, with up to 7 probes designed for each gene and up to 2 probes for each unique intergenic region. An additional 720 negative-control probes were included for quality control.

Strain description and DNA array analysis. A total of 13 *S. baltica* strains which had been previously described in reference 4 were used in this study. DNA extraction and comparative genomic hybridization protocols are described in reference 16. In short, cells were grown aerobically in Trypticase soy broth at 22°C. Genomic DNA was extracted according to reference 25, followed by sonication to produce DNA fragments less than 3 kbp in size. DNA samples were labeled with the fluorescent Cy5 dye. Hybridization reactions were conducted at 50°C for 18 h. Arrays were scanned using an Axon GenePix

4000B scanner (one-channel hybridization). Values representing the mean intensity of signals from the negative-control probes were subtracted from those representing the signals of all spots. For each gene, the median value of normalized signals from all probes corresponding to the gene was taken to represent the gene signal. The presence or absence of genes was determined based on an arbitrary cutoff value, optimization of which yielded minimum false-discovery rates (FDR) based on hybridization profiles from four reference genomes (OS185, OS195, OS223, and OS155). Numerically, the optimized cutoff was 850 signal units, at which value the FDR was 2.2%. The full list of gene presence/absence data for the investigated *S. baltica* strains is included in the supplemental material.

Expression array analysis. Cell growth conditions and RNA extraction and array hybridization protocols are described in reference 16. Briefly, three independent cultures of strains OS185 and OS195 were grown in anaerobic HEPES medium (26) with 10 mM sodium fumarate at 22°C for 18 to 19 h before cells were pelleted and transferred to HEPES medium lacking electron acceptors (aerobic) or containing 5 mM sodium nitrate or 10 mM sodium thiosulfate. All cultures were incubated at 22°C. The aerobic cultures were aerated by shaking on an orbital shaker at 150 rpm. RNA was extracted from cell pellets using a Qiagen RNeasy kit, followed by reverse transcription performed with random primers at 25°C for 10 min, 42°C for 70 min, and 70°C for 15 min (Invitrogen). Labeled cDNAs were purified with QiaQuick MinElute PCR purification columns and eluted in 12 μl of RNase-free water (Qiagen). In the meantime, a reference cDNA pool was constructed using RNA from cells of all four of strains OS155, OS185, OS195, and OS223 grown in HEPES medium containing 10 mM sodium fumarate, 5 mM sodium nitrate, or 10 mM sodium thiosulfate under anoxic conditions or 10 mM sodium chloride under oxic conditions.

For each hybridization experiment, 10 μl of labeled cDNA from strain OS185 or OS195 was mixed with an equal volume of labeled reference cDNA. The same oligoarray as that used as described above for DNA-DNA studies was used for transcriptomic analysis. For each array, Cy3 and Cy5 signals were normalized to the arithmetic mean of ratios using GenePix software. Features having fewer than 50% of pixels with signal more than 2 standard deviations above background in both channels were excluded in downstream analysis. Genes showing significantly different expression levels between pairwise comparisons among the three growth conditions were identified using significance analysis of microarrays (27). Experiments were repeated in triplicate, with each point plotted in the principal-component analysis (PCA). The full list of differentially regulated genes for investigated *S. baltica* strains is included in the supplemental material.

Accession number(s). Original microarray data are saved in the GEO repository under accession no. GSE100228.

SUPPLEMENTAL MATERIAL

Supplemental material for this article may be found at https://doi.org/10.1128/AEM.02015-17.

SUPPLEMENTAL FILE 1, PDF file, 0.6 MB.
SUPPLEMENTAL FILE 2, XLSX file, 0.3 MB.
SUPPLEMENTAL FILE 3, XLSX file, 0.3 MB.
SUPPLEMENTAL FILE 4, XLSX file, 0.2 MB.

ACKNOWLEDGMENTS

We thank the Joint Genome Institute for sequencing and annotating all *S. baltica* genomes and Erdogan Gulari and Jean-Marie Rouillard for assistance in microarray design and production of the arrays. We also thank Margaret Romine and Gretta Seres for helpful discussion and sharing of their data and Amanda Geaslin for assisting with laboratory work.

This work was partially supported by the U.S. Department of Energy Office of Science under grant DE-FG02-07ER64389.

REFERENCES

1. Konstantinidis KT, Tiedje JM. 2005. Genomic insights that advance the species definition for prokaryotes. Proc Natl Acad Sci U S A 102:2567–2572. https://doi.org/10.1073/pnas.0409727102.
2. Dobrindt U, Agerer F, Michaelis K, Janka A, Buchrieser C, Samuelson M, Svanborg C, Gottschalk G, Karch H, Hacker J. 2003. Analysis of genome plasticity in pathogenic and commensal *Escherichia coli* isolates by use of DNA arrays. J Bacteriol 185:1831–1840. https://doi.org/10.1128/JB.185.6.1831-1840.2003.
3. Luo C, Walk ST, Gordon DM, Feldgarden M, Tiedje JM, Konstantinidis KT. 2011. Genome sequencing of environmental *Escherichia coli* expands understanding of the ecology and speciation of the model bacterial species. Proc Natl Acad Sci U S A 108:7200–7205. https://doi.org/10.1073/pnas.1015622108.
4. Deng J, Brettar I, Luo C, Auchtung J, Konstantinidis KT, Rodrigues JLM, Höfle M, Tiedje JM. 2014. Stability, genotypic and phenotypic diversity of *Shewanella baltica* in the redox transition zone of the Baltic Sea. Environ Microbiol 16:1854–1866. https://doi.org/10.1111/1462-2920.12344.
5. Konstantinidis KT, Ramette A, Tiedje JM. 2006. The bacterial species definition in the genomic era. Philos Trans R Soc Lond B Biol Sci 361:1929–1940. https://doi.org/10.1098/rstb.2006.1920.
6. Cohan Frederick M, Kopac Sarah M. 2011. Microbial genomics: *E. coli* relatives out of doors and out of body. Curr Biol 21:R587–R589. https://doi.org/10.1016/j.cub.2011.06.011.
7. Rosselló-Mora R, Amann R. 2001. The species concept for prokaryotes. FEMS Microbiol Rev 25:39–67. https://doi.org/10.1111/j.1574-6976.2001.tb00571.x.
8. Hambright WS, Deng J, Tiedje JM, Brettar I, Rodrigues JLM. 2016. *Shewanella baltica* ecotypes have wide transcriptional variation under the same growth conditions. mSphere 1:e00158-16. https://doi.org/10.1128/mSphere.00158-16.

9. Vital M, Chai B, Ostman B, Cole J, Konstantinidis KT, Tiedje JM. 2015. Gene expression analysis of *E. coli* strains provides insights into the role of gene regulation in diversification. ISME J 9:1130–1140. https://doi.org/10.1038/ismej.2014.204.
10. Brem RB, Kruglyak L. 2005. The landscape of genetic complexity across 5,700 gene expression traits in yeast. Proc Natl Acad Sci U S A 102:1572–1577. https://doi.org/10.1073/pnas.0408709102.
11. Jung J, Park W. 2013. Comparative genomic and transcriptomic analyses reveal habitat differentiation and different transcriptional responses during pectin metabolism in *Alishewanella* species. Appl Environ Microbiol 79:6351–6361. https://doi.org/10.1128/AEM.02350-13.
12. Parkinson JE, Baumgarten S, Michell CT, Baums IB, LaJeunesse TC, Voolstra CR. 2016. Gene expression variation resolves species and individual strains among coral-associated dinoflagellates within the genus *Symbiodinium*. Genome Biol Evol 8:665–680. https://doi.org/10.1093/gbe/evw019.
13. Carreto L, Eiriz MF, Domingues I, Schuller D, Moura GR, Santos MAS. 2011. Expression variability of co-regulated genes differentiates *Saccharomyces cerevisiae* strains. BMC Genomics 12:201. https://doi.org/10.1186/1471-2164-12-201.
14. Ziemke F, Brettar I, Höfle MG. 1997. Stability and diversity of the genetic structure of a *Shewanella putrefaciens* population in the water column of the central Baltic. Aquat Microb Ecol 13:63–74. https://doi.org/10.3354/ame013063.
15. Brettar I, Rheinheimer G. 1991. Denitrification in the Central Baltic: evidence for H_2S-oxidation as motor of denitrification at the oxic-anoxic interface. Mar Ecol Prog Ser 77:157–169. https://doi.org/10.3354/meps077157.
16. Caro-Quintero A, Deng J, Auchtung J, Brettar I, Höfle MG, Klappenbach J, Konstantinidis KT. 2011. Unprecedented levels of horizontal gene transfer among spatially co-occurring *Shewanella bacteria* from the Baltic Sea. ISME J 5:131–140. https://doi.org/10.1038/ismej.2010.93.
17. Gao H, Barua S, Liang Y, Wu L, Dong Y, Reed S, Chen J, Culley D, Kennedy D, Yang Y, He Z, Nealson KH, Fredrickson JK, Tiedje JM, Romine M, Zhou J. 2010. Impacts of *Shewanella oneidensis* c-type cytochromes on aerobic and anaerobic respiration. Microb Biotechnol 3:455–466. https://doi.org/10.1111/j.1751-7915.2010.00181.x.
18. Liang Y-L, He Z-L, Gao H-C, Qiu G-Z, Zhou J-Z, Liu X-D. 2009. Characterization of cytochrome mutants for pellicle formation in *Shewanella onedensis* MR-1. Trans Nonferrous Met Soc China 19:700–706. https://doi.org/10.1016/S1003-6326(08)60336-2.
19. Brandt U, Trumpower B. 1994. The protonmotive q-cycle in mitochondria and bacteria. Crit Rev Biochem Mol Biol 29:165–197. https://doi.org/10.3109/10409239409086800.
20. Shirodkar S, Reed S, Romine M, Saffarini D. 2011. The octahaem SirA catalyses dissimilatory sulfite reduction in *Shewanella oneidensis* MR-1. Environ Microbiol 13:108–115. https://doi.org/10.1111/j.1462-2920.2010.02313.x.
21. Beliaev AS, Klingeman DM, Klappenbach JA, Wu L, Romine MF, Tiedje JM, Nealson KH, Fredrickson JK, Zhou J. 2005. Global transcriptome analysis of *Shewanella oneidensis* MR-1 exposed to different terminal electron acceptors. J Bacteriol 187:7138–7145. https://doi.org/10.1128/JB.187.20.7138-7145.2005.
22. Yoder-Himes DR, Chain PSG, Zhu Y, Wurtzel O, Rubin EM, Tiedje JM, Sorek R. 2009. Mapping the *Burkholderia cenocepacia* niche response via high-throughput sequencing. Proc Natl Acad Sci U S A 106:3976–3981. https://doi.org/10.1073/pnas.0813403106.
23. Philippe N, Crozat E, Lenski RE, Schneider D. 2007. Evolution of global regulatory networks during a long-term experiment with *Escherichia coli*. Bioessays 29:846–860. https://doi.org/10.1002/bies.20629.
24. Stern DL, Orgogozo V. 2008. The loci of evolution: how predictable is genetic evolution? Evolution 62:2155–2177. https://doi.org/10.1111/j.1558-5646.2008.00450.x.
25. Richards E, Reichardt M, Rogers S. 1994. Preparation of genomic DNA from plant tissue, p 2.3.1–2.3.7. *In* Ausubel M, Brent R, Kingston RE, Moore DD, Seidman JG, John A, Smith JA, Struhl K (ed), Current protocols in molecular biology, supplement 27. John Wiley & Sons, Hoboken, NJ.
26. Cruz-García C, Murray AE, Klappenbach JA, Stewart V, Tiedje JM. 2007. Respiratory nitrate ammonification by *Shewanella oneidensis* MR-1. J Bacteriol 189:656–662. https://doi.org/10.1128/JB.01194-06.
27. Tusher VG, Tibshirani R, Chu G. 2001. Significance analysis of microarrays applied to the ionizing radiation response. Proc Natl Acad Sci U S A 98:5116–5121. https://doi.org/10.1073/pnas.091062498.