



# Why Close a Bacterial Genome? The Plasmid of *Alteromonas Macleodii* HOT1A3 is a Vector for Inter-Specific Transfer of a Flexible Genomic Island

Eduard Fadeev<sup>1</sup>, Fabio De Pascale<sup>2</sup>, Alessandro Vezzi<sup>2</sup>, Sariel Hübner<sup>3,4</sup>, Dikla Aharonovich<sup>1</sup> and Daniel Sher<sup>1\*</sup>

<sup>1</sup> Department of Marine Biology, Leon H. Charney School of Marine Sciences, University of Haifa, Haifa, Israel,

<sup>2</sup> Department of Biology and CRIBI Biotechnology Centre, University of Padua, Padova, Italy, <sup>3</sup> Department of Botany and Biodiversity Research Centre, University of British Columbia, Vancouver, Canada, <sup>4</sup> The Department of Evolutionary and Environmental Biology, University of Haifa, Haifa, Israel

## OPEN ACCESS

### Edited by:

Rex Malmstrom,  
U.S. Department of Energy Joint  
Genome Institute, USA

### Reviewed by:

Kyung-Hoon Shin,  
Hanyang University, South Korea  
Jim Tripp,  
University of California, Davis, USA

### \*Correspondence:

Daniel Sher  
dsheer@univ.haifa.ac.il

### Specialty section:

This article was submitted to  
Aquatic Microbiology,  
a section of the journal  
Frontiers in Microbiology

Received: 15 November 2015

Accepted: 15 February 2016

Published: 08 March 2016

### Citation:

Fadeev E, De Pascale F, Vezzi A,  
Hübner S, Aharonovich D and Sher D  
(2016) Why Close a Bacterial  
Genome? The Plasmid  
of *Alteromonas Macleodii* HOT1A3 is  
a Vector for Inter-Specific Transfer of a  
Flexible Genomic Island.  
*Front. Microbiol.* 7:248.  
doi: 10.3389/fmicb.2016.00248

Genome sequencing is rapidly becoming a staple technique in environmental and clinical microbiology, yet computational challenges still remain, leading to many draft genomes which are typically fragmented into many contigs. We sequenced and completely assembled the genome of a marine heterotrophic bacterium, *Alteromonas macleodii* HOT1A3, and compared its full genome to several draft genomes obtained using different reference-based and *de novo* methods. In general, the *de novo* assemblies clearly outperformed the reference-based or hybrid ones, covering >99% of the genes and representing essentially all of the gene functions. However, only the fully closed genome (~4.5 Mbp) allowed us to identify the presence of a large, 148 kbp plasmid, pAM1A3. While HOT1A3 belongs to *A. macleodii*, typically found in surface waters (“surface ecotype”), this plasmid consists of an almost complete flexible genomic island (fGI), containing many genes involved in metal resistance previously identified in the genomes of *Alteromonas mediterranea* (“deep ecotype”). Indeed, similar to *A. mediterranea*, *A. macleodii* HOT1A3 grows at concentrations of zinc, mercury, and copper that are inhibitory for other *A. macleodii* strains. The presence of a plasmid encoding almost an entire fGI suggests that wholesale genomic exchange between heterotrophic marine bacteria belonging to related but ecologically different populations is not uncommon.

**Keywords:** *Alteromonas*, genomic island, plasmid, genome assembly, metal resistance

## INTRODUCTION

The development of next generation sequencing (NGS) technologies has introduced new and much cheaper methods for prokaryotic whole genome sequencing (Quail et al., 2012). In combination with the decreasing price for computational resources needed for bioinformatic analyses, this has resulted in the publication of a large number of new bacterial genomes (Metzker, 2010). More and more “non-bioinformatic” labs are now sequencing their own prokaryotic genomes and facing questions regarding the project such as which sequencing platform to choose, how many libraries should be generated, which assembly method should be used for the libraries, etc. Some of these

questions have been addressed in the literature (Loman et al., 2012; Magoc et al., 2013; Ekblom and Wolf, 2014), while others remain unanswered.

One of the main questions facing each new bacterial sequencing project is whether or not the planned use for the genome requires it to be closed into a single, high quality sequence for each DNA molecule (genome or plasmid, defined as “complete” status). The alternative is often a draft genome, which is fragmented into multiple contigs or has many ambiguous bases (N’s). For example, if the genome in question is to be compared to the genomes of other organisms to infer the genetic basis for different physiological characteristics or ecological niches, will such a comparison be biased, for example through fragmented representation of function-related genomic regions such as operons, genomic islands or extra-chromosomal elements such as plasmids (Fernández-Gómez et al., 2012)? Notably, as of October 2015 the curated genome database of the National Center for Biotechnology Information (NCBI) contained 49,204 bacterial genomes, yet only 10% of them are defined as “complete genomes”. All the rest are defined as “draft genomes” which range from one contig with many N’s (scaffold) to a large number of contigs whose order is unknown. The difficulty in assembling a closed genome despite the high coverage typically obtained with NGS techniques is usually caused by repetitive sequences including both highly conserved core genes (e.g., ribosomal and tRNA genes) and non-core (accessory) genes which are often harbored on mobile genetic elements such as plasmids and transposons (Loman et al., 2012). Such sequences often produce “break points” in the assembly, resulting in multiple contigs (Hunt et al., 2014). Alternatively, when assembled, such multicopy genes might be collapsed into a single gene sequence (Ekblom and Wolf, 2014), thus missing genetic microdiversity which may be functionally important (Macrander et al., 2015).

Here we present a comparison of four commonly used approaches which we applied for the sequencing and assembly of the medium-sized genome of a heterotrophic marine bacterium, *Alteromonas macleodii* HOT1A3 (Sher et al., 2011): (I) Reference-based assembly using short Illumina reads. In this approach, which is likely the most simple and inexpensive one and can be performed using user-friendly programs with graphic user interface such as Geneious (Kearse et al., 2012), the genome of a closely-related organism is used as a template for assembly. Underlying this approach is the assumption that the reference genome is highly similar to the one being sequenced in gene content and gene order (although this is usually rigorously tested by comparing the generated contigs to the reference). This approach is often used for closely-related strains of pathogenic bacteria (e.g., Mutreja et al., 2011; Loman et al., 2012; Powers et al., 2013); (II) *De novo* assembly using short paired-end reads produced using Illumina platforms. This approach requires more computational power and bioinformatics expertise, but does not rely on a reference genome (see Kisand and Lettieri, 2013; Magoc et al., 2013, for overviews and detailed bioinformatic comparisons of different *de novo* assembly methods). We present results using two different libraries, each assembled

using a different program (Table 1); (III) A hybrid approach combining the reference-based and *de novo* approaches, aimed at utilizing the *de novo* assembled contigs to close major gaps in the reference-based scaffold. These gaps often contain functionally and ecologically important genetic material (e.g., Genomic islands, Juhas et al., 2009). This approach has been successfully used to assemble the genomes of other strains of *Alteromonas*, our model organism (see below, López-Pérez et al., 2012, 2014); (IV) *De novo* assembly using long sequencing reads produced by the Pacific Biosciences (PacBio) RS II platform (English et al., 2012; Koren et al., 2013). PacBio requires more DNA and of higher quality than Illumina sequencing and relatively advanced computational resources which nevertheless are available on the Amazon Machine Image (<http://aws.amazon.com>, see details in the materials and methods section).

*Alteromonas macleodii*, the clade to which our model organism belongs, are opportunistic copiotrophic bacteria belonging to the  $\gamma$ -Proteobacteria, and represent approximately 3–5% of the bacterial population in ocean surface waters (Baumann et al., 1972; Ivanova et al., 2004; Nelson and Wear, 2014). They are highly specialized in adaptation to a wide range of carbon sources, and can be found both free-living and particle associated (on marine snow, biofilms *etc.*; Acinas et al., 1999; García-Martínez et al., 2002). *A. macleodii* are often among the first organisms to respond to perturbations, e.g., during experiments aimed to simulate mixing or DOM addition (McCarren et al., 2010; Shi et al., 2012) and have been suggested to play an important role in carbon cycling (Baker et al., 2013; Nelson and Wear, 2014; Pedler et al., 2014). *A. macleodii* and related strains also affect the growth of other microorganism, inhibiting the growth of several species of algae (Mayali and Azam, 2004) and taking advantage of the dying cells (Sheik et al., 2014). Our specific model organism, *A. macleodii* strain HOT1A3, was selected due its contrasting effect on the growth of two different *Prochlorococcus* strains in laboratory co-culture (Sher et al., 2011).

Until recently the *A. macleodii* strains whose genomes have been sequenced were divided into two major groups, “surface ecotypes” and “deep ecotypes” (López-López et al., 2005). Strains belonging to the surface ecotype have been isolated from many locations around the globe, whereas those belonging to the deep ecotype have all been isolated from the Mediterranean Sea in which the temperature of the deep waters never drops below 12°C (López-Pérez et al., 2013). Further phylogenetic analysis of the strains resulted in the re-naming of the “deep ecotype” strains as a new species, *Alteromonas mediterranea* (Ivanova et al., 2015).

While the genomes of these species are highly syntenous (i.e., they share very similar gene organization; Ivars-Martinez et al., 2008), they differ in the gene content of several flexible genomic islands (fGIs). These genomic islands, located at equivalent genomic locations (López-Pérez et al., 2013, 2014), potentially contribute to adaptation and evolution of the organism by gene exchange processes (Klockgether et al., 2004; Ullrich et al., 2005; Rodríguez-Valera et al., 2009; Avrani et al., 2011).

With this background, as we sequenced and assembled the genome of *A. macleodii* HOT1A3, we compared different

**TABLE 1 | General features of all genome assemblies for *Alteromonas macleodii* strain HOT1A3.**

Sequencing platform	PacBio RS II	Illumina HiSeq 2000	MIRA <i>de novo</i>	Illumina GAIIX	Reference based assembly
Assembly method	HGAP <i>de novo</i>	VELVET <i>de novo</i>	MIRA <i>de novo</i>	Hybrid assembly	Reference based assembly
Genome size (bp)	Genome 4,652,873 [Plasmid 148,912]	4,715,275	4,707,271	4,371,362	4,654,302
Library	399,576 SE long reads	1,810,538 150 bp PE reads		19,859,116 72 bp PE reads	
Mean Illumina library insert size (bp)	–	475 bp		228 bp	
GC%	44.67 [42.3]	44.6	44.63	44.86	44.95
Contigs	1 [1]	122	834	12	114
N50	–	138,989	8891	682,590	80,406
Subsystems	477 [13]	483	470	469	467
CDS	4074 [178]	4178	4893	3796	3679
Hypothetical proteins	946 [95]	1032	1234	824	778
RNAs	87 [0]	80	65	76	68
Complete rDNA operons	5 [0]	0	1	4	5

\*Brackets [] refer to features related to plasmid pAM1A3.

\*\*All annotation features are based on automatic prokaryotic genome annotation pipeline (RAST, Overbeek et al., 2005).

sequencing and assembly approaches, aiming to identify the costs and benefits of each technique that would provide guidance to our lab and others when embarking on future bacterial genome sequencing projects. Our results show that, compared to the fully closed genome of *Alteromonas* HOT1A3 obtained using PacBio, the genomes assembled *de novo* from short Illumina reads were fragmented but still covered almost all of the genetic data on the full genome. In contrast, using a reference-based assembly resulted in much higher “loss” of genomic data. Nevertheless, the advantage of closing a genome were evident in that, only in this assembly, we could bioinformatically identify the presence of a 148 kbp plasmid, and link this plasmid to almost an entire genomic island previously observed exclusively in the genomes of *A. mediterranea* strains.

## MATERIALS AND METHODS

### *Alteromonas Macleodii* HOT1A3 Isolation and Maintenance

*Alteromonas* HOT1A3 was previously isolated in 2007 from a water sample collected at the HOT station ALOHA (22°45' N 158° W) at a depth of 5 m during the C-More BLOOMER cruise on August 18, 2007 (Sher et al., 2011). The water sample was transported back to the Chisolm lab at MIT and maintained for 1 month in clear polycarbonate bottles at 21°C under 18 μE constant cold white illumination. The strain was isolated by plating on solid ProMM plates followed by six rounds of streaking and colony picking on the same media. ProMM is based on the Pro99 media used to culture *Prochlorococcus* (Moore et al., 2007), with the addition of a set of defined organic compounds (0.05% w/v each of Lactate, Pyruvate, Acetate and Glycerol) and vitamins. The isolate was initially identified as *A. macleodii* based on the sequence analysis of 16s rRNA gene fragment. The strain

is routinely cultured in ProMM media and maintained as frozen stocks at –80°C in 25% (v/v) glycerol in ProMM.

### DNA Extraction and Sequencing

DNA was extracted using the phenol chloroform protocol (Neumann et al., 1992). The genome was sequenced separately in three sequencing runs: (1) Using an Illumina GAIIX platform (72 bp paired-end read). The libraries for this run were produced as follows: DNA was sheared by sonication using a Bio-Ruptor ultrasonicator (Diagenode), end-repaired using the Quick Blunting kit from New England Biolabs, and Illumina adaptors were ligated using the quick ligate kit from Enzymatics followed by PCR enrichment. Size selection was performed using SPRI beads to obtain fragments with a mean size of 228 bp; (2) Using an Illumina HiSeq2000 platform (150 bp paired-end read). The libraries for this sequencing run were produced at the University of Haifa Bioinformatics Support Unit using the NEBNext Ultra DNA Library Prep Kit for Illumina followed by size selection using SPRI beads to obtain fragments with a mean size of 475 bp; (3) Using a Pacific Biosciences RS II platform, with the libraries produced by the DNA template prep kit 3.0 from Pacific Biosciences, at the Yale Center for Genome Analysis.

### Genome Assembly

Each of the different genome assemblies was obtained using a combination of sequencing libraries and assembly programs as shown in **Table 1**: (1) The Illumina GA library was assembled *de novo* using MIRA short-reads assembler version 3.4.0 with default parameters (Chevreux et al., 2004). In addition, we attempted to assemble the genome using a reference genome of closely related strain. We chose *A. macleodii* ATCC27126 strain as a reference due to the highest percentage of mapped reads and the highest coverage across the entire genome, compared to other published genomes of *A. macleodii*.

In order to close the gaps in the produced scaffold, we mapped the contigs generated by MIRA to the reference-based assembly using the reference mapping function of Geneious 7 (Kearse et al., 2012), followed by manual validation of the results (termed as “hybrid assembly”). (2) The Illumina HiSeq2000 *de novo* assembly was obtained using a custom pipeline. The initial assembly was performed using VELVET assembler (v. 1.2.10; Zerbino and Birney, 2008) with a *k*-mer size of 81 bases, considering the reads as paired. The assembly was evaluated using the REAPR tool (v. 1.0.16) with standard parameters (Hunt et al., 2013). The resulting corrected contigs were arranged into scaffolds using SSPACE scaffolder algorithm (v. 3.0) with default options (Boetzer et al., 2011). Additional evaluation using REAPR was performed to make sure that no mis-assembly was introduced. Gaps within the scaffolds were filled using GapFiller (v. 1.10; Boetzer and Pirovano, 2012), specifying the following options -m 80 and -o 5 (these options specify for: m, the minimum number of overlapping bases with the sequences around the gap, and o the minimum number of reads needed to call a base during an extension). As final step after the gap filling procedure we ran REAPR to assess that the final assembly was consistent and no error was introduced by GapFiller. (3) The PacBio library was assembled using the Hierarchical Genome Assembly Process (HGAP, Chin et al., 2013), integrated in the SMRT analysis software (version 2.3.0) from Pacific Biosciences. The contigs were closed into circular molecules by alignment of 10 kbp from both ends of the contig to each other, using the Geneious 7 alignment tool (Kearse et al., 2012).

## Gene Prediction and Annotation

All assemblies were annotated using Rapid Annotation using Subsystem Technology (RAST) prokaryotic genome annotation server (Aziz et al., 2008; Brettin et al., 2015). The genome and plasmid generated by PacBio sequencing were additionally annotated based on all published genomes of *A. macleodii* in NCBI genome database (NC\_011138.3, NC\_018632.1, NC\_018678.1, NC\_018692.1, NC\_018679.1, NC\_019393.1, NC\_023045.1, NC\_021717.1, NC\_021712.1, NC\_021713.1, NC\_021714.1, CP004846.1, CP004849.1) using Prokka (version 1.10) genome annotation tool (Seemann, 2014). Both annotations were merged together by comparison of location of the genes (coordinates on the genome). The final annotation consisted of 4215 genes in the genome and 184 genes in the plasmid. Gene ontology (GO) terms were assigned to the CDS using InterProScan5 (Jones et al., 2014). In addition, due to the discovery of several metal resistance related operons on the plasmid, the annotation of all its genes was verified using BLASTP (Moreno-Hagelsieb and Latimer, 2008) to BacMet dataset – a high quality, manually curated database of experimentally confirmed antibacterial biocide- and metal-resistance genes (Pal et al., 2014). All genomic maps of genome and plasmid annotations and comparisons were generated using BRIG visualization tool (Alikhan et al., 2011) and Circos visualization tool (Krzywinski et al., 2009).

## Comparison of Genome Assembly Products

In order to assess the quality of each one of the assembly approaches we used, the products of each assembly was compared to the final genome and plasmid sequences. Genome fraction calculations of each assembly method was conducted using the QUality ASsessment Tool for genome assembly (Gurevich et al., 2013). The number of fully contained proteins in the contigs of the fragmented assemblies was calculated aligning all the annotated proteins from the final genome and plasmid using tBLASTn to all ORFs of each one of other assemblies, where only completely aligned proteins were counted (Magoc et al., 2013). A statistical comparison between the assemblies was conducted using Fischer's exact test of gene counts assigned to different subsystems according to SEED categorization system, which organizes gene functional category into 5 hierarchical levels (Overbeek et al., 2005).

## Comparison of *Alteromonas* HOT1A3 to published *A. macleodii* Genomes

All published genomes of *A. macleodii* were re-annotated using RAST annotation pipeline. Using BLASTn alignment (coverage of at least 80% of the total length of the gene with an *e*-value  $\leq 0.0001$ ) we detected the core genes which were present in all the genomes, as well as “surface ecotype” and “deep ecotype” related genes. All alignments were conducted using Geneious 7 aligner (Kearse et al., 2012) and phylogenetic trees were generated using MEGA6 software (Tamura et al., 2013). Functional comparison between the genomes was conducted using principal component analysis of gene counts assigned to different subsystems according to SEED categorization system.

## Pulse Field Gel Electrophoresis

To validate the presence of the bioinformatically-identified pAM1A3 plasmid, pulse field gel electrophoresis (PFGE) was carried out (Hightower et al., 1987; Mathew et al., 1988). In order to extract only the plasmid DNA, DNA was extracted using the NucleoBond<sup>®</sup> BAC 100 kit (Blaas et al., 2012). In order to verify the actual size of the plasmid, it was cut using three different restriction enzymes: NotI, SgrAI, PacI (Barth et al., 2009). PFGE of the extracted DNA was performed using CHEF DRIII device with 1% agarose gels and 0.5X TBE buffer at 14°C for 11.5 h. Lambda ladder PFGE marker, low range ladder PFGE marker and 1 KB DNA ladder marker were used as molecular size markers. The gel was stained using GelRed stain (Biotum, Hayward, CA 94545) and photographed over an UV transilluminator (López-Pérez et al., 2012; Marasini and Fakhr, 2014; Ozdemir and Acar, 2014).

## Heavy Metal Resistance Assays

Heavy metal resistance was assessed by growth experiments in ProMM medium with heavy metals (mercury chloride, zinc acetate, or copper chloride) added at different concentrations (0.1, 0.25, 0.5, 0.75, 1 nM for zinc and copper and 0.001, 0.01, 0.05, 0.1 mM for mercury). The cultures were grown in 96 well plates at 20 ± 2°C inside a Perkin-Elmer Enspire multimode

plate reader with the optical density at 600 nm measured every 1 h. The minimal inhibitory concentration was defined as the concentration where no growth was observed after 36 h.

## Accession Numbers

The sequences have been deposited in GenBank under the genomic accession numbers CP012202 for the *A. macleodii* HOT1A3 genome and CP0122203 for the plasmid pAM1A3.

## RESULTS

### The Fully Assembled Genome of *A. macleodii* HOT1A3

Aiming to obtain a completely closed genome for *Alteromonas* HOT1A3, we conducted several assembly attempts using different DNA-seq libraries and different assembly approaches (Table 1). The assembly of the PacBio library generated two contigs, sized 4.6 Mbp and 148 kbp, which we could close into two circular DNA molecules (Figure 1). Pulse-field gel electrophoresis combined with restriction enzyme analysis revealed the presence of two closely-related plasmids, differing by about 6 kbp (Supplementary Figure S1). Two regions in the plasmid exhibited increased sequencing coverage and a high number of single nucleotide polymorphisms (SNPs), suggesting that these regions are duplicated, potentially in tandem, on some of the plasmids (Supplementary Figure S1). The number of total SNPs identified in the sequences was 697 for the genome and 65 for the plasmid, attesting to the high quality and homogeneity of the genome. The genome too had a region of ~70 kbp with approximately double coverage, which did not have a higher density of SNPs compared to the rest of the genome. This region may also represent a tandem repeat.

Using the 16S rDNA gene as a phylogenetic marker, *Alteromonas* HOT1A3 belongs to the *A. macleodii* (Figure 2A). This was supported by a phylogenetic analysis of eleven concatenated core genes (Supplementary Figure S2) as well as by whole-genome comparison, revealing that the genome of HOT1A3 has 97% nucleotide identity with the genome of *A. macleodii* ATCC 27126, which is considered to be the type strain for *A. macleodii* (Figure 2B). Notably, these two strains were both isolated near Hawaii > 35 years apart (Baumann et al., 1972; Sher et al., 2011). Finally, the classification of *Alteromonas* HOT1A3 as belonging to the *A. macleodii* species rather than to *A. mediterranea* was also evident by the number of genes shared by *Alteromonas* HOT1A3 with one of these species but not the other (see red and blue gene layers in Figure 1). We use the fully assembled genome, produced from PacBio reads, as a benchmark to assess the results of all other assembly methods.

### Comparison of the Genomes Obtained Using Different Assembly Methods

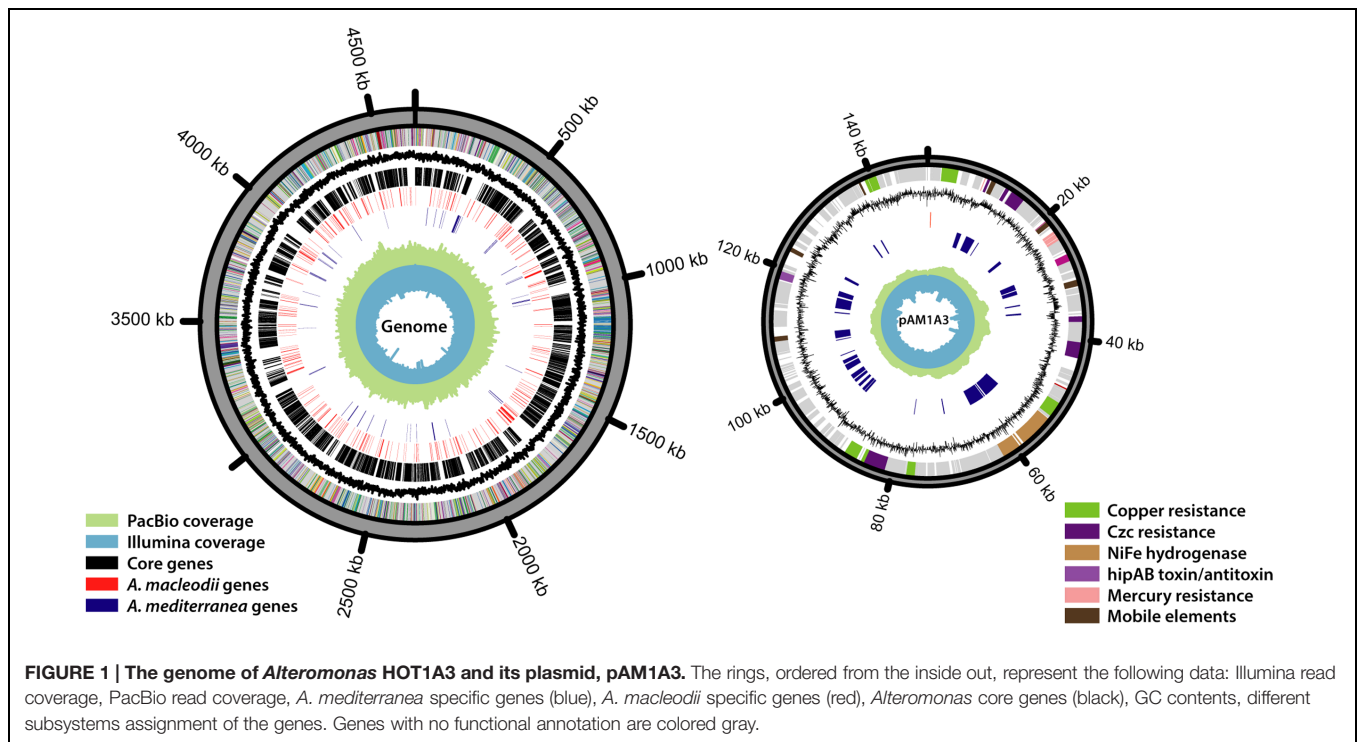
The predicted genome size of the various *de novo* assemblies was quite close to that of the closed genome with its plasmid, despite their fragmentation into hundreds of contigs (Table 1). In contrast, the reference-based assembly and the hybrid

assembly both predicted somewhat smaller genomes. Of all the fragmented assemblies, the one produced using longer Illumina reads with a larger insert size, and assembled *de novo* using VELVET, had the closest number of genomic features (coding sequences [CDSs], hypothetical proteins and RNAs) to the final PacBio-sequenced genome and plasmid (Table 1). However, the fragmented nature of this assembly resulted in it not containing completely assembled ribosomal operons.

We next assessed what specific and functional genetic information had been lost in each of the draft genomes compared to the closed PacBio genome. We used three metrics for this comparison: a) the length of the nucleotide sequence of the genome and plasmid combined (a total of 4.8 Mbp, Figure 3A); (b) the number of complete predicted genes (Coding DNA Sequences or CDSs, Figure 3A); (c) the representation of functional pathways, defined using the SEED subsystems approach (Figure 3B; Aziz et al., 2008). The *de novo* assemblies clearly outperformed the hybrid and reference-based ones, covering almost the entire genome both in terms of nucleotide sequence and in terms of CDSs (Figure 3A). In all of the assemblies the fraction of complete CDSs was higher than the fraction of the nucleotide sequence itself, suggesting that at least some of the sequences hindering assembly are found in intergenic regions. While the representation of most of the functional groups was very similar between the different assemblies, there were five functional subsystems which were significantly underrepresented within the reference based assemblies but not in the *de novo* assemblies (Figure 3B, Supplementary Figure S3, Fisher's Exact Test,  $p = 0.05$ ). The missing genes in these subsystems are involved in metal resistance, hydrogenase activity and RNA metabolism, and the majority of them are found on the plasmid (see below). Thus, while the *de novo* assemblies capture almost all of the genes and pathways present in the closed genome, the reference-based and hybrid assemblies had reduced representation of several potentially important genes and pathways. This could lead to biases or errors in attempting to interpret how the functions encoded in the genome relate to the organism's life history or ecology, especially in reference to the effect of heavy metals.

### *Alteromonas* HOT1A3 Hosts a Plasmid – pAM1A3

One unique genomic feature, which emerged from the closed genome but not from the draft ones, was the presence in *Alteromonas* HOT1A3 of a plasmid. The plasmid, termed pAM1A3, is ~148 kbp long and encodes 184 genes (Figure 1). Among these genes there are several plasmid-related features such as genes encoding the bacterial *parA*- partitioning protein, which ensures proper distribution of plasmid replicates to daughter cells during cell division (Bignell and Thomas, 2001) and the methylation subunit of a restriction modification system which may protect the plasmid from specific restriction enzymes (Tock and Dryden, 2005). The plasmid also encodes a putative *hipAB* toxin-antitoxin gene pair, which consists of two genes,



*hipA* which encodes a stable toxin and *hipB* which encodes an unstable antitoxin in form of a repressor for the *hipAB* operon (Wen et al., 2014). This mechanism develops an “addiction” of the cell by constant need of antitoxin expression, and may represent a mechanism for the maintenance of the plasmid despite the related metabolic cost (Yamaguchi et al., 2011). The gene *hipA* was not found in any of the *A. macleodii* published genomes, whereas a *hipB*-like gene is found on a plasmid which was isolated from *A. mediterranea* AltDE1 strain (Gonzaga et al., 2012). The plasmid does not contain genes involved in conjugation, but many genes involved in the biosynthesis of type IV secretion systems (pili) are found on the genome.

There are several transcriptional regulators found on the plasmid, among them member of the *marR* family of regulatory proteins, which are known to regulate the expression of multidrug efflux pumps (Nikaido, 1998; Alekshun et al., 2001), and three copies of *merR*-family transcriptional regulator, which is a common transcription regulation system often associated with plasmids (Silver and Walderhaug, 1992).

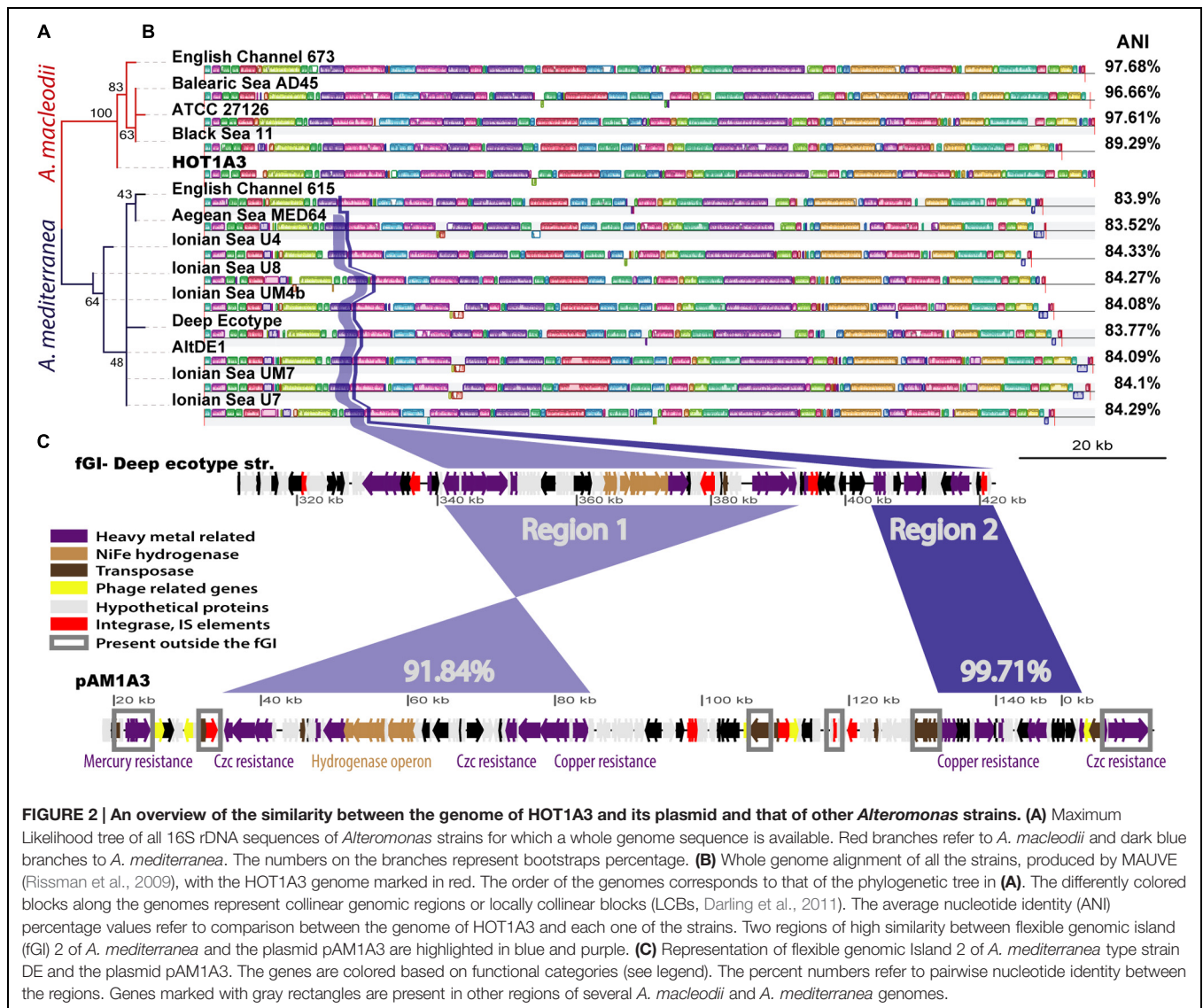
Among the functional genes encoded on the plasmid, many belong to several operons associated with metal resistance (Figure 2C). The mercury resistance operon *merRTPCA*, includes the regulatory protein *merR*, and three mercury transport proteins *merT*, *merB* and *merC* (Summers, 1992). The operon found on the plasmid also contains *merA*, a key enzyme with a role in mercury detoxification, yet lacks the lyase subunit *merB* which may be necessary for the activity of *merA* (Huang et al., 1999). The plasmid also encodes the copper resistance genes *copABCDG*, as well as a two-component regulatory module *cusRS*, which typically regulates the *cop* operon (Grass and Rensing, 2001; Bouzat and Hoostal, 2013). Finally, there are

two operons of cobalt-zinc-cadmium resistance (*czcABC*) which encode the components of heavy metal efflux pumps (Silver and Walderhaug, 1992).

In addition to the heavy metal resistance operons, the plasmid also hosts a complete gene cluster encoding a *NiFe* hydrogenase complex which is completely absent from all published *A. macleodii* genomes, with an exception of the draft genome of *A. macleodii* MIT1002 (Weyman et al., 2011; Biller et al., 2015). pAM1A3 also has several transposable elements which are associated with horizontal gene transfer such as Tn3-related resolvase and transposase genes, IS66-family and Tn21-family transposases (Liebert et al., 1999; Han et al., 2001; Parks and Peters, 2007).

## Plasmid pAM1A3 Contains Almost an Entire Flexible Genomic Island

Some *A. macleodii* and *A. mediterranea* strains harbor plasmids, however, as shown in **Supplementary Figure S4**, none of these plasmids revealed high similarity to pAM1A3, with an exception of a small region of 7 kbp which encode Tn3-related resolvase and transposase genes. Instead, more than half of the functional genes on the pAM1A3 plasmid have close homologs in all the genomes of *A. mediterranea*, and almost no representation in the genomes of *A. macleodii* strains. These functional genes, which include many of the metal resistance genes as well as the hydrogenase genes described above, belong to two almost contiguous regions (Figure 2 and **Supplementary Figure S4**), both of which correspond to a fGI previously identified in strains belonging to the *A. mediterranea* species (described as the “Deep ecotype” in Figure 2C; Ivars-Martinez et al., 2008; Gonzaga et al., 2012) and in one of two strains of *A. australica* (López-Pérez

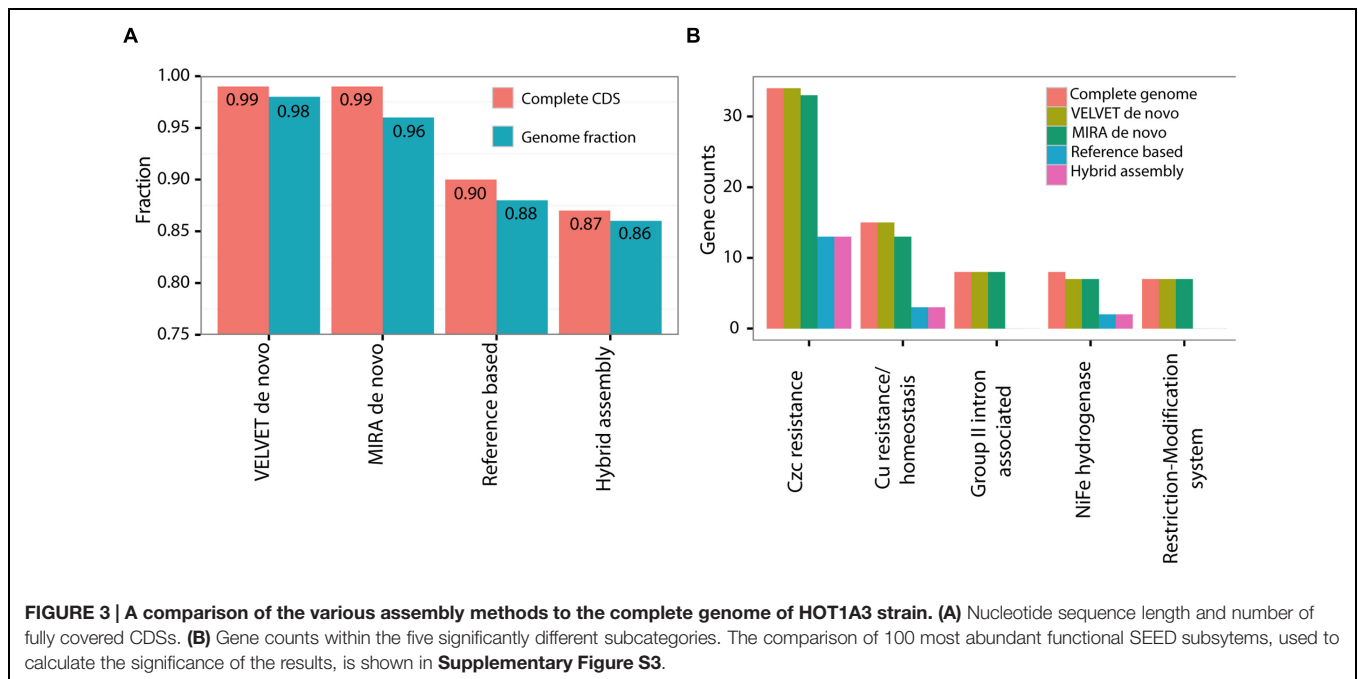


et al., 2014). Remarkably, a similar genomic island was also found on a plasmid hosted by *Glaciecola* sp. strain 4H-3-7 + YE-5 (pGLAAG01, **Supplementary Figure S5**, Klippel et al., 2011; López-Pérez et al., 2014).

In terms of its gene content and organization, the region encoding these genes on the plasmid is closest to fGI-2 of the *A. mediterranea* type strain DE, containing ~75% of the genes (~80% of the fGI sequence) and revealing ~91.8–99.7% nucleotide sequence identity (**Figure 2C**). Thus, it is tempting to speculate that the genes on the plasmid were acquired almost wholesale from a strain closely related to DE strain. Nevertheless, phylogenetic analysis of several of the genes on the plasmid, e.g., *copA*, *merR*, and *czcA*, showed that the phylogeny of the genes on the plasmid is not homogenous, and that some are phylogenetically closer to other strains of *A. mediterranea* and not to DE (**Supplementary Figure S6**). It is currently unclear whether this reflects the evolutionary history of the fGI on this hypothetical donor strain or whether the acquisition of novel

genetic information and/or its rearrangement occurred on the plasmid.

Given that, with its plasmid, *Alteromonas* HOT1A3 has a mixture of features associated with *A. macleodii* and *A. mediterranea*, we asked to which of these groups strain *Alteromonas* HOT1A3 is more closely related in terms of functional metabolic capacity. We re-annotated the genomes of the previously described *A. macleodii* and *A. mediterranea* strains using the same pipeline we used for *Alteromonas* HOT1A3, and performed a principle component analysis (PCA) of the genomes, based on their functional capacity (the SEED subsystems, see Materials and Methods). As shown in **Figure 4A**, a strong clustering of the genomes into two distinct groups is observed, one containing *A. mediterranea* strains and one containing *A. macleodii* strains. In this clustering *Alteromonas* HOT1A3 clearly partitions with *A. macleodii*, in agreement with the phylogenetic analysis and the genome-wide comparisons (**Figures 1 and 2**). However, when the PCA was performed using



only the subsystems associated with the plasmid, *Alteromonas* HOT1A3 clustered with the *A. mediterranea* (Figure 4B). This suggests that, while clearly belonging to the *A. macleodii* species, for some specific functional features, e.g., metal resistance, *A. macleodii* HOT1A3 may have acquired through its plasmid some of the functional potential of *A. mediterranea* strain.

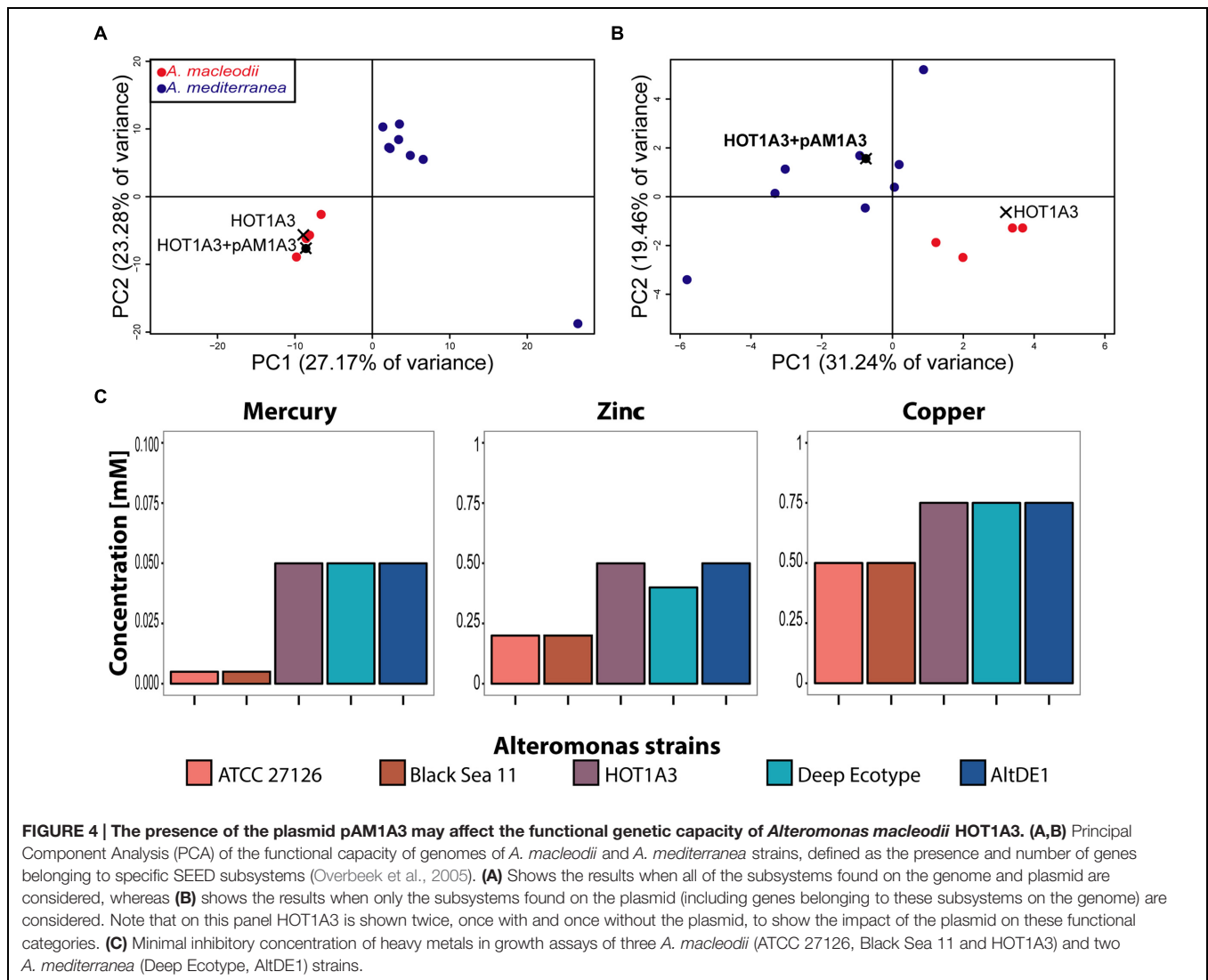
To assess this hypothesis, we grew representative strains of both *A. macleodii* (ATCC 27126 and Black Sea 11 – BS11; López-Pérez et al., 2012) and *A. mediterranea* (DE, AltDE1; Gonzaga et al., 2012) in ProMM media with different concentrations of heavy metals (zinc, mercury, and copper). As shown in Figure 4C, the two *A. macleodii* strains, ATCC 21726 and BS11 were more sensitive to heavy metals (a lower minimum inhibitory concentrations, MIC) than the *A. mediterranea* strains, in agreement with previously published metal resistance comparison (Ivars-Martinez et al., 2008). In contrast, *A. macleodii* HOT1A3, which contains the plasmid with the metal resistance genes, was as resistant to the heavy metals as the *A. mediterranea* strains, supporting the hypothesis that the metal resistance genes on the plasmid are functional.

## DISCUSSION

Horizontal gene exchange is of importance in the adaptation and evolution of bacteria, and the involvement of plasmids in exchange of genetic material and shaping of fGIs in pathogenic bacteria is well recognized (recently reviewed by Juhas et al., 2009; Morales-Espinosa et al., 2012). In marine bacteria, the role of phage-mediated lateral gene transfer in shaping bacterial genomes has been extensively explored (reviewed by Rodriguez-Valera et al., 2009; Rohwer and Thurber,

2009; Breitbart, 2012), perhaps due to the expectation that in dilute seawater transduction, rather than conjugation, is a more likely form of genetic exchange. Nevertheless, the presence of plasmids has long been recognized in marine bacteria (Sizemore and Colwell, 1977), and they are involved in ecologically-important aspects such as growth under low iron conditions, adherence of heterotrophic bacteria to algae and pathogenicity (e.g., Frank et al., 2015; Lee et al., 2015; Osorio et al., 2015). Additionally, the exchange of plasmids via conjugation has been demonstrated in batch seawater (e.g., Dahlberg et al., 1998). In this regard, previous studies have suggested that many of the genomic islands in *A. macleodii* and *A. mediterranea*, and especially the metal resistance related island fGI2, are constantly being re-shaped by local insertion and deletion events, and are thus termed “additive” (Gonzaga et al., 2012). Our observation of a plasmid, pAM1A3, that contains almost this entire genomic island, suggests an alternative mechanism, whereby large amounts of genetic material are transferred wholesale via plasmids. Indeed, previous studies of the genomes of *A. macleodii* and *A. mediterranea* strains have suggested plasmid mediated insertions of three genomic regions, two relatively small (19–21 kbp) and one large (65 kbp), the latter encoding a hybrid polyketide/non-ribosomal peptide synthases gene cluster (López-Pérez et al., 2013). Kalhoefer et al. (2011) recently observed a similar case in a different group of marine bacteria, whereby a relatively large plasmid (~94 kbp) from *Roseobacter litoralis* was highly similar to a genomic region from a related strain, *R. denitrificans* (Kalhoefer et al., 2011). This suggests that plasmid-mediated gene exchange, including transfer of large amounts of genetic material from the genomes into plasmids and vice-versa, is common in at least some lineages of heterotrophic marine bacteria. Whether such genetic exchange occurs when the cells





**FIGURE 4 | The presence of the plasmid pAM1A3 may affect the functional genetic capacity of *Alteromonas macleodii* HOT1A3. (A,B)** Principal Component Analysis (PCA) of the functional capacity of genomes of *A. macleodii* and *A. mediterranea* strains, defined as the presence and number of genes belonging to specific SEED subsystems (Overbeek et al., 2005). **(A)** Shows the results when all of the subsystems found on the genome and plasmid are considered, whereas **(B)** shows the results when only the subsystems found on the plasmid (including genes belonging to these subsystems on the genome) are considered. Note that on this panel HOT1A3 is shown twice, once with and once without the plasmid, to show the impact of the plasmid on these functional categories. **(C)** Minimal inhibitory concentration of heavy metals in growth assays of three *A. macleodii* (ATCC 27126, Black Sea 11 and HOT1A3) and two *A. mediterranea* (Deep Ecotype, AltDE1) strains.

are free-living (planktonic) or particle associated is currently unknown.

How common is the presence of a pAM1A3-like plasmid, with the genetic information it encodes, in Alteromonads in nature? Previously, it has been shown that the global ocean survey (GOS) metagenome contains many sequences related to the metal resistance genes found in the metal resistance fGI, and in fact these genes were more common than the core genes of this strain (Ivars-Martinez et al., 2008). Recently, another *A. macleodii* strain, MIT1002, has been shown to contain some genes found in *A. mediterranea*, including metal resistance genes and a hydrogenase (Biller et al., 2015). Indeed, the genome of this strain contains a single 70 kbp contig with almost complete representation of region 1 of the pAM1A3 plasmid, as well as an additional contig which consists of 12 kbp of pAM1A3 region 2 (Supplementary Figure S7). Due to the fragmented nature of the draft genome assembly of MIT1002 it is currently unclear whether this strain contains a plasmid or whether the relevant genes have been integrated into its

genome. *Alteromonas macleodii* strain MIT1002 originated from an enrichment culture of *Prochlorococcus* NATL2A collected in 1990 from the North Atlantic, but was physically isolated from it only 20 years later (Biller et al., 2015). Strain HOT1A3 was isolated from the Pacific Ocean near Hawaii in 2007 (Sher et al., 2011). Both of the strains originate from surface waters. Taken together, the GOS metagenomic results, the presence of *A. mediterranea* related genes in *A. macleodii* strains likely originating from two separate oceans and isolated almost two decades apart, and the presence of similar genes in *A. australica* and in a plasmid from *Glaciecola* sp., suggest a global distribution of genes involved in heavy metal resistance and hydrogenase activity among the *Alteromonadaceae*. It also suggests that plasmids such as pAM1A3 are likely vectors for genetic transfer of such information among distantly related *Alteromonas* species or even among *Alteromonadaceae* genera.

Maintaining extra-chromosomal amplicons, such as plasmids, incurs significant metabolic cost (Platt et al., 2012). What

advantage does *Alteromonas macleodii* HOT1A3 gain, if at all, by maintaining the pAM1A3 plasmid? Similar to other plasmids (Cytryn et al., 2008; Kalhoefer et al., 2011), pAM1A3 encodes many metal resistance genes, and these may be responsible for its increased resistance to Zn, Hg, and Cu compared to other *A. macleodii* strains (Figure 4). The oceanographic profiles of many heavy metals are similar to those of nutrients such as nitrate and phosphate, with maxima in deep waters (Bruland and Lohan, 2003; Morel et al., 2003). The concentrations of Zn, Hg, and Cu in deep waters are several orders of magnitude below the MIC we measured (up to 10 nM for Zn, e.g., in the North Pacific, Bruland et al., 2014, compared to a MIC of 200 nM, Figure 4), and most of these metal ions are complexed to organic compounds and not directly bioavailable. Nevertheless, even very low levels of heavy metals (140-fold lower than the MIC) have been shown to sufficiently affect bacterial physiology to offset the maintenance cost of a plasmid with a similar size to that of pAM1A3 (Gullberg et al., 2014). Additionally, heavy metal concentrations may be significantly higher on particles such as dust or zooplankton (e.g., Wang, 2002), and as copiotrophic bacteria *Alteromonas macleodii* and *A. mediterranea* have both been suggested to colonize such particles (Acinas et al., 1999; García-Martínez et al., 2002; Ivars-Martinez et al., 2008). Alternatively, other genes on the plasmid may be providing some fitness advantage, such as the hydrogenase operon (Vargas et al., 2011; Weyman et al., 2011) or any of the other >150 genes on the plasmid, most of which have unknown functions. Finally, the plasmid may provide no fitness advantage at all, rather being maintained by the putative toxin-antitoxin *hipAB* pair. Notably, we have observed expression of multiple genes on the plasmid in RNA-seq experiments (Fadееv et al., manuscript in preparation), suggesting that at least some of the genes are functional.

Plasmids and other extra-chromosomal elements can be identified experimentally (e.g., using PFGE, Supplementary Figure S1, López-Pérez et al., 2012), however, these methods require specific equipment, and are often not performed without initial evidence of such elements. Such evidence was provided by the efforts we put into obtaining a high-quality, fully sequenced and closed genome. Nevertheless, many of the draft assemblies could also have been useful, for example to provide reference genomes for transcriptomic analyses. Furthermore, with the high coverage obtained using Illumina sequencing, the *de novo* assemblies contained almost all of the coding genes (Figure 3A). Thus, with the appropriate caveats, a comparison of the functional capacities of the different *Alteromonas* strains could have been performed using a draft genome. Finally, phylogenetic analysis of the draft genome might have indicated horizontal transfer of some genes, yet only the revelation of a plasmid during genome closing shows a mechanism for horizontal transfer and identifies the transferrable contents. The reference-based and hybrid assemblies, in contrast, lack a significant part of the genomes and thus both comparative genomic and transcriptomic assays will clearly be biased, unless additional measures are taken to close the genome, e.g., through the use of complimentary methods such as optical

mapping or gap-closing by PCR (e.g., López-Pérez et al., 2013).

## AUTHOR CONTRIBUTIONS

EF and DS designed research. EF, DA, and DS isolated the DNA, produced in-house sequencing libraries and performed PFGE. EF, FP, AV, and SH assembled the genome, EF and DS analyzed the data. All authors participated in writing the manuscript.

## FUNDING

This study was supported by BSF grant 2010183 and Marie Curie Career Reintegration Grant “Microbes-2-Model” (to DS).

## ACKNOWLEDGMENTS

We thank Mario López-Pérez and Francisco Rodriguez-Valera for kindly providing *Alteromonas* strains for the metal resistance assays, Gazalah Sabehi, Tzvia Gildor, and Oded Beja for help with the PFGE, and Eddie Cytryn for critical review of the manuscript. Strain HOT1A3 was isolated in the laboratory of Prof. Sallie W. Chisholm at MIT, and the Illumina GA sequencing was supported by grants to SWC from the Gordon and Betty Moore Foundation (Grant award letters #495 and #495.01).

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <http://journal.frontiersin.org/article/10.3389/fmicb.2016.00248>

**FIGURE S1 | Validation of the presence of plasmid pAM1A3 using PFGE.** (A) Schematic illustration of the plasmid ring with the restriction digest sites; the table within the ring refers to the expected band sizes using each one of the enzymes. Two regions with higher sequencing coverage, which might be duplicated (see below), are shown. (B) PFGE gel of the plasmid isolated using NucleoBond BAC 100 kit, either uncut or after digesting with the three restriction enzymes shown in (A). The ladder markers refer to a low range ladder PFGE marker, a lambda ladder PFGE and a 1 KB DNA ladder, with the sizes of the markers on the right of the gel image. (C) Inset showing the doubled bands observed when pAM1A3 was digested by two of the restriction enzymes. (D) Read coverage of the plasmid (from the Illumina HiSeq2000 library, see Table 1). “Hot spots” for single nucleotide polymorphisms (SNPs) are highlighted. The numbers below marked regions refer to the number of SNP’s in that region.

**FIGURE S2 | Phylogenetic analysis of *Alteromonas* strains.** Maximum Likelihood tree of all *Alteromonas* strains for which a whole genome sequence is available. The numbers on the branches represent bootstraps percentage. (A) According to 16S rDNA sequences. (B) According to concatenated alignment of the following housekeeping genes: adenylsulfate kinase, DNA adenine methylase, *dnaK*, *flgG*, *gyrB*, lipoyl synthase, malate dehydrogenase, octanoyltransferase, phosphoglucosyltransferase, porphobilinogen deaminase, sulfate adenyltransferase subunit 2.

**FIGURE S3 | Abundance of the 100 most abundant functional SEED subsystems representation in different assembly methods compared to the final genome and plasmid.** Asterisk marked columns showed significant difference ( $p = 0.05$ , Fishers Exact Test with Benjamini-Hochberg False Discovery Rate).

**FIGURE S4 | Sequence comparison using BLAST ring image generator (Alikhan et al., 2011), of pAM1A3 to four isolated *A. macleodii* plasmids and pGLAAG01 isolated from *Glaciecola* sp. 4H-3-7 + YE-5 (outer rings, colored according to the legend) to pAM1A3 (inner black thin ring). The regions in the figure represent at least 70% identity and an *e*-value  $\leq$  0.00001.**

**FIGURE S5 | Pairwise alignment of region 1 (see Figure 2) of metal resistance fGI and pAM1A3.** The large gap in the alignment refers to a mobile element which is present in the fGI but absent in the pAM1A3 sequence. The alignment was conducted using Geneious aligner with default parameters.

**FIGURE S6 | Phylogeny of selected genes on the pAM1A3 plasmid.** The gray rectangles refer to genes placed within the metal resistance fGI. The phylogenetic trees were built using the Neighbor-Joining method, with bootstrap representation of >80% in all of the branches in all of the trees. The numbers in round brackets refer to the position of the gene in the plasmid sequence.

**FIGURE S7 | BlastN results of draft *A. macleodii* MIT1002 genome search against the sequence of pAM1A3 plasmid (*e*-value  $\leq$  10<sup>-6</sup>).** Regions 1 and 2 refer to the mutual regions of the pAM1A3 plasmid and metal resistance fGI (see Figure 2).

## REFERENCES

- Acinas, S. G., Antón, J., and Rodríguez-Valera, F. (1999). Diversity of free-living and attached bacteria in offshore western Mediterranean waters as depicted by analysis of genes encoding 16S rRNA. *Appl. Environ. Microbiol.* 65, 514–522.
- Alekshun, M. N., Levy, S. B., Mealy, T. R., Seaton, B. A., and Head, J. F. (2001). The crystal structure of MarR, a regulator of multiple antibiotic resistance, at 2.3 Å resolution. *Nat. Struct. Biol.* 8, 710–714. doi: 10.1038/90429
- Alikhan, N.-F., Petty, N. K., Ben Zakour, N. L., and Beatson, S. A. (2011). BLAST ring image generator (BRIG): simple prokaryote genome comparisons. *BMC Genomics* 12:402. doi: 10.1186/1471-2164-12-402
- Avrani, S., Wurtzel, O., Sharon, I., Sorek, R., and Lindell, D. (2011). Genomic island variability facilitates *Prochlorococcus*-virus coexistence. *Nature* 474, 604–608. doi: 10.1038/nature10172
- Aziz, R. K., Bartels, D., Best, A. A., DeJongh, M., Disz, T., Edwards, R. A., et al. (2008). The RAST server: rapid annotations using subsystems technology. *BMC Genomics* 9:75. doi: 10.1186/1471-2164-9-75
- Baker, B. J., Sheik, C. S., Taylor, C. A., Jain, S., Bhasi, A., Cavalcoli, J. D., et al. (2013). Community transcriptomic assembly reveals microbes that contribute to deep-sea carbon and nitrogen cycling. *ISME J.* 7, 1962–1973. doi: 10.1038/ismej.2013.85
- Barth, M. C., Dederich, D. A., and Dedon, P. C. (2009). An improved method for large-scale preparation of negatively and positively supercoiled plasmid DNA. *Biotechniques* 47, 633–635. doi: 10.2144/000113172
- Baumann, L., Baumann, P., Mandel, M., and Allen, R. D. (1972). Taxonomy of aerobic marine eubacteria. *J. Bacteriol.* 110, 402–429.
- Bignell, C., and Thomas, C. M. (2001). The bacterial ParA-ParB partitioning proteins. *J. Biotechnol.* 91, 1–34. doi: 10.1016/S0168-1656(01)00293-0
- Billier, S. J., Coe, A., Martin-Cuadrado, A.-B., and Chisholm, S. W. (2015). Draft Genome sequence of *Alteromonas macleodii* strain MIT1002, isolated from an enrichment culture of the marine cyanobacterium *Prochlorococcus*. *Genome Announc.* 3, e967–15. doi: 10.1128/genomeA.00967-15
- Blaas, L., Musteanu, M., Grabner, B., Eferl, R., Bauer, A., and Casanova, E. (2012). The use of bacterial artificial chromosomes for recombinant protein production in mammalian cell lines. *Methods Mol. Biol.* 824, 581–593. doi: 10.1007/978-1-61779-433-9\_31
- Boetzer, M., Henkel, C. V., Jansen, H. J., Butler, D., and Pirovano, W. (2011). Scaffolding pre-assembled contigs using SSPACE. *Bioinformatics* 27, 578–579. doi: 10.1093/bioinformatics/btq683
- Boetzer, M., and Pirovano, W. (2012). Toward almost closed genomes with GapFiller. *Genome Biol.* 13:R56. doi: 10.1186/gb-2012-13-6-r56
- Bouzat, J. L., and Hoostal, M. J. (2013). Evolutionary analysis and lateral gene transfer of two-component regulatory systems associated with heavy-metal tolerance in bacteria. *J. Mol. Evol.* 76, 267–279. doi: 10.1007/s00239-013-9558-z
- Breitbart, M. (2012). Marine viruses: truth or dare. *Ann. Rev. Mar. Sci.* 4, 425–448. doi: 10.1146/annurev-marine-120709-142805
- Brettin, T., Davis, J. J., Disz, T., Edwards, R. A., Gerdes, S., Olsen, G. J., et al. (2015). RASTtk: a modular and extensible implementation of the RAST algorithm for building custom annotation pipelines and annotating batches of genomes. *Sci. Rep.* 5:8365. doi: 10.1038/srep08365
- Bruland, K. W., and Lohan, M. C. (2003). “Controls of trace metals in seawater,” in *Treatise on Geochemistry*, eds H. D. Holland and K. K. Turekian (Oxford: Elsevier), 23–47. doi: 10.1016/B0-08-043751-6/06105-3
- Bruland, K. W., Middag, R., and Lohan, M. C. (2014). “Controls of trace metals in seawater,” in *Treatise on Geochemistry*, 2nd Edn, eds H. D. Holland and K. K. Turekian (Oxford: Elsevier), 19–51. doi: 10.1016/B978-0-08-095975-7.00602-1
- Chevreux, B., Pfisterer, T., Drescher, B., Driesel, A. J., Müller, W. E. G., Wetter, T., et al. (2004). Using the miraEST assembler for reliable and automated mRNA transcript assembly and SNP detection in sequenced ESTs. *Genome Res.* 14, 1147–1159. doi: 10.1101/gr.1917404
- Chin, C.-S., Alexander, D. H., Marks, P., Klammer, A. A., Drake, J., Heiner, C., et al. (2013). Nonhybrid, finished microbial genome assemblies from long-read SMRT sequencing data. *Nat. Methods* 10, 563–569. doi: 10.1038/nmeth.2474
- Cytryn, E. J., Jitackorn, S., Giraud, E., and Sadowsky, M. J. (2008). Insights learned from pBTAi1, a 229-kb accessory plasmid from *Bradyrhizobium* sp. strain BTAi1 and prevalence of accessory plasmids in other *Bradyrhizobium* sp. strains. *ISME J.* 2, 158–170. doi: 10.1038/ismej.2007.105
- Dahlberg, C., Bergström, M., and Hermansson, M. (1998). In situ detection of high levels of horizontal plasmid transfer in marine bacterial communities. *Appl. Environ. Microbiol.* 64, 2670–2675.
- Darling, A. E., Tritt, A., Eisen, J. A., and Facciotti, M. T. (2011). Mauve assembly metrics. *Bioinformatics* 27, 2756–2757. doi: 10.1093/bioinformatics/btr451
- Eklblom, R., and Wolf, J. B. W. (2014). A field guide to whole-genome sequencing, assembly and annotation. *Evol. Appl.* 7, 1026–1042. doi: 10.1111/eva.12178
- English, A. C., Richards, S., Han, Y., Wang, M., Vee, V., Qu, J., et al. (2012). Mind the gap: upgrading genomes with Pacific Biosciences RS long-read sequencing technology. *PLoS ONE* 7:e47768. doi: 10.1371/journal.pone.0047768
- Fernández-Gómez, B., Fernández-Guerra, A., Casamayor, E. O., González, J. M., Pedrós-Alió, C., and Acinas, S. G. (2012). Patterns and architecture of genomic islands in marine bacteria. *BMC Genomics* 13:347. doi: 10.1186/1471-2164-13-347
- Frank, O., Michael, V., Päuker, O., Boedeker, C., Jogler, C., Rohde, M., et al. (2015). Plasmid curing and the loss of grip – The 65-kb replicon of *Phaobacter inhibens* DSM 17395 is required for biofilm formation, motility and the colonization of marine algae. *Syst. Appl. Microbiol.* 38, 120–127. doi: 10.1016/j.syapm.2014.12.001
- García-Martínez, J., Acinas, S. G., Massana, R., and Rodríguez-Valera, F. (2002). Prevalence and microdiversity of *Alteromonas macleodii*-like microorganisms in different oceanic regions. *Environ. Microbiol.* 4, 42–50. doi: 10.1046/j.1462-2920.2002.00255.x
- Gonzaga, A., Martin-Cuadrado, A.-B. B., López-Pérez, M., Mizuno, C. M., García-Heredia, I., Kimes, N. E., et al. (2012). Polyclonality of concurrent natural populations of *Alteromonas macleodii*. *Genome Biol. Evol.* 4, 1360–1374. doi: 10.1093/gbe/evs112
- Grass, G., and Rensing, C. (2001). Genes involved in copper homeostasis in *Escherichia coli* genes involved in copper homeostasis in *Escherichia coli*. *J. Bacteriol.* 183, 2145–2147. doi: 10.1128/JB.183.6.2145
- Gullberg, E., Albrecht, L. M., Karlsson, C., Sandegren, L., and Andersson, D. I. (2014). Selection of a multidrug resistance plasmid by sublethal levels of antibiotics and heavy metals. *MBio* 5, 19–23. doi: 10.1128/mBio.01918-14.Editor
- Gurevich, A., Saveliev, V., Vyahhi, N., and Tesler, G. (2013). QUASt: quality assessment tool for genome assemblies. *Bioinformatics* 29, 1072–1075. doi: 10.1093/bioinformatics/btt086
- Han, C., Shiga, Y., Tobe, T., and Sasakawa, C. (2001). Structural and functional characterization of IS 679 and IS 66 -family elements. *J. Bacteriol.* 183, 4296–4304. doi: 10.1128/JB.183.14.4296

- Hightower, R. C., Metge, D. W., and Santi, D. V. (1987). Plasmid migration using orthogonal-field-alternation gel electrophoresis. *Nucleic Acids Res.* 15, 8387–8398. doi: 10.1093/nar/15.20.8387
- Huang, C. C., Narita, M., Yamagata, T., and Endo, G. (1999). Identification of three merB genes and characterization of a broad-spectrum mercury resistance module encoded by a class II transposon of *Bacillus megaterium* strain MB1. *Gene* 239, 361–366. doi: 10.1016/S0378-1119(99)00184-5
- Hunt, M., Kikuchi, T., Sanders, M., Newbold, C., Berriman, M., and Otto, T. D. (2013). REAPR: a universal tool for genome assembly evaluation. *Genome Biol.* 14:R47. doi: 10.1186/gb-2013-14-5-r47
- Hunt, M., Newbold, C., Berriman, M., and Otto, T. D. (2014). A comprehensive evaluation of assembly scaffolding tools. *Genome Biol.* 15:R42. doi: 10.1186/gb-2014-15-3-r42
- Ivanova, E. P., Flavier, S., and Christen, R. (2004). Phylogenetic relationships among marine *Alteromonas*-like *Proteobacteria*: emended description of the family Alteromonadaceae and proposal of Pseudoalteromonadaceae fam. nov., Colwelliaceae fam. nov., Shewanellaceae fam. nov., Moritellaceae fam. nov., Ferri. *Int. J. Syst. Evol. Microbiol.* 54, 1773–1788. doi: 10.1099/ij.s.0.02997-0
- Ivanova, E. P., López-Pérez, M., Zabalos, M., Nguyen, S. H., Webb, H. K., Ryan, J., et al. (2015). Ecophysiological diversity of a novel member of the genus *Alteromonas*, and description of *Alteromonas mediterranea* sp. nov. *Antonie Van Leeuwenhoek* 107, 119–132. doi: 10.1007/s10482-014-0309-y
- Ivars-Martinez, E., Martin-Cuadrado, A.-B., D'Auria, G., Mira, A., Ferrera, S., Johnson, J., et al. (2008). Comparative genomics of two ecotypes of the marine planktonic copiotroph *Alteromonas macleodii* suggests alternative lifestyles associated with different kinds of particulate organic matter. *ISME J.* 2, 1194–1212. doi: 10.1038/ismej.2008.74
- Jones, P., Binns, D., Chang, H.-Y., Fraser, M., Li, W., McAnulla, C., et al. (2014). InterProScan 5: genome-scale protein function classification. *Bioinformatics* 30, 1236–1240. doi: 10.1093/bioinformatics/btu031
- Juhas, M., van der Meer, J. R., Gaillard, M., Harding, R. M., Hood, D. W., and Crook, D. W. (2009). Genomic islands: tools of bacterial horizontal gene transfer and evolution. *FEMS Microbiol. Rev.* 33, 376–393. doi: 10.1111/j.1574-6976.2008.00136.x
- Kalhoefer, D., Thole, S., Voget, S., Lehmann, R., Liesegang, H., Wollher, A., et al. (2011). Comparative genome analysis and genome-guided physiological analysis of *Roseobacter litoralis*. *BMC Genomics* 12:324. doi: 10.1186/1471-2164-12-324
- Kearse, M., Moir, R., Wilson, A., Stones-Havas, S., Cheung, M., Sturrock, S., et al. (2012). Geneious basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics* 28, 1647–1649. doi: 10.1093/bioinformatics/bts199
- Kisand, V., and Lettieri, T. (2013). Genome sequencing of bacteria: sequencing, de novo assembly and rapid analysis using open source tools. *BMC Genomics* 14:211. doi: 10.1186/1471-2164-14-211
- Klippel, B., Lochner, A., Bruce, D. C., Walston Davenport, K., Detter, C., Goodwin, L. A., et al. (2011). Complete genome sequence of the marine cellulose- and Xylan-degrading bacterium *Glaciecola* sp. Strain 4H-3-7+YE-5. *J. Bacteriol.* 193, 4547–4548. doi: 10.1128/JB.05468-11
- Klockgether, J., Reva, O., Larbig, K., and Tümmler, B. (2004). Sequence analysis of the mobile genome island pKLC102 of *Pseudomonas aeruginosa* C. *J. Bacteriol.* 186, 518–534. doi: 10.1128/JB.186.2.518
- Koren, S., Harhay, G. P., Smith, T. P., Bono, J. L., Harhay, D. M., Mcvey, S. D., et al. (2013). Reducing assembly complexity of microbial genomes with single-molecule sequencing. *Genome Biol.* 14:R101. doi: 10.1186/gb-2013-14-9-r101
- Krzywinski, M., Schein, J., Birol, I., Connors, J., Gascoyne, R., Horsman, D., et al. (2009). Circo: an information aesthetic for comparative genomics. *Genome Res.* 19, 1639–1645. doi: 10.1101/gr.092759.109
- Lee, C.-T., Chen, I.-T., Yang, Y.-T., Ko, T.-P., Huang, Y.-T., Huang, J.-Y., et al. (2015). The opportunistic marine pathogen *Vibrio parahaemolyticus* becomes virulent by acquiring a plasmid that expresses a deadly toxin. *Proc. Natl. Acad. Sci. U.S.A.* 112, 10798–10803. doi: 10.1073/pnas.1503129112
- Liebert, C. A., Hall, R. M., and Summers, A. O. (1999). Transposon Tn21, flagship of the floating genome. *Microbiol. Mol. Biol. Rev.* 63, 507–522.
- Loman, N. J., Constantinidou, C., Chan, J. Z. M., Halachev, M., Sergeant, M., Penn, C. W., et al. (2012). High-throughput bacterial genome sequencing: an embarrassment of choice, a world of opportunity. *Nat. Rev. Microbiol.* 10, 599–606. doi: 10.1038/nrmicro2850
- López-López, A., Bartual, S. G., Stal, L., Onyshchenko, O., and Rodríguez-Valera, F. (2005). Genetic analysis of housekeeping genes reveals a deep-sea ecotype of *Alteromonas macleodii* in the Mediterranean Sea. *Environ. Microbiol.* 7, 649–659. doi: 10.1111/j.1462-2920.2004.00733.x
- López-Pérez, M., Gonzaga, A., Ivanova, E. P., and Rodríguez-Valera, F. (2014). Genomes of *Alteromonas australica*, a world apart. *BMC Genomics* 15:483. doi: 10.1186/1471-2164-15-483
- López-Pérez, M., Gonzaga, A., Martin-Cuadrado, A.-B., Onyshchenko, O., Ghavidel, A., Ghai, R., et al. (2012). Genomes of surface isolates of *Alteromonas macleodii*: the life of a widespread marine opportunistic copiotroph. *Sci. Rep.* 2, 1–11. doi: 10.1038/srep00696
- López-Pérez, M., Gonzaga, A., and Rodríguez-Valera, F. (2013). Genomic diversity of “deep ecotype” *Alteromonas macleodii* isolates: evidence for pan-mediterranean clonal frames. *Genome Biol. Evol.* 5, 1220–1232. doi: 10.1093/gbe/evt089
- Macrander, J., Broe, M., and Daly, M. (2015). Multi-copy venom genes hidden in de novo transcriptome assemblies, a cautionary tale with the snakelocks sea anemone *Anemonia sulcata* (Pennant, 1777). *Toxicon* 108, 184–188. doi: 10.1016/j.toxicon.2015.09.038
- Magoc, T., Pabinger, S., Canzar, S., Liu, X., Su, Q., Puiu, D., et al. (2013). GAGE-B: an evaluation of genome assemblers for bacterial organisms. *Bioinformatics* 29, 1718–1725. doi: 10.1093/bioinformatics/btt273
- Marasini, D., and Fakhr, M. K. (2014). Exploring PFGE for detecting large plasmids in *Campylobacter jejuni* and *Campylobacter coli* isolated from various retail meats. *Pathogens* 3, 833–844. doi: 10.3390/pathogens3040833
- Mathew, M. K., Smith, C. L., and Cantor, C. R. (1988). High-resolution separation and accurate size determination in pulsed-field gel electrophoresis of DNA. 2. Effect of pulse time and electric field strength and implications for models of the separation process. *Biochemistry* 27, 9210–9216. doi: 10.1021/bi00426a022
- Mayali, X., and Azam, F. (2004). Algicidal bacteria in the sea and their impact on algal blooms. *J. Eukaryot. Microbiol.* 51, 139–144. doi: 10.1111/j.1550-7408.2004.tb00538.x
- McCarren, J., Becker, J. W., Repeta, D. J., Shi, Y., Young, C. R., Malmstrom, R. R., et al. (2010). Microbial community transcriptomes reveal microbes and metabolic pathways associated with dissolved organic matter turnover in the sea. *Proc. Natl. Acad. Sci. U.S.A.* 107, 16420–16427. doi: 10.1073/pnas.1010732107
- Metzker, M. L. (2010). Sequencing technologies - the next generation. *Nat. Rev. Genet.* 11, 31–46. doi: 10.1038/nrg2626
- Moore, L. R., Coe, A., Zinser, E. R., Saito, M. A., Sullivan, M. B., Lindell, D., et al. (2007). Culturing the marine cyanobacterium *Prochlorococcus*. *Limnol. Oceanogr. Methods* 5, 353–362. doi: 10.4319/lom.2007.5.353
- Morales-Espinosa, R., Soberón-Chávez, G., Delgado-Sapién, G., Sandner-Miranda, L., Méndez, J. L., González-Valencia, G., et al. (2012). Genetic and phenotypic characterization of a *Pseudomonas aeruginosa* population with high frequency of genomic islands. *PLoS ONE* 7:e37459. doi: 10.1371/journal.pone.0037459
- Morel, F. M. M., Milligan, A. J., and Saito, M. A. (2003). “Marine bioinorganic chemistry: the role of trace metals in the oceanic cycles of major nutrients,” in *Treatise on Geochemistry*, eds H. D. Holland and K. K. Turekian (Amsterdam: Elsevier), 113–143.
- Moreno-Hagselsieb, G., and Latimer, K. (2008). Choosing BLAST options for better detection of orthologs as reciprocal best hits. *Bioinformatics* 24, 319–324. doi: 10.1093/bioinformatics/btm585
- Mutreja, A., Kim, D. W., Thomson, N. R., Connor, T. R., Lee, J. H., Kariuki, S., et al. (2011). Evidence for several waves of global transmission in the seventh cholera pandemic. *Nature* 477, 462–465. doi: 10.1038/nature10392
- Nelson, C. E., and Wear, E. K. (2014). Microbial diversity and the lability of dissolved organic carbon. *Proc. Natl. Acad. Sci. U.S.A.* 111, 7166–7167. doi: 10.1073/pnas.1405751111
- Neumann, B., Pospiech, A., and Schairer, H. U. (1992). Rapid isolation of genomic DNA from gram-negative bacteria. *Trends Genet.* 8, 332–333. doi: 10.1016/0168-9525(92)90146-U
- Nikaido, H. (1998). Antibiotic resistance caused by gram-negative multidrug efflux pumps. *Clin. Infect. Dis.* 27(Suppl. 1), S32–S41. doi: 10.1086/514920
- Osoiro, C. R., Rivas, A. J., Balado, M., Fuentes-Monteverde, J. C., Rodríguez, J., Jiménez, C., et al. (2015). A transmissible plasmid-borne pathogenicity

- island confers piscibactin biosynthesis in the fish pathogen *Photobacterium damsela* subsp. piscicida. *Appl. Environ. Microbiol.* 81, 5867–5879. doi: 10.1128/AEM.01580-15
- Overbeek, R., Begley, T., Butler, R. M., Choudhuri, J. V., Chuang, H.-Y., Cohoon, M., et al. (2005). The subsystems approach to genome annotation and its use in the project to annotate 1000 genomes. *Nucleic Acids Res.* 33, 5691–5702. doi: 10.1093/nar/gki866
- Ozdemir, K., and Acar, S. (2014). Plasmid profile and pulsed-field gel electrophoresis analysis of *Salmonella enterica* isolates from humans in Turkey. *PLoS ONE* 9:e95976. doi: 10.1371/journal.pone.0095976
- Pal, C., Bengtsson-Palme, J., Rensing, C., Kristiansson, E., and Larsson, D. G. J. (2014). BacMet: antibacterial biocide and metal resistance genes database. *Nucleic Acids Res.* 42, 737–743. doi: 10.1093/nar/gkt1252
- Parks, A. R., and Peters, J. E. (2007). Transposon Tn7 is widespread in diverse bacteria and forms genomic islands. *J. Bacteriol.* 189, 2170–2173. doi: 10.1128/JB.01536-06
- Pedler, B. E., Aluwihare, L. I., and Azam, F. (2014). Single bacterial strain capable of significant contribution to carbon cycling in the surface ocean. *Proc. Natl. Acad. Sci. U.S.A.* 111, 7202–7207. doi: 10.1073/pnas.1401887111
- Platt, T. G., Bever, J. D., and Fuqua, C. (2012). A cooperative virulence plasmid imposes a high fitness cost under conditions that induce pathogenesis. *Proc. Biol. Sci.* 279, 1691–1699. doi: 10.1098/rspb.2011.2002
- Powers, J. G., Weigman, V. J., Shu, J., Pufky, J. M., Cox, D., and Hurban, P. (2013). Efficient and accurate whole genome assembly and methylome profiling of *E. coli*. *BMC Genomics* 14:675. doi: 10.1186/1471-2164-14-675
- Quail, M., Smith, M. E., Coupland, P., Otto, T. D., Harris, S. R., Connor, T. R., et al. (2012). A tale of three next generation sequencing platforms: comparison of Ion torrent, pacific biosciences and illumina MiSeq sequencers. *BMC Genomics* 13:341. doi: 10.1186/1471-2164-13-341
- Rissman, A. I., Mau, B., Biehl, B. S., Darling, A. E., Glasner, J. D., and Perna, N. T. (2009). Reordering contigs of draft genomes using the mauve aligner. *Bioinformatics* 25, 2071–2073. doi: 10.1093/bioinformatics/btp356
- Rodriguez-Valera, F., Martin-Cuadrado, A.-B., Rodriguez-Brito, B., Pasić, L., Thingstad, T. F., Rohwer, F., et al. (2009). Explaining microbial population genomics through phage predation. *Nat. Rev. Microbiol.* 7, 828–836. doi: 10.1038/nrmicro2235
- Rohwer, F., and Thurber, R. V. (2009). Viruses manipulate the marine environment. *Nature* 459, 207–212. doi: 10.1038/nature08060
- Seemann, T. (2014). Prokka: rapid prokaryotic genome annotation. *Bioinformatics* 30, 2068–2069. doi: 10.1093/bioinformatics/btu153
- Sheik, A. R., Brussaard, C. P. D., Lavik, G., Lam, P., Musat, N., Krupke, A., et al. (2014). Responses of the coastal bacterial community to viral infection of the algae *Phaeocystis globosa*. *ISME J.* 8, 212–225. doi: 10.1038/ismej.2013.135
- Sher, D., Thompson, J. W., Kashtan, N., Croal, L., and Chisholm, S. W. (2011). Response of *Prochlorococcus* ecotypes to co-culture with diverse marine bacteria. *ISME J.* 5, 1125–1132. doi: 10.1038/ismej.2011.1
- Shi, Y., McCarren, J., and Delong, E. F. (2012). Transcriptional responses of surface water marine microbial assemblages to deep-sea water amendment. *Environ. Microbiol.* 14, 191–206. doi: 10.1111/j.1462-2920.2011.02598.x
- Silver, S., and Walderhaug, M. (1992). Gene regulation of plasmid- and chromosome-determined inorganic ion transport in bacteria. *Microbiol. Rev.* 56, 195–228.
- Sizemore, R. K., and Colwell, R. R. (1977). Plasmids carried by antibiotic-resistant marine bacteria. *Antimicrob. Agents Chemother.* 12, 373–382. doi: 10.1128/AAC.12.3.373
- Summers, A. O. (1992). Untwist and shout: a heavy metal-responsive transcriptional regulator. *J. Bacteriol.* 174, 3097–3101.
- Tamura, K., Stecher, G., Peterson, D., Filipski, A., and Kumar, S. (2013). MEGA6: molecular evolutionary genetics analysis version 6.0. *Mol. Biol. Evol.* 30, 2725–2729. doi: 10.1093/molbev/mst197
- Tock, M. R., and Dryden, D. T. F. (2005). The biology of restriction and anti-restriction. *Curr. Opin. Microbiol.* 8, 466–472. doi: 10.1016/j.mib.2005.06.003
- Ullrich, S., Ullrich, S., Kube, M., Kube, M., Schu, S., Schu, S., et al. (2005). A hypervariable 130-kilobase genomic region of *Microbiology* 187, 7176–7184. doi: 10.1128/JB.187.21.7176
- Vargas, W. A., Weyman, P. D., Tong, Y., Smith, H. O., and Xu, Q. (2011). [NiFe] hydrogenase from *Alteromonas macleodii* with unusual stability in the presence of oxygen and high temperature. *Appl. Environ. Microbiol.* 77, 1990–1998. doi: 10.1128/AEM.01559-10
- Wang, W. X. (2002). Interactions of trace metals and different marine food chains. *Mar. Ecol. Prog. Ser.* 243, 295–309. doi: 10.3354/meps243295
- Wen, Y., Behiels, E., Felix, J., Elegheert, J., Vergauwen, B., Devreese, B., et al. (2014). The bacterial antitoxin HipB establishes a ternary complex with operator DNA and phosphorylated toxin HipA to regulate bacterial persistence. *Nucleic Acids Res.* 42, 10134–10147. doi: 10.1093/nar/gku665
- Weyman, P. D., Smith, H. O., and Xu, Q. (2011). Genetic analysis of the *Alteromonas macleodii* [NiFe]-hydrogenase. *FEMS Microbiol. Lett.* 322, 180–187. doi: 10.1111/j.1574-6968.2011.02348.x
- Yamaguchi, Y., Park, J.-H., and Inouye, M. (2011). Toxin-antitoxin systems in bacteria and archaea. *Annu. Rev. Genet.* 45, 61–79. doi: 10.1146/annurev-genet-110410-132412
- Zerbino, D. R., and Birney, E. (2008). Velvet: algorithms for de novo short read assembly using de Bruijn graphs. *Genome Res.* 18, 821–829. doi: 10.1101/gr.074492.107

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2016 Fadееv, De Pascale, Vezzi, Hübner, Aharonovich and Sher. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.