



# Discovery of mosaic genomic islands in *Pseudomonas* spp.

Mehul Jani<sup>1</sup> · Rajeev K. Azad<sup>1,2</sup>

Received: 18 October 2020 / Revised: 4 January 2021 / Accepted: 17 February 2021 / Published online: 1 March 2021  
© The Author(s), under exclusive licence to Springer-Verlag GmbH, DE part of Springer Nature 2021

## Abstract

Genomic islands, defined as large clusters of genes mobilized through horizontal gene transfer, have a profound impact on evolution of prokaryotes. Recently, we developed a new program, IslandCafe, for identifying such large localized structures in bacterial genomes. A unique attribute of IslandCafe is its ability to decipher mosaic structures within genomic islands. Mosaic genomic islands have generated immense interest due to novel traits that have been attributed to such islands. To provide the *Pseudomonas* research community a catalogue of mosaic islands in *Pseudomonas* spp., we applied IslandCafe to decipher genomic islands in 224 completely sequenced genomes of *Pseudomonas* spp. We also performed comparative genomic analysis using BLAST to infer potential sources of distinct segments within genomic islands. Of the total 4271 genomic islands identified in *Pseudomonas* spp., 1036 were found to be mosaic. We also identified drug-resistant and pathogenic genomic islands and their potential donors. Our analysis provides a useful resource for *Pseudomonas* research community to further examine and interrogate mosaic islands in the genomes of interest and understand their role in the emergence and evolution of novel traits.

**Keywords** Genomic islands · *Pseudomonas* spp. · Genome segmentation · Clustering · Comparative genomics

## Introduction

Clusters of functionally related genes mobilized across distinct lineages through horizontal gene transfer (HGT) (Koonin et al. 2001) are referred to as genomic islands (GIs) and are known to have a profound influence on the evolution of prokaryotes (Juhás et al. 2009; van der Meer et al. 2003). Of immense importance is to understand how by acquiring GIs, bacteria gain versatile novel traits, including the capability to degrade novel metabolites, become antibiotic resistance, or become pathogenic. Identification and characterization of GIs are central to the long-term goal of inferring and understanding the factors that modulate bacterial genome evolution.

A GI may arise because of a single insertion of a large genetic element with several contiguous genes from a donor organism. Alternatively, it may represent multiple independently acquired genetic elements at a genomic locus (Osborn and Boltner 2002). The latter, thus, represents different genomic contexts, yet it is characterized as a single “mosaic” GI due to the physical association of the acquired DNAs from different sources within a recipient genome. Even a single event of HGT can transform non-virulent bacteria to virulent or drug-susceptible bacteria to drug resistant (Noto et al. 2008).

Mosaic GIs may thus arise due to several independent insertion events, recombination events, and transposition, reflecting different genomic contexts (Mathee et al. 2008; Qiu et al. 2006). A mosaic GI has DNA elements from multiple sources. The regions with different origins within a mosaic GI may have differential contributions to the complex function that it may confer. For example, the Hrp PAI in *P. syringae* has a tripartite mosaic structure with a cluster of type III secretion genes bounded by effector genes that cooperate to provide parasitic fitness and virulence function (Alfano et al. 2000). On the other hand, a virulence determinant localized in a certain region in a mosaic GI may have been acquired in a single transfer

---

Communicated by Erko Stackebrandt.

✉ Rajeev K. Azad  
Rajeev.Azad@unt.edu

<sup>1</sup> Department of Biological Sciences and BioDiscovery Institute, University of North Texas, Denton, TX, USA

<sup>2</sup> Department of Mathematics, University of North Texas, Denton, TX, USA

event, and the other regions represent additional HGT events. Genetic elements acquired from different sources within a mosaic island may also code for unrelated functions. The deconstruction of mosaic structure of GIs is thus a first step in understanding the complex functions imparted by their disparate components resulting in better fitness and adaptability (Jani et al. 2016).

Recently, we developed a new software, IslandCafe (Jani and Azad 2019), for more robust identification of GIs as well as for characterization of their mosaic organizational structures via compositional anomaly and feature enrichment assessment. IslandCafe compared favorably with other programs in identifying GIs in both simulated genomes and well-curated bacterial genomes (Jani and Azad 2019). At the core of this approach is the integration of marker enrichment and phyletic pattern analyses within a framework of recursive segmentation and agglomerative clustering. This enabled not only the screening of atypical non-island segments but also the identification of islands lacking markers by virtue of their association with islands with markers. Since this method first disassembles the disparate segments of apparently different ancestries via recursive segmentation and then reassembles segments of apparently same ancestry within the genome via agglomerative clustering, the compositional structures of GIs are revealed. Of great interest are the mosaic GIs that are revealed by this process—contiguous segments that show signatures of horizontal acquisition but each segment potentially representing a lineage different from those of the neighboring segments. This underlying segmentation and clustering algorithm used by IslandCafe has been used previously for identifying GIs in *P. aeruginosa* (Jani et al. 2016). It was able to identify verified islands in *P. aeruginosa* as well as delineate novel GIs (Jani et al. 2016). Importantly, it was also able to identify mosaic islands. Here, we applied IslandCafe to decipher mosaic GIs in 224 completely sequenced genomes of *Pseudomonas* spp., which are Gram-negative rod-shaped bacteria of genus *Pseudomonas* and family Pseudomonadaceae.

*Pseudomonas* spp. represents a diverse group of bacteria that dwell in many different environments and display high metabolic diversity and genome plasticity. These include human opportunistic pathogen *Pseudomonas aeruginosa* (Emerson et al. 2002), plant pathogen *Pseudomonas syringae* (Buell et al. 2003), plant growth promoting *Pseudomonas fluorescens* (Paulsen et al. 2005), and water and soil dwelling *Pseudomonas putida* (Nelson et al. 2002). Due to its nearly ubiquitous presence and diverse functions, this group of organisms has drawn immense interest and attention among microbiologists. Since the sequencing of first *Pseudomonas* genome in 2000 (the *P. aeruginosa* PAO1 strain) (Stover et al. 2000), genomes of hundreds of *Pseudomonads* have been completely sequenced, enabling comparative studies to understand factors or mechanisms driving

*Pseudomonas* spp. evolution and their contributions to the metabolic versatility of this important group.

While mosaic GIs have been studied previously in *Pseudomonas* (Jani et al. 2016; Mathee et al. 2008), this has been limited to a few individual strains. The *pha*-GI island of *Pseudomonas* sp. 14-3 was found to contain regions with dissimilar G+C content and was thus identified a mosaic island (Ayub et al. 2007). Likewise, in *P. syringae*, Hrp pathogenicity island was found to have tripartite structure (Alfano et al. 2000). pKLC102 of *Pseudomonas aeruginosa* C was identified as having phage and plasmid sequences (Klockgether et al. 2004). Similarly, PFGI-1 island of *Pseudomonas fluorescens* Pf-5 was found to have sequences of phage and plasmid origins (Mavrodi et al. 2009).

Contribution of mosaic GIs to the evolution of *Pseudomonas* genomes at species and genus level remained one of the yet unexplored areas in *Pseudomonas* genomics, though they are now recognized to be one of the potential factors underlying differential pathogenicity or other traits (Jani et al. 2016). GI prediction methods have rarely been used to identify the underlying mosaic structure of GIs. IslandCafe was also previously used to just identify GIs and not determine the disparate phylogenetic ancestries of the genomic islets comprising the GIs (Jani and Azad 2019). Here, we exploit the ability of IslandCafe (Jani and Azad 2019) in deciphering regions of distinct compositions within a GI to interrogate *Pseudomonas* genomes for the presence of compositionally composite GIs and offer the *Pseudomonas* community a catalogue of mosaic GIs for further investigation in the strains of their interest. In contrast to the previous studies on GI identification in *Pseudomonas* spp., we have considered here multiple species of *Pseudomonas*, with the focus on mosaic GIs, thus, helping understand the role of mosaic GIs in the evolution of *Pseudomonas* genomes.

## Materials and methods

### *Pseudomonas* genomes

The complete genome sequences of 224 *Pseudomonas* spp. were downloaded from Pseudomonas Genome Database (pseudomonas.com) (Winsor et al. 2016). These genomes represented six *Pseudomonas* species, *P. aeruginosa*, *P. fluorescens*, *P. putida*, *P. chlororaphis*, *P. syringae* and *P. stutzeri* with *Pseudomonas aeruginosa* being most abundant with 137 genomes. Complete list of genomes that were used in this study are provided in Table S1.

## Identifying mosaic GIs using IslandCafe

IslandCafe (Jani and Azad 2019) is based on a statistical framework that allows incorporation of biological and phylogenetic information to identify GIs. Clusters of similar genomic segments are first generated within a statistical hypothesis testing framework of segmentation and clustering at a stringent setting. This results in generation of pure clusters that harbor either vertically inherited (native) or horizontally acquired (alien) segments but not both. However, the generation of numerous small native clusters along with the large native cluster complicates the detection of GIs as GIs are identified as residents of smaller clusters whereas the native segments are identified as residents of the largest cluster. Attempts to coalesce native clusters into a single native cluster by relaxing the stringency result in undesirable cluster mergers. IslandCafe addresses this by performing GI specific feature enrichment analysis of the smaller clusters as well as the phyletic pattern analysis of the genes harbored by these clusters. Enrichment of a cluster in GI specific markers, such as genes associated with integration/recombination and transposition, renders all segments within the cluster deemed GIs or parts of GIs by IslandCafe. Segments of clusters lacking marker enrichment and containing genes that are well distributed in the close relatives are deemed native. This includes weakly typical native segments as well as the segments that show atypicality for reasons other than horizontal acquisition (e.g., highly expressing native genes such as ribosomal protein genes). Notably, this procedure, in contrast to other methods, allows identification of marker devoid or deficient GIs by their association with GIs enriched in markers in the “alien” clusters (i.e., via sharing of similar composition). Once the GIs are established through this procedure, a GI composed of contiguous atypical segments that differ in composition from their immediate neighboring segments (within the GI) is designated as a mosaic GI. Mosaic GIs thus have segments from multiple clusters that represent potentially different evolutionary origins.

## Comparative genomic analysis of mosaic GIs

Each of the distinct segments of a mosaic GI was queried against a non-redundant nucleotide database using BLASTn (Altschul et al. 1990) to identify the potential donor of the segment. Here, we particularly looked for best hits with unusually very high similarity (>90% identity and query coverage) in distant taxa. The potential donors of mosaic GI segments were thus identified.

## Clustering analysis to identify unique GIs

To identify GIs that are shared between different genomes (GIs having the same nucleotide sequence) we used a sequence-clustering tool, Linclust (Steinegger and Söding 2018). If the sequences of GIs present in different strains had 90% or more similarity, they were grouped together in a cluster.

## Functional annotation

Gene functional annotations were obtained from GenBank files downloaded from *Pseudomonas* Genome Database ([pseudomonas.com](http://pseudomonas.com)) (Winsor et al. 2016).

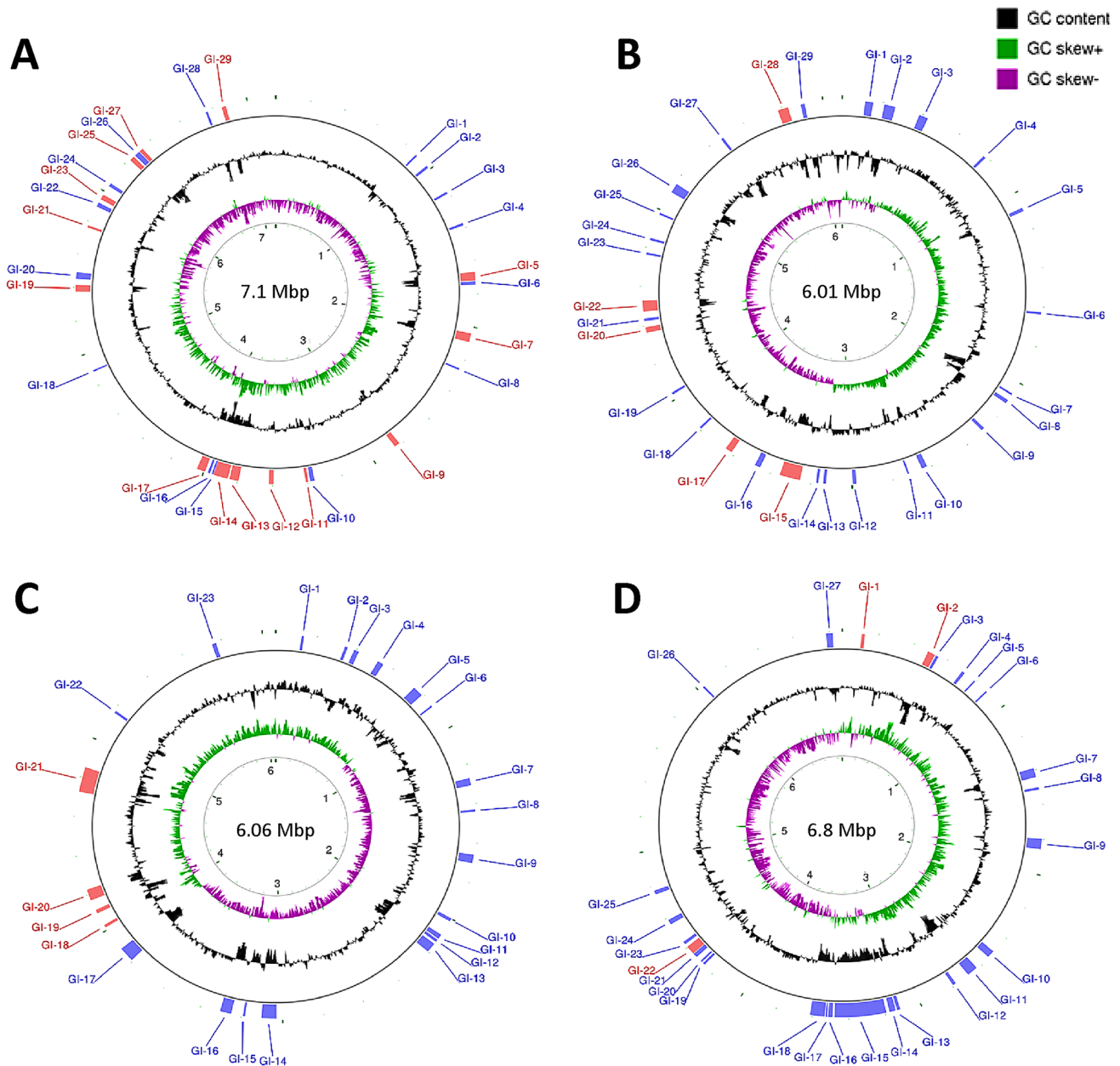
## Results and discussion

### Identification of mosaic islands

In application to 224 completely sequenced *Pseudomonas* spp. genomes, IslandCafe identified 4271 GIs, of which 1036 were found to be mosaic (Table 1). There were ~19 GIs (~5 mosaic) per genome on average, ranging in size from 8001 to 286,303 bp (8001–174,596 bp for mosaic) and with mean and median sizes 23,642 bp and 16,895 bp (34,292 bp and 26,263 bp for mosaic), respectively. Of the 224 genomes, 137 were those of *P. aeruginosa*. Circular maps of GIs in four representative *Pseudomonas* species, namely *P. aeruginosa*, *P. syringae*, *P. fluorescens*, and *P. putida* are shown in Fig. 1 (See Supplementary Figures

**Table 1** Distribution of GIs, GI clusters, and mosaic GIs in *Pseudomonas* spp

Species	Number of strains	Number of GIs	Number of GI clusters	Number of single GI clusters (unique GIs)	Number of mosaic GIs
<i>Pseudomonas aeruginosa</i>	137	2957	1725	1307	836
<i>Pseudomonas chlororaphis</i>	8	186	164	153	40
<i>Pseudomonas fluorescens</i>	18	349	334	319	57
<i>Pseudomonas putida</i>	24	401	374	349	62
<i>Pseudomonas stutzeri</i>	13	161	154	149	19
<i>Pseudomonas syringae</i>	24	217	182	161	22
Total	224	4271	2933	2438	1036



**Fig. 1** Genomic map of **a** *Pseudomonas aeruginosa* AR 0357, **b** *Pseudomonas syringae* pv. *syringae* B301D, **c** *Pseudomonas fluorescens* UK4, and **d** *Pseudomonas putida* H8234 showing GIs predicted by IslandCafe. Islands shown in red are mosaic GIs

and Tables for more details; coordinates of GIs and segments comprising mosaic GIs in respective genomes are provided in Tables S1 and S2).

Among *Pseudomonas* spp. considered here, GIs are most abundant in *Pseudomonas chlororaphis*, with ~ 3 GIs per Mbp. Mosaic GIs are most prevalent in *P. aeruginosa* with ~ 1 mosaic GI per Mbp. Our results show that acquisition of GIs in *Pseudomonas* spp. is frequent and mosaic GIs are widespread in the *Pseudomonas* genomes.

### Identification of unique GIs

To identify GIs that are unique to a strain or are present in multiple strains, we used Linclust, a sequence-clustering tool (Steinegger and Söding 2018). Based on sequence similarity, we identified 2933 clusters of GIs (Table 1). If the sequences of GIs present in different strains had 90% or more similarity, they were grouped together in a cluster. Expectedly, some GIs were present in multiple strains (Table S3). The

largest cluster (cluster #2639) is composed of GIs from 35 *P. aeruginosa* strains. The strains sharing a GI could all have acquired the shared GI from the same donor in separate transfer events or could have acquired via intraspecies dissemination following acquisition in a strain from the donor, or this could be a consequence of both, i.e., multiple interspecies and intraspecies transfers. We observed that at > 90% sequence similarity, all GIs within a cluster belong to a single species (Table S3). It is possible that *Pseudomonas* spp. may have shared GIs, however, because of different host evolutionary pressures, these GIs might have diverged.

### Gene content of mosaic islands

Functional characterization of mosaic GIs provided insights into their potential roles in adaptation. We scanned the annotation of the genes and identified genes often associated with GIs. Of the 2147 segments identified as composing

the mosaic GIs in *P. aeruginosa*, 774 harbored genes that are often associated with horizontal transfer (Table S4). These genes included those encoding transposase, integrase, recombinase, and integration host factor, as well as insertion element, phage and plasmid genes. We analyzed the gene content of the mosaic GI segments that were found to harbor genes with unusually high similarity in distant taxa by our BLAST analyses (discussed further in the next section). We provide here examples of mosaic GIs in the *P. aeruginosa* VRFPA04 and *P. fluorescens* UK4 genomes. These GIs not only harbor genes often associated with HGT but also have unusually high sequence similarity with distantly related genomes.

GI-11 of *P. aeruginosa* VRFPA04, located at 906,202–935,582 bp, is comprised of two distinct segments, Segment I from 906,202 to 923,513 bp and Segment II from 923,514 to 935,582 (Fig. 2a). Segment I encode proteins involved in multidrug resistance and virulence. This segment



**Fig. 2** **a** Gene map of a mosaic GI (906,202–935,582 bp) of *Pseudomonas aeruginosa* VRFPA04. Genes in the first segment (906,202–923,513 bp) are shown by block arrows with vertical lines and genes in the second segment (923,514–935,582 bp) are shown by block arrows with checkered boxes. **b** Gene map of a mosaic GI (4,167,919–4,191,679 bp) of *Pseudomonas fluorescens* UK4. Genes in the first segment (4,167,919–4,178,307 bp) are shown by

block arrows with vertical lines and genes in the second segment (4,178,308–4,191,679 bp) are shown by block arrows with checkered boxes. Genes annotated as hypothetical protein genes are shown in gray. Genes often associated with GI transfer are shown in red, virulence genes in pink, and genes involved in metabolism are shown in yellow

harbors genes encoding proteins such as multidrug transporter belonging to small multidrug resistance (SMR) family, Resistance-Nodulation-Division (RND) transporter, and acriflavin-resistance TetR family transcriptional regulator, which are involved in drug resistance. This segment also contains genes encoding virulence factors such as type VI secretion protein and iron ABC transporter ATP-binding protein. In addition to the resistance and virulence genes, this segment harbors genes encoding transposase and conjugal transfer protein that may be involved in horizontal transfer. Similarly, Segment II harbors genes associated with GI transfer, such as transposase and conjugal transfer protein-encoding genes, in addition to the genes that code for proteins involved in virulence such as peptidase and proteins potentially involved in drug resistance such as ABC transporter permease. Sequence similarity suggests that Segment I originates from *Serratia* sp. SSN1H1 and Segment II from *Delftia tsuruhatensis* (Table S5). This island was identified in a single island cluster (Cluster 904, Table S3) and thus, is unique to this strain.

Likewise, in *P. fluorescens* UK4, a mosaic GI (GI-20 in Fig. 1c) is comprised of two compositionally distinct segments, Segment I located at 4,167,919–4,178,307 bp and Segment II 4,178,308–4,191,679 bp; these segments display gene content that indicates their shared and differential functions (Fig. 2b). Each segment carries genes often associated with GI transfer; the Segment I harbors an integrase gene and the Segment II has genes encoding transposase and conjugal transfer protein (Fig. 2b). While Segment I harbors genes involved in metabolism, Segment II contains a type VI secretion protein-encoding gene known to be involved in virulence. These segments also likely have different origins. Segment I has high similarity with a sequence in *Orrella dioscoreae* and Segment II displays high similarity with a sequence in *Comamonas testosteroni* TK (Table S5). This GI belongs to a single island cluster (Cluster 1613, Table S5), and thus is a biomarker of this strain.

These examples suggest that segments comprising a mosaic GI may work synergistically to contribute to a complex trait, e.g., virulence and resistance, or they may have disparate functions to add to the metabolic repertoire of the host genome.

### Analysis of mosaic islands

For each distinct segment in a mosaic GI, best BLAST hits with unusually very high similarity in distant taxa were obtained. These are strong cases of potentially very recent transfers. Instances of such transfers that likely led to the emergence of mosaic GIs are enlisted in Table S5, with putative donor taxa and BLAST alignment scores (identity, coverage, bit score, and e-value) indicated. We further investigated whether the donors have cohabited the same

environment as of the recipient, which allows opportunities to exchange DNA among phylogenetically distant organisms. Indeed, we found many instances of shared ecology, which provides insights into the co-evolution of the organisms via gene sharing. The donors reflect the diversity observed in the habitats of *Pseudomonas* spp.

For *P. aeruginosa*, the most frequent donor from our analyses is *Azotobacter chroococcum*. 80 segments of *P. aeruginosa*'s mosaic GIs had their best BLAST hits in *A. chroococcum*. *A. chroococcum* is an obligately aerobic nitrogen-fixing bacterium living mainly in soil (Robson et al. 2015). *P. aeruginosa* can also be found in soil (Schroth et al. 2018), the shared ecology may be facilitating the DNA transfer among these bacteria. Likewise, the potential donor of GIs in *P. aeruginosa* that had the next highest number of BLAST hits, *Stenotrophomonas rhizophila*, is also found in soil (Wolf et al. 2002). BLAST analysis also revealed donors that are potentially pathogenic. Amongst these, *Escherichia coli* was prominent, along with pathogens associated with cystic fibrosis such as *Bordetella hinzii* (Funke et al. 1996), *Stenotrophomonas maltophilia* (Brooke 2012; Demko et al. 1998), *Bordetella bronchiseptica* (Spilker et al. 2008) and *Bordetella petrii* (Spilker et al. 2008). *E. coli* is often observed in the human gastrointestinal tract (Savageau 1983), where *P. aeruginosa* is known to cause gastrointestinal infections (Ohara and Itoh 2003). *P. aeruginosa* is also commonly found in cystic fibrosis patients (Emerson et al. 2002). Lungs of cystic fibrosis patients are known to act as reservoirs of bacteria with frequent gene transfers (Jani et al. 2016).

*Pseudomonas chlororaphis* (Bodelier et al. 1997), *P. fluorescens* (Paulsen et al. 2005), *P. putida* (Fernández et al. 2012), and *P. stutzeri* (Lalucat et al. 2006) are found in soil. This is also reflected in our donor analyses of their mosaic GIs. *Paucimonas lemoignei* is the most common donor of mosaic GI segments in *P. chlororaphis*. High sequence similarity between segments from a betaproteobacterium (*P. lemoignei*) and a gammaproteobacterium (*P. chlororaphis*) further lends credence to our predictions of GIs based on compositional anomaly.

Likewise, mosaic GI segments of *P. fluorescens* and *P. putida* showed unusually high sequence similarity with DNA segments of distantly related *Stenotrophomonas rhizophila*. *S. rhizophila* belongs to the class *Xanthomonadales* whereas the recipients *P. fluorescens* and *P. putida* belong to the class *Pseudomonadales*. *P. stutzeri*'s mosaic GI segments had the highest sequence similarity with those in *Pseudomonadaceae* bacterium SI-3.

We did not find high sequence similarity between mosaic GI segments of *P. syringae* and DNA segments of distantly related bacteria, perhaps indicating that these GI segments have been long time residents of the recipient bacteria and thus their compositional signatures have ameliorated to that

of its host, *P. syringae*. Other plausible scenarios could be transfers from closely related donors or rapid evolution of the GI segments since the acquisition.

### Acquisition of alien gene clusters

Using BLAST, we also found several instances (Table S6) where segments comprising a mosaic GI displayed high sequence similarities to potential donors from different lineages. For example, a mosaic GI located at 5,947,763–5,978,518 bp in the *P. aeruginosa* AR 0357 genome has its first segment (5,947,763–5,962,553 bp) highly similar to a genomic segment in *Bordetella bronchiseptica* and its other segment (5,962,554–5,978,518 bp) highly similar to a region in the *Serratia sp.* SSNIH1 genome. Both *B. bronchiseptica*, a betaproteobacterium, and *Serratia sp.* SSNIH1, a gammaproteobacterium, are distantly related to each other and to the recipient *P. aeruginosa*. While *P. aeruginosa* and *B. bronchiseptica* are found in patients with cystic fibrosis (Emerson et al. 2002; Spilker et al. 2008), *Serratia* has also been occasionally found in cystic fibrosis patients (Coenye et al. 2002). The cohabitation thus might have allowed transfer of genetic materials even among distantly related bacteria. We observed similar instances in *P. fluorescens* (*Orrella dioscoreae* and *Comamonas testosteroni TK102* as potential donors of disparate segments of a mosaic GI) and *P. putida* (*Klebsiella michiganensis* and *Salmonella enterica* as the potential donors) (Table S6). The mosaicism of these islands, deciphered based on compositional bias, was thus supported by sequence comparison via alignment as well. These instances where usually high-nucleotide sequence similarity was observed in otherwise distantly related bacteria indicate recent horizontal transfers. While here we focused only on recent transfers, further studies that may utilize conservation of amino acids are needed to infer potential donors for ancient DNA acquisitions.

### Conclusions

Our analyses show that *Pseudomonas* spp. are capable of acquiring genetic elements from a diverse set of potential donors. The extent of putative foreign genes in *Pseudomonas* genomes, as revealed by this study, elucidates the propensity of *Pseudomonas* spp. to acquire alien gene clusters from different lineages. In several cases, the donors inferred by our analyses are known to cohabit the same environment as the recipient *Pseudomonas* spp. The diversity of the donors and their phylogenetic distance from the recipient *Pseudomonas* spp. suggests that in several instances, horizontal DNA exchange is driven by shared ecology that facilitates DNA exchange among even

distantly related organisms, supporting previous studies on gene exchange. Our results also highlight selection for co-localization of DNA segments from different donors, as in mosaic GIs that may be conferring novel traits to increase the fitness under stressful conditions.

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.1007/s00203-021-02253-2>.

**Funding** The authors did not receive support from any organization for the submitted work.

### Declarations

**Conflict of interest** The authors have no relevant financial or non-financial interests to disclose.

### References

- Alfano JR, Charkowski AO, Deng W, Badel JL, Petnicki-Ocwieja T, Van Dijk K et al (2000) The *Pseudomonas syringae* Hrp pathogenicity island has a tripartite mosaic structure composed of a cluster of type III secretion genes bounded by exchangeable effector and conserved effector loci that contribute to parasitic fitness and pathogenicity in plants. *Proc Natl Acad Sci USA* 97(9):4856–4861
- Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic local alignment search tool. *J Mol Biol* 215(3):403–410
- Ayub ND, Pettinari MJ, Méndez BS, Lopez NI (2007) The polyhydroxyalkanoate genes of a stress resistant Antarctic *Pseudomonas* are situated within a genomic island. *Plasmid* 58(3):240–248
- Bodelier PL, Wijnhuizen AG, Blom CW, Laanbroek HJ (1997) Effects of photoperiod on growth of and denitrification by *Pseudomonas chlororaphis* in the root zone of *Glyceria maxima*, studied in a gnotobiotic microcosm. *Plant Soil* 190(1):91–103
- Brooke JS (2012) *Stenotrophomonas maltophilia*: an emerging global opportunistic pathogen. *Clin Microbiol Rev* 25(1):2–41
- Buell CR, Joardar V, Lindeberg M, Selengut J, Paulsen IT, Gwinn ML et al (2003) The complete genome sequence of the Arabidopsis and tomato pathogen *Pseudomonas syringae* pv. tomato DC3000. *Proc Natl Acad Sci USA* 100(18):10181–10186
- Coenye T, Goris J, Spilker T, Vandamme P, LiPuma JJ (2002) Characterization of unusual bacteria isolated from respiratory secretions of cystic fibrosis patients and description of *Inquilinus limosus* gen. nov. sp. nov. *J Clin Microbiol* 40(6):2062–2069
- Demko CA, Stern RC, Doershuk CF (1998) *Stenotrophomonas maltophilia* in cystic fibrosis: incidence and prevalence. *Pediatr Pulmonol* 25(5):304–308
- Emerson J, Rosenfeld M, McNamara S, Ramsey B, Gibson RL (2002) *Pseudomonas aeruginosa* and other predictors of mortality and morbidity in young children with cystic fibrosis. *Pediatr Pulmonol* 34(2):91–100
- Fernández M, Niqui-Arroyo JL, Conde S, Ramos JL, Duque E (2012) Enhanced tolerance to naphthalene and enhanced rhizoremediation performance for *Pseudomonas putida* KT2440 via the NAH7 catabolic plasmid. *Appl Environ Microbiol* 78(15):5104–5110
- Funke G, Hess T, von Graevenitz A, Vandamme P (1996) Characteristics of *Bordetella hinzii* strains isolated from a cystic fibrosis patient over a 3-year period. *J Clin Microbiol* 34(4):966–969

- Jani M, Azad RK (2019) IslandCafe: compositional anomaly and feature enrichment assessment for delineation of genomic Islands. *G3 (Bethesda)* 9(10):3273–3285
- Jani M, Mathee K, Azad RK (2016) Identification of novel genomic islands in liverpool epidemic strain of *Pseudomonas aeruginosa* using segmentation and clustering. *Front Microbiol* 7:1210
- Juhas M, van der Meer JR, Gaillard M, Harding RM, Hood DW, Crook DW (2009) Genomic islands: tools of bacterial horizontal gene transfer and evolution. *FEMS Microbiol Rev* 33(2):376–393
- Klockgether J, Reva O, Larbig K, Tümmler B (2004) Sequence analysis of the mobile genome island pKLC102 of *Pseudomonas aeruginosa* C. *J Bacteriol* 186(2):518–534
- Koonin EV, Makarova KS, Aravind L (2001) Horizontal gene transfer in prokaryotes: quantification and classification. *Annu Rev Microbiol* 55(1):709–742
- Lalucut J, Bennisar A, Bosch R, García-Valdés E, Palleroni NJ (2006) Biology of *Pseudomonas stutzeri*. *Microbiol Mol Biol Rev* 70(2):510–547
- Mathee K, Narasimhan G, Valdes C, Qiu X, Mawish JM, Koehrsen M et al (2008) Dynamics of *Pseudomonas aeruginosa* genome evolution. *Proc Natl Acad Sci USA* 105(8):3100–3105
- Mavrodi DV, Loper JE, Paulsen IT, Thomashow LS (2009) Mobile genetic elements in the genome of the beneficial rhizobacterium *Pseudomonas fluorescens* Pf-5. *BMC Microbiol* 9(1):8
- Nelson KE, Weinel C, Paulsen IT, Dodson RJ, Hilbert H, Martins dos Santos V et al (2002) Complete genome sequence and comparative analysis of the metabolically versatile *Pseudomonas putida* KT2440. *Environ Microbiol* 4(12):799–808
- Noto MJ, Kreiswirth BN, Monk AB, Archer GL (2008) Gene acquisition at the insertion site for SCCmec, the genomic island conferring methicillin resistance in *Staphylococcus aureus*. *J Bacteriol* 190(4):1276–1283
- Ohara T, Itoh K (2003) Significance of *Pseudomonas aeruginosa* colonization of the gastrointestinal tract. *Intern Med* 42(11):1072–1076
- Osborn AM, Boltner D (2002) When phage, plasmids, and transposons collide: genomic islands, and conjugative-and mobilizable-transposons as a mosaic continuum. *Plasmid* 48(3):202–212
- Paulsen IT, Press CM, Ravel J, Kobayashi DY, Myers GS, Mavrodi DV et al (2005) Complete genome sequence of the plant commensal *Pseudomonas fluorescens* Pf-5. *Nat Biotechnol* 23(7):873–878
- Qiu X, Gurkar AU, Lory S (2006) Interstrain transfer of the large pathogenicity island (PAPI-1) of *Pseudomonas aeruginosa*. *Proc Natl Acad Sci USA* 103(52):19830–19835
- Robson RL, Jones R, Robson RM, Schwartz A, Richardson TH (2015) *Azotobacter* genomes: the genome of *Azotobacter chroococcum* NCIMB 8003 (ATCC 4412). *PLoS ONE* 10(6):e0127997
- Savageau MA (1983) *Escherichia coli* habitats, cell types, and molecular mechanisms of gene control. *Am Nat* 122(6):732–744
- Schroth MN, Cho JJ, Green SK, Kominos SD (2018) Epidemiology of *Pseudomonas aeruginosa* in agricultural areas. *J Med Microbiol* 67(8):1191–1201
- Spilker T, Liwiński AA, LiPuma JJ (2008) Identification of *Bordetella* spp. in respiratory specimens from individuals with cystic fibrosis. *Clin Microbiol Infect* 14(5):504–506
- Steinegger M, Söding J (2018) Clustering huge protein sequence sets in linear time. *Nat Commun* 9(1):1–8
- Stover CK, Pham XQ, Erwin AL, Mizoguchi SD, Warrener P, Hickey MJ et al (2000) Complete genome sequence of *Pseudomonas aeruginosa* PAO1, an opportunistic pathogen. *Nature* 406(6799):959
- van der Meer JR, Sentschilo V (2003) Genomic islands and the evolution of catabolic pathways in bacteria. *Curr Opin Biotechnol* 14(3):248–254
- Winsor GL, Griffiths EJ, Lo R, Dhillon BK, Shay JA, Brinkman FS (2016) Enhanced annotations and features for comparing thousands of *Pseudomonas* genomes in the *Pseudomonas* genome database. *Nucleic Acids Res* 44(D1):D646–D653
- Wolf A, Fritze A, Hagemann M, Berg G (2002) *Stenotrophomonas rhizophila* sp. nov., a novel plant-associated bacterium with antifungal properties. *Int J Syst Evol Microbiol* 52(6):1937–1944

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.