

# Comparative genomics of the *Pseudomonas corrugata* subgroup reveals high species diversity and allows the description of *Pseudomonas ogarae* sp. nov.

Daniel Garrido-Sanz<sup>1,2</sup>, Miguel Redondo-Nieto<sup>1</sup>, Marta Martin<sup>1</sup> and Rafael Rivilla<sup>1,\*</sup>

## ABSTRACT

*Pseudomonas corrugata* constitute one of the phylogenomic subgroups within the *Pseudomonas fluorescens* species complex and include both plant growth-promoting rhizobacteria (PGPR) and plant pathogenic bacteria. Previous studies suggest that the species diversity of this group remains largely unexplored together with frequent misclassification of strains. Using more than 1800 sequenced *Pseudomonas* genomes we identified 121 genomes belonging to the *P. corrugata* subgroup. Intergenomic distances obtained using the genome-to-genome blast distance (GBDP) algorithm and the determination of digital DNA–DNA hybridization values were further used for phylogenomic and clustering analyses, which revealed 29 putative species clusters, of which only five correspond to currently named species within the subgroup. Comparative and functional genome-scale analyses also support the species status of these clusters. The search for PGPR and plant pathogenic determinants showed that approximately half of the genomes analysed could have a pathogenic behaviour based on the presence of a pathogenicity genetic island, while all analysed genomes possess PGPR traits. Finally, this information together with the characterization of phenotypic traits, allows the reclassification proposal of *Pseudomonas fluorescens* F113 as *Pseudomonas ogarae* sp. nov., nom rev., type strain F113<sup>T</sup> (=DSM 112162<sup>T</sup>=CECT 30235<sup>T</sup>), which is substantiated by genomic, functional genomics and phenotypic differences with their closest type strains.

## DATA SUMMARY

The genomic and proteomic sequences used throughout this study are publicly accessible through the National Center for Biotechnology Information (NCBI), under the assembly accession numbers provided in File S1 (available in the online version of this article).

The authors confirm all supporting data, code and protocols have been provided within the article or through supplementary data files.

## INTRODUCTION

*Pseudomonas* is a highly diverse and metabolically versatile genus of *Gammaproteobacteria* [1] ubiquitously found in the environment. Members of this genus thrive from tropics to polar and desertic habitats [2–5] and have been found in fresh and sea waters [6, 7] and soils worldwide [8]. *Pseudomonas* can also colonize multiple niches, including humans, animals and plant tissues [9–12]. Among plant-associated *Pseudomonas*, both commensal and pathogenic lifestyles are reported across different species. For example, *P. syringae* is a known pathogen of a wide range of plant

Received 16 February 2021; Accepted 26 April 2021; Published 29 June 2021

**Author affiliations:** <sup>1</sup>Departamento de Biología, Facultad de Ciencias, Universidad Autónoma de Madrid, Darwin 2, 28049 Madrid, Spain; <sup>2</sup>Department of Fundamental Microbiology, University of Lausanne, CH-1015 Lausanne, Switzerland.

**\*Correspondence:** Rafael Rivilla, rafael.rivilla@uam.es

**Keywords:** *Pseudomonas corrugata*; *Pseudomonas ogarae*; phylogenomics; comparative genomics; PGPR.

**Abbreviations:** ACC, 1-aminocyclopropane-1-carboxylic acid; ANI, average nucleotide identity; CECT, Spanish Type Culture Collection; DAPG, 2,4-diacetylphloroglucinol; dDDH, digital DNA–DNA hybridization; DSMZ, German Collection of Microorganisms and Cell Cultures; GBDP, genome-to-genome blast distance; KEGG, Kyoto Encyclopedia of Genes and Genomes; LPQ, lipopeptide/quorum sensing; MCL, Markov Cluster Algorithm; ML, maximum-likelihood; MLSA, multilocus sequence analysis; MRI, modified Rand Index; NCBI, National Center for Biotechnology Information; NJ, neighbor-joining; NMDS, A non-metric multidimensional scaling; OG, orthologous groups; PCA, principal component analysis; PGPR, plant growth-promoting rhizobacteria; T3SS, type III secretion system; T6SS, type VI secretion system; TYGS, Type (Strain) Genome Server.

The genome sequence of *Pseudomonas ogarae* RDP1, sequenced in this study, is publicly available at the NCBI under the BioProject accession number PRJNA697838.

**Data statement:** All supporting data, code and protocols have been provided within the article or through supplementary data files. Supplementary material is available with the online version of this article.

000593 © 2021 The Authors



This is an open-access article distributed under the terms of the Creative Commons Attribution License.

species [13]. Other plant pathogens include *P. corrugata* and *P. mediterranea* species, the causing agents of pith necrosis in tomato [14]. In contrast, multiple *Pseudomonas* species are recognized for their beneficial plant role, known as plant growth-promoting rhizobacteria (PGPR), and are mainly found within isolates and species from the *Pseudomonas fluorescens* complex of species [15, 16], including *P. brassicacearum*, *P. chlororaphis* and *P. protegens* [17–19].

Phylogenetic analyses have divided this diverse genus into two main lineages: *P. aeruginosa* and *P. fluorescens* [15, 20]. The *P. fluorescens* lineage is composed of three large groups: *P. fluorescens*, *P. syringae* and *P. putida*, commonly referred to as species complexes [15, 21, 22] due to the multiple species that are found within each of these groups. This complexity is further reflected in the number of different subgroups in which each of these larger groups can be subdivided. For instance, the *P. fluorescens* species complex is composed of nine distinct phylogenomic subgroups: *P. fluorescens*, *P. gesardii*, *P. fragi*, *P. mandelii*, *P. jessenii*, *P. korensis*, *P. chlororaphis*, *P. protegens* and *P. corrugata* [15, 23]. The delineation of these different subgroups is supported by multilocus sequence analysis (MLSA), phylogenomics using average nucleotide identity (ANI) [24] and using the genome-to-genome blast distance phylogeny (GBDP) algorithm [25], which also digitally replicates the DNA–DNA hybridization ‘gold standard’ for species delineation and therefore provides a robust and accurate delineation of species and phylogenomic groups [15, 26]. In addition, comparative genomic analysis of the *P. fluorescens* species complex has shown that certain traits involved in the PGPR and biocontrol abilities of strains from these subgroups follow a phylogenomic distribution [15].

The *P. corrugata* subgroup within the *P. fluorescens* species complex is composed of five named species: *P. corrugata*, *P. mediterranea*, *P. thivervalensis*, *P. kilonensis* and *P. brassicacearum* [15, 23], although the number of species of this subgroup is probably higher, as suggested in multiple studies [15, 26–28], which also emphasize frequent misclassification of strains. In addition, the *P. corrugata* subgroup is particularly baffling, since it includes recognized PGPR members together with plant-pathogenic species and strains. Among the PGPR traits, the production of the antibiotic 2,4-diacetylphloroglucinol (DAPG) and 1-aminocyclopropane-1-carboxylic acid (ACC) deaminase are probably the most studied characters within PGPR pseudomonads. The biosynthesis of the broad-spectrum antibiotic DAPG [29] has also shown to elicit the plant-induced systemic resistance [30] and stimulate lateral root growth by interacting with the signalling pathway of the phytohormone auxin [31]. On the other hand, ACC deaminase contributes to phytostimulation by promoting root growth [32] and also alleviate abiotic stresses [33, 34]. Conversely, the *P. corrugata* subgroup also includes the causal agents of pith necrosis in tomato, *P. corrugata* [35], which gives the name to this subgroup by oldest species description, and *P. mediterranea* [36], together with other plant pathogens, including *Pseudomonas* sp. N2C3, which

### Impact Statement

The *Pseudomonas corrugata* subgroup within the *Pseudomonas fluorescens* species complex comprise both phytobeneficial and phytopathogenic species. Therefore, their use as inoculants in agriculture requires clear identification at the species level. Currently, there are only five recognized species, including *P. corrugata* and *P. mediterranea*, both described as phytopathogens. The analysis of a large number of sequenced genomes belonging to this group could provide a better understanding of how opposed lifestyles differ within these closely related species. Using genome-wide comparative genomic analyses, we have determined that the *P. corrugata* subgroup contains 29 species, most of them so far unacknowledged. PGPR traits are distributed across all the species groups, while a plant-pathogenic determinant is found in half of the genomes analysed. No plant-pathogenic determinants are found within the *P. brassicacearum*, *P. kilonensis*, *P. thivervalensis* and the newly described species *P. ogarae*, validating their use as agricultural inoculants. This new species type strain, F113<sup>T</sup> (formerly *P. fluorescens* F113), is a model rhizobacterium for competitive root colonization and a well-known biocontrol agent. The results presented here also provide the basis for further analysis of the remaining uncharacterized species and expose an enormous bacterial diversity within the *P. corrugata* phylogenomic subgroup.

kills or stunt plants from the families *Brassicaceae* and *Papaveroideae* [37]. Within the plant-associated pathogenic pseudomonads of this subgroup, a genetic island containing non-ribosomal peptide synthetase genes and genes similar to the acyl-homoserine lactone quorum-sensing system in *Proteobacteria* [named lipopeptide/quorum sensing (LPQ) island] is found and reportedly responsible for their plant-pathogenic performance [37, 38]. Surprisingly, *P. corrugata* is used for the biocontrol of a number of plant-pathogenic bacteria and fungi [39], and they also harbour multiple PGPR traits, including hydrogen cyanide and auxin biosynthesis [27]. This ambivalent behaviour could be explained by other parameters, such as the genotype and physiological state of the plant or the ability of bacteria to manipulate the plant immune system [27].

In this work, we analysed more than a hundred genomes belonging to the *P. corrugata* subgroup within the *P. fluorescens* species complex in order to determine their current species status. Using phylogenomics and comparative and functional genomics we determined a high number of species groups that were previously unacknowledged. Using a polyphasic approach we propose the reclassification of *P. fluorescens* F113 into the novel species *P. ogarae* sp. nov., type strain F113, which is substantiated by phylogenomics, functional genomics and phenotypic analyses.

## METHODS

### Identification of genomes belonging to the *P. corrugata* subgroup

From the 10785 *Pseudomonas* genomes listed in the NCBI RefSeq database in February 2020, those of species and lineages outside *P. fluorescens* phylogenetic groups (as previously described in [15]) were removed. These include genomes assigned to *P. aeruginosa*, *P. putida*, *P. syringae* complex of species and *P. stutzeri*. The remaining 1811 *Pseudomonas* genomes were downloaded from the NCBI RefSeq database in February 2020.

Genomes belonging to the *P. corrugata* phylogenomic subgroup were determined based on the GBDP distance threshold of 0.132915 previously established for the *P. fluorescens* complex of species [15] using the genome of *P. ogarae* F113<sup>T</sup> (this study, formerly *P. fluorescens* F113) as a reference, which has been previously classified into the *P. corrugata* subgroup in several previous studies [15, 26, 40, 41]. Overall, 121 genomes were shown to belong to the *P. corrugata* subgroup (File S1) and therefore were selected for further analyses.

### Genome sequencing, assembly and annotation of *P. ogarae* RDP1

*P. ogarae* RDP1 was isolated from the rhizosphere of pepper (*Capsicum annuum*) as previously described [26]. Total DNA was extracted using the Realpure Genomic DNA extraction Kit (Durviz, Spain) and sequenced by paired-end Illumina MiSeq 2×150 at Parque Científico de Madrid (Spain). Reads were quality-filtered using Trimmomatic software v0.38 [42]. The genome of *P. ogarae* RDP1 was assembled using SPAdes v3.14.1 software [43], read error correction and careful parameters. Prokka v1.14.6 software [44] was used to annotate the genome, using the genetic code number 11 and default parameters. The genome sequence has been deposited in the NCBI and it is publicly available under the BioProject accession number PRJNA697838.

### Phylogenomic analysis and clustering

Intergenomic distances among the 121 identified *P. corrugata* subgroup genomes and representative type strain sequenced genomes of the remaining *P. fluorescens* subgroups (*P. mandelii* subgroup: *P. mandelii* LMG 21607<sup>T</sup> and *P. lini* CCUG 51522<sup>T</sup>; *P. jessenii* subgroup: *P. jessenii* DSM 17150<sup>T</sup> and *P. vacuoverensis* CCUG 49675<sup>T</sup>; *P. koreensis* subgroup: *P. koreensis* CCUG 51519<sup>T</sup> and *P. baetica* LMG 25716<sup>T</sup>; *P. chlororaphis* subgroup: *P. chlororaphis* subsp. *chlororaphis* DSM 50083<sup>T</sup>; *P. protegens* subgroup: *P. protegens* CHA0<sup>T</sup>; *P. fluorescens* subgroup: *P. veronii* DSM 11331<sup>T</sup>, *P. azotoformans* DSM 18862<sup>T</sup>, *P. fluorescens* ATCC 13525<sup>T</sup> and *P. synxantha* DSM 18928<sup>T</sup>; *P. gesardii* subgroup: *P. gesardii* DSM 17152<sup>T</sup> and *P. mucidolens* NBRC 103159<sup>T</sup>; *P. fragi* subgroup: *P. fragi* NBRC 2458<sup>T</sup>) and type strains genomes of other *Pseudomonas* main groups (*P. aeruginosa* ATCC 10145<sup>T</sup>, *P. resinovorans* DSM 21078<sup>T</sup>, *P. putida* NBRC 14164<sup>T</sup> and *P. syringae* KCTC 12500<sup>T</sup>) were calculated using the GBDP algorithm [25],

via the genome-to-genome distance calculator (GGDC: <http://ggdc.dsmz.de/>) web service. The obtained sets of intergenomic distances (File S2) were transformed into a matrix and imported to MEGA X software [45] to build the neighbour-joining (NJ) phylogenomic tree. *Escherichia coli* ATCC 11775<sup>T</sup> was used as the outgroup. GBDP was also used to calculate digital DNA–DNA hybridization (dDDH) values among all genome pairwise comparisons. In addition, the Type (Strain) Genome Server (TYGS: <https://tygs.dsmz.de/>) web service was used with the identified novel unnamed species-level groups for a whole genome-based taxonomic analysis [46], using default parameters.

Clustering of intergenomic distances at species level (dDDH≥70%) and into phylogenomic groups was examined using the OPTSIL clustering software v1.5 [47] as previously described [48]. Briefly, distance threshold (*T*) values from 0 to 0.2, using a step size of 0.0005 were evaluated using single-, average- and complete-linkage clustering (*F*=0, 0.5 and 1, respectively). The best *T* and *F* for both, species-level and phylogenomic group clusters were chosen based on the highest modified Rand Index (MRI) score.

### Orthologous groups identification and genome fractions

Proteomes of the 121 genomes belonging to the *P. corrugata* subgroup were analysed with OrthoFinder software v2.3.3 [49], using diamond v0.9.131 [50] searches and the MCL graph-clustering algorithm [51]. Resulting orthologous groups (OGs) were analysed with a previously designed R script [48] to obtain the core, group-specific and pan genome fractions over 1000 randomly sampled genomes. The mean, Q1 and Q3 statistics were represented using the ggplot2 R package [52].

### Phylogeny of single-copy genes

Orthologous sequences of 1891 single-copy genes present in the 121 *P. corrugata* subgroup genomes were aligned with Clustal Omega [53] and further concatenated. The resulting concatenated alignment was examined with gblocks v0.91 [54] to remove poorly aligned columns and highly divergent regions that might not be homologous or may have been saturated by multiple substitutions [55], using a minimum block length of two amino acids. The resulting matrix was imported into the RaxML-NG v0.9.0 [56] to build a maximum-likelihood (ML) phylogenetic tree, using the LG model of amino acid evolution [57] combined with gamma-distributed substitution rates and empirical frequencies of amino acids. Fast bootstrap and subsequent search for the best-scoring tree [58] and the autoMRE criterium [59] were applied.

### Functional annotation of orthologs and distribution

The functional annotation of orthologous groups was achieved using eggNOG-Mapper v2.0 [60] based on eggNOG v5.0 orthology data [61]. Sequence searches were performed using diamond. KEGG BRITE hierarchical classification was retrieved using the KEGGREST R package

[62] and represented using ggplot2. In addition, the UpSet R package [63] was used to represent specific intersections of the orthologous groups' matrix and aggregate number of grouped elements.

A principal component analysis (PCA) was performed with the prcomp function of the stats R package, using the relative abundance of the last-rank BRITE categories assigned to OGs in each genome (considering singletons and paralogues) as input variable. Forty-four genomes consisting of more than 75 contigs (File S1) were discarded to avoid misrepresentation of categories. Separation of PCA observation groups was measured and quantified as previously described [64]. Briefly, Mahalanobis distances [65] between groups, using the first two principal components' centroids of the PCA were used to calculate Hotelling's two-sample  $T^2$  statistic, which was then used to obtain the  $F$ -value. Finally,  $F$ -critical values were calculated using an  $\alpha=0.05$  and compared against the obtained  $F$ -value for significance. All the mathematical formulas applied are specified in [64].

#### Phenotypic characterization of *P. ogarae*, *P. kilonensis*, *P. brassicacearum* and *P. thivervalensis*

Type strains of *P. brassicacearum* DSM 13227<sup>T</sup>, *P. thivervalensis* DSM 13194<sup>T</sup> [66] and *P. kilonensis* DSM 13647<sup>T</sup> [67] were acquired from the German Collection of Microorganisms and Cell Cultures (DSMZ). *P. ogarae* F113<sup>T</sup> (this study, formerly *P. fluorescens*) was first isolated by the group of O'Gara in 1992 [68] and received in our laboratory in 1997, where its whole genome was sequenced [69]. The four strains together with *P. ogarae* RDP1 (this study), *Pseudomonas* sp. WCS365, *P. brassicacearum* NFM421 and *P. fluorescens* Q8r1-96 (the three belonging to *P. brassicacearum* species) were characterized at the phenotypic level by the determination of enzymatic activities and oxidation of carbon substrates by using API 20 NE and API 50 CHB/E (bioMérieux). API tests were done following the manufacturer specifications and read after 48 h of incubation at 28 °C. Assimilation of carbon substrates was determined using Biolog GEN III MicroPlate (Biolog) with some modifications. Fresh (overnight growth) cells were resuspended in 10 ml of minimal medium [70] until an optical density at 600 nm ( $OD_{600}$ ) of 0.05 was reached. In total, 100  $\mu$ l of each cell resuspension was transferred into the plate's wells and incubated at 28 °C in a rotary shaker (135 r.p.m.) for 16 h. Assimilation of carbon substrates was evaluated as the ability of the cells to grow by measuring the  $OD_{600}$  and its comparison to the negative control for the assimilation tests and the positive control for Biolog's sensitivity assays. Positive results were considered when the  $DO_{600}$  obtained was at least twice that of the negative control (for assimilation) and at least 75% of the positive control (for sensitivity).

A non-metric multidimensional scaling (NMDS) ordination analysis was used to plot the dissimilarity of the strains tested based on API and Biolog results. A positive result was given a value of 1, a negative result a value of 0 and a weak growth a value of 0.5. Only characters that show differences among the strains were selected. NMDS ordination analysis based

on Jaccard distances were calculated using the metaMDS function of the vegan R package [71], using a number of dimensions,  $k=2$ .

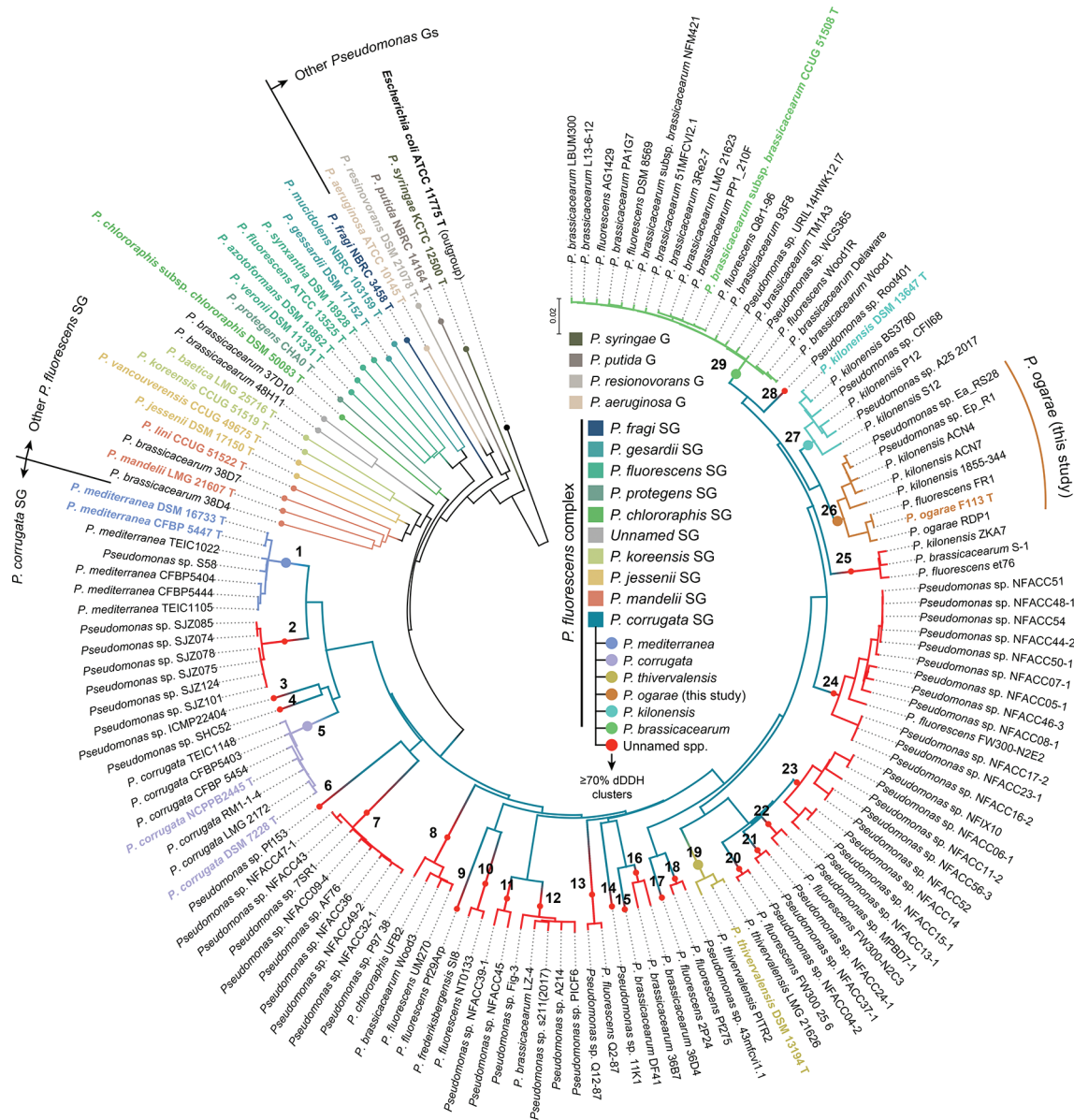
*P. ogarae* F113<sup>T</sup> (this study, formerly *P. fluorescens*) was deposited in the DSMZ culture collection and the Spanish Type Culture Collection (CECT) under the collection numbers DSM 112162<sup>T</sup> and CECT 30235<sup>T</sup>, respectively.

## RESULTS AND DISCUSSION

### Phylogenomic analysis and clustering of *P. corrugata* subgroup genomes

The 121 genomes identified as putatively belonging to the *P. corrugata* subgroup, together with representative genomes of the remaining *P. fluorescens* subgroups and other major *Pseudomonas* groups were used to calculate intergenomic distances and digital DNA–DNA hybridization values using the GBDP algorithm [25]. These distances were further used to build a phylogenomic tree and for clustering analysis. The phylogenomic tree obtained (Fig. 1) agrees with accepted phylogenies of *Pseudomonas* and the *P. fluorescens* complex of species [15, 23, 72]. The clustering analysis (Fig. 2, File S3) shows that, at the phylogenomic-group level, a distance threshold of 0.1465 and using a complete-linkage clustering, results in total consistency (i.e. MRI=1) with the reference partition, unequivocally assigning 121 genomes as belonging to the *P. corrugata* subgroup and discriminating the remaining *Pseudomonas* subgroups. The distance threshold obtained is slightly higher than the one previously established for eight *P. fluorescens* subgroups. In a previous study, using only 93 genomes belonging to the *P. fluorescens* complex of species, the authors identified an intergenomic distance threshold of 0.132915 [15], while here we report a distance of 0.1465 as the threshold for genomes belonging to the *P. corrugata* subgroup. These differences could be due to the larger dataset used here, which allows a more precise threshold determination, and to the fact that one subgroup and a single representative genome of each of the remaining *P. fluorescens* subgroups have been considered. Nonetheless, this value threshold assigned 121 *Pseudomonas* genomes over more than *c.a.* 1800 genomes, as belonging to the *P. corrugata* subgroup (File S1).

At the species level, the clustering of intergenomic distances revealed a far greater diversity than previously thought. At present, there are only five validly named species within the *P. corrugata* subgroup: *P. brassicacearum*, *P. kilonensis*, *P. thivervalensis*, *P. corrugata* and *P. mediterranea* [15, 23, 72]. However, based on the clustering of the intergenomic distances at a threshold  $T=0.036$ , which equals 70% of digital DNA–DNA hybridization (dDDH, i.e. the 'gold standard' for species delineation), there are 29 species clusters within the *P. corrugata* subgroup (Fig. 1), yielding total consistency (i.e. MRI=1) with the reference partition when using an average-linkage clustering (i.e.  $F=0.5$ , Fig. 2a). Among them, the largest species-cluster is *P. brassicacearum* (group 29, Fig. 1), composed of 19 genomes, including the type strain CCUG



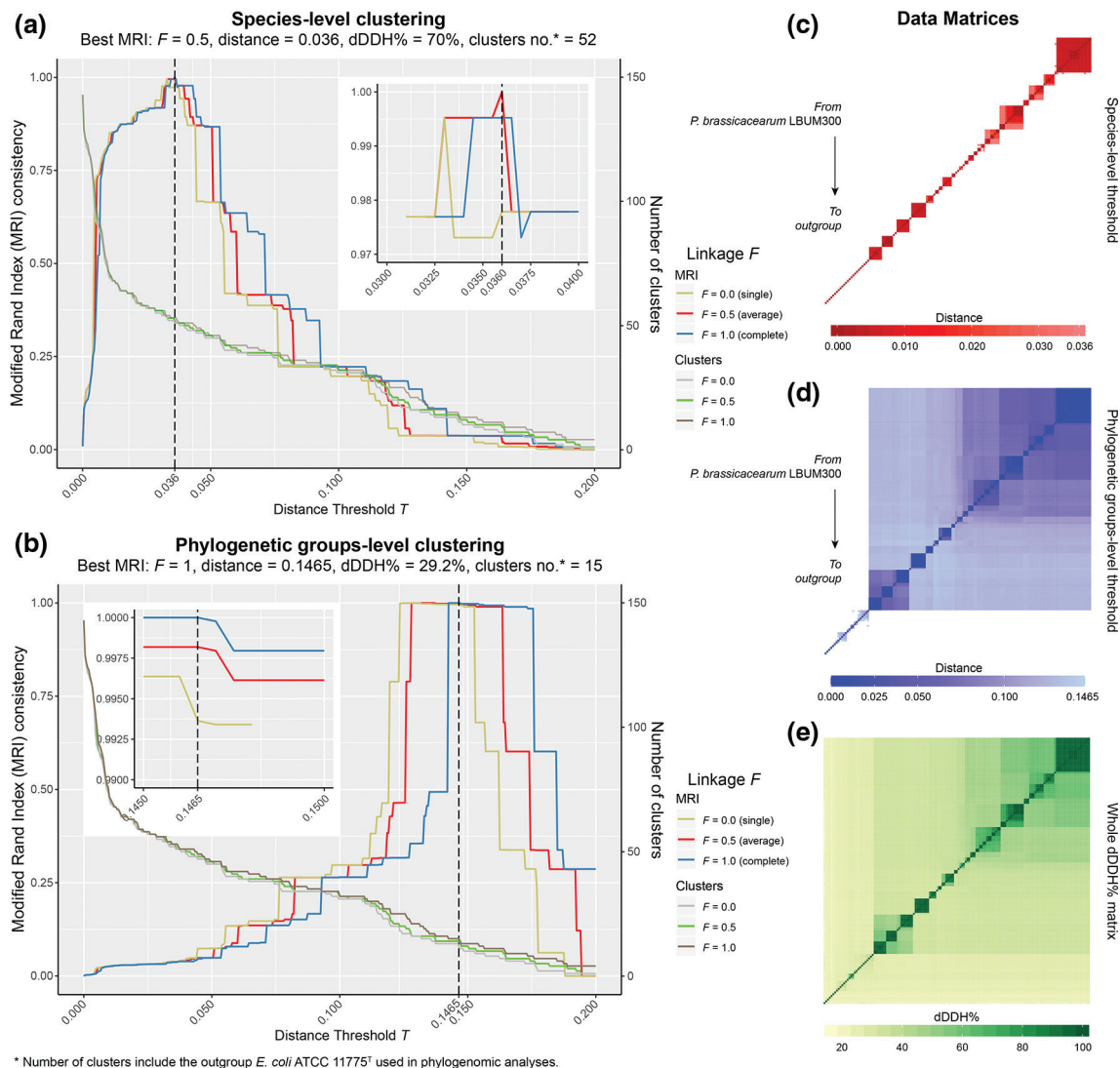
**Fig. 1.** GBDP-based phylogeny of 121 genomes belonging to *P. corrugata* and other major subgroups within the *P. fluorescens* complex of species and other major *Pseudomonas* groups. Colours according to the phylogenomic group (G), subgroup (SG) or species identified by clustering analyses of GBDP intergenomic distances (Fig. 2, File S3). Numbers in nodes leading to species-clusters are used throughout the study. Bold T, and coloured taxa name indicate type strain.

51508<sup>T</sup>, while eight species-level clusters contain only one representative genome.

Surprisingly, there are multiple genomes misclassified at the species level. These include *P. brassicacearum* strains S-1, 36D4, 36B7, DF41, Lz-4 and Wood3, and four others that do not even belong to the *P. corrugata* subgroup: 28D4 and 38D7 within the *P. mandelii* subgroup, and 48H11 and 37D10 strains that according to our analysis consist in a novel phylogenomic group within the *P. fluorescens* complex of species (Fig. 1), *P. kilonensis* strains ACN4, ACN7, 1855–344 and ZKA7 that also do not cluster together within the type strain *P. kilonensis*

DSM 13647<sup>T</sup>, and numerous strains wrongly assigned to *P. fluorescens*: AG1429, DSM 8569, Q8r1-96, Wood1R, FR1, F113 (*P. ogarae* this study, formerly *P. fluorescens*), et76, FW300-N2E2, FW300-N2C3, FW300\_25\_6, Pf275, 2P24, Q2-87, NT0133, Pf29Arp and UM270. There are also two strains wrongly assigned to *P. frederiksbergensis* SI8 and *P. chlororaphis* UFB2. Nonetheless, the majority of genomes belonging to the *P. corrugata* subgroup are still unclassified at the species level.

In addition, the Type (Strain) Genome Server (TYGS) [46], was used for a whole genome-based taxonomic analysis to

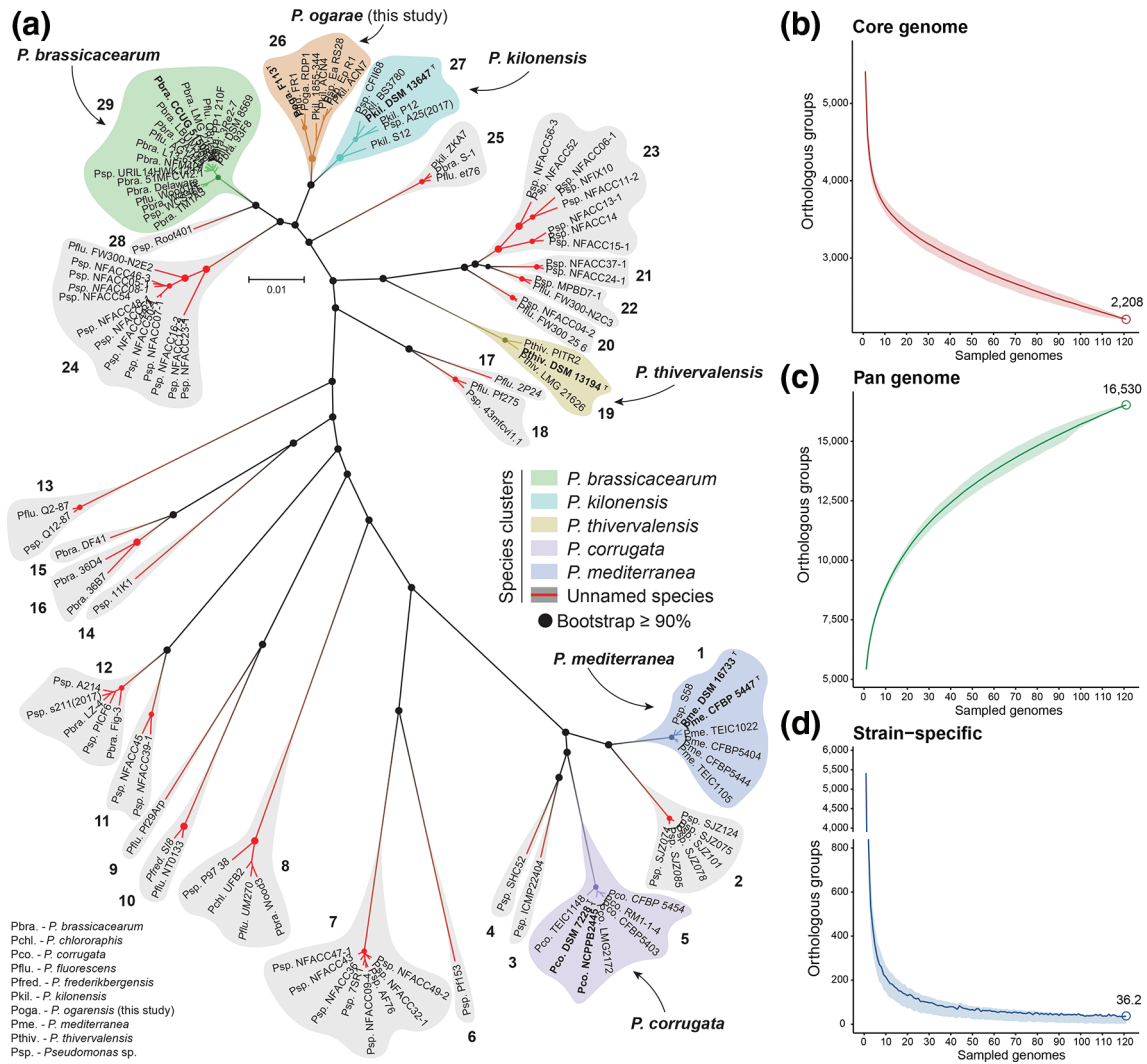


**Fig. 2.** Clustering analysis of 145 genomes belonging to the *P. corrugata* subgroup and the other main subgroups and groups within the *Pseudomonas* genus (Fig. 1) using a range of distance thresholds  $T$ . Total cluster consistency (i.e. MRI=1) was achieved using an average linkage (i.e.  $F=0.5$ ) at species-level (a) or complete linkage (i.e.  $F=1$ ) at phylogenomic group-level (b) compared to the reference partition. Distance matrices (c, d) and digital DNA–DNA hybridization (dDDH) matrix (e), showing these clusters.

further evaluate the 29 species-level clusters identified. The results are shown in the File S4 and validate species status for all the species-level clusters identified in this study. Only strain genomes clustered within the five named species (*P. brassicacearum*, *P. kilonensis*, *P. thivervalensis*, *P. corrugata* and *P. mediterranea*, Fig. 1) belong to these species, while the remaining clusters at the species level represent novel species.

One of these novel species cluster includes, among others, a widely used PGPR model of plant-root colonization: *Pseudomonas ogorae* F113<sup>T</sup> (this study, formerly *P. fluorescens* F113, F113 hereafter), isolated in 1992 from the rhizosphere of sugar beet [68]. Although F113 was one of the first genomes belonging to the *P. fluorescens* complex of species to be completely sequenced [69], it is long known that F113 indeed does not belong to the *P. fluorescens* species [15, 26, 40]. A

previous study tried to reclassify this strain as belonging to *P. kilonensis* [40]. Nonetheless, our analysis evidence that neither *P. fluorescens* nor *P. kilonensis* is the proper species name of F113, which constitutes a novel species together with strain RDP1 (this study), unclassified isolates Ea\_RS28 and Ep\_R1, misclassified *P. kilonensis* strains ACN4, ACN7 and 1855–344, and misclassified *P. fluorescens* strain FR1 (Fig. 1). Indeed, the comparison of F113 against *P. kilonensis* DSM 13647<sup>T</sup> achieved a dDDH of 64.5% (confidence interval of 61.6–67.3%, File S2), clearly demonstrating that F113 does not belong to *P. kilonensis* species. This was also substantiated by the TYGS analysis, which categorically shows that the eight mentioned strains belong to the same novel species, with no type strain of any other described species close enough (File S4) among more than 12000 type strains currently included



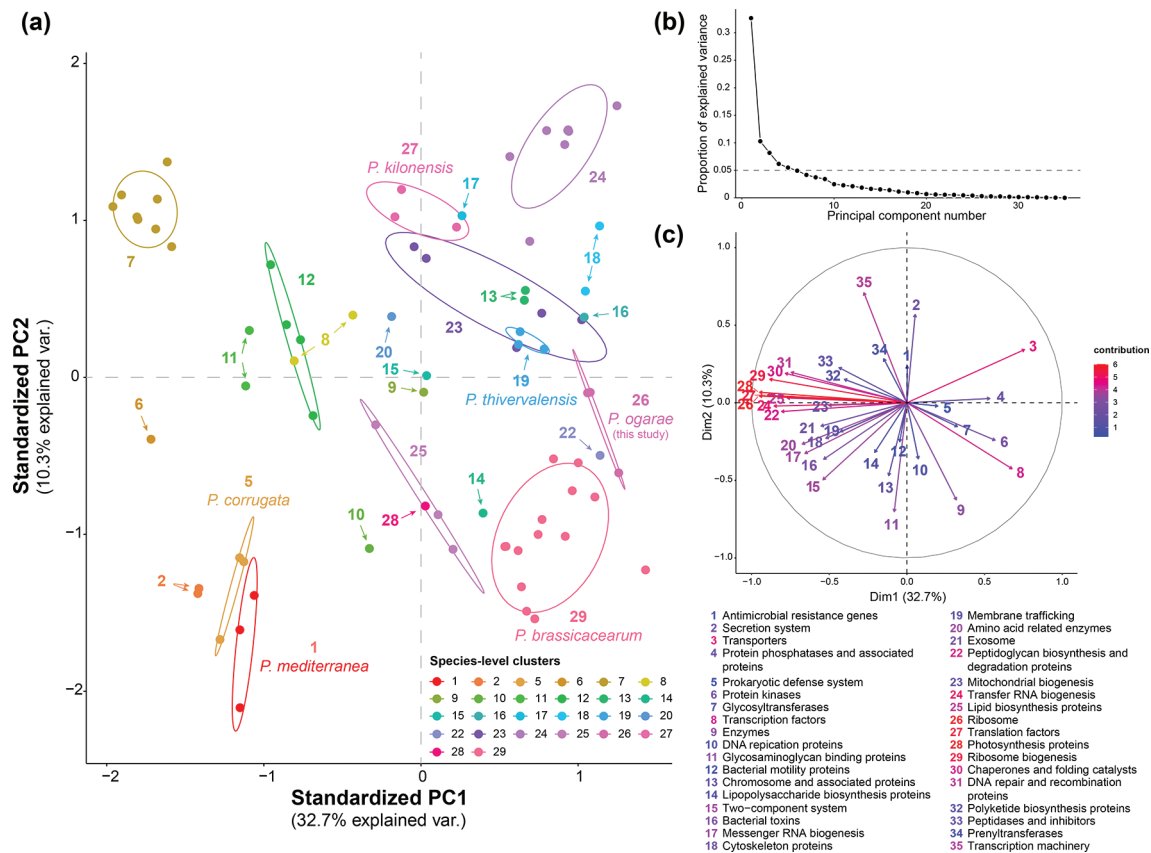
**Fig. 3.** Analysis based on the identification of orthologous groups. (a) Unrooted maximum-likelihood tree of 1891 single-copy amino acid sequences present in the 121 genomes belonging to the *P. corrugata* subgroup. Bold and † indicate type strain. Species groups are shadowed in colour according to the legend and numbers according to Fig. 1. Dots indicate bootstrap support  $\geq 90\%$  (not shown in nodes within a species cluster). Core genome (b), pan genome (c), and strain-specific genome (d) fractions over the sampled genomes. Mean values (line) and Q1 and Q3 quantiles (shadow) over 1000 replicates of randomly sampled genomes are represented. Empty circles and the number above indicate the mean genome fraction value achieved at 121 genomes sampled.

in the server. Among the 228 *Pseudomonas* species validly published to date (August 2020), the TYGS server included 201. The remaining 27 *Pseudomonas* type strains not included in the TYGS server do not belong to the *P. corrugata* subgroup according to 16S rRNA, MLSA, whole-genome analysis or a combination of these analyses (File S5).

### Single-copy proteins phylogeny and genome fractions

The determination of orthologous groups of amino acid sequences among the 121 *P. corrugata* subgroup genomes identified 1891 single-copy proteins present in all the genomes. These single-copy proteins were used to build a maximum-likelihood phylogenetic tree based on their

alignment and concatenation. The results (Fig. 3a) show a clustering pattern identical to that previously observed based on whole-genome analysis at species-level, with bootstrap support  $\geq 90\%$  in the nodes leading to these clusters. This result further validates the presence of 29 species within the *P. corrugata* subgroup. Nonetheless, the relative position of certain clusters differs compared to the phylogenomic tree based on intergenomic distances (Fig. 1). For instance, group 24 (composed of 13 unclassified isolates) is closer to *P. brassicacearum* in the core-proteins' tree (Fig. 3a) than in the whole-genome phylogeny (Fig. 1), which could imply a different evolutionary pressure of the core genome versus the whole genome or a bias in codon usage.



**Fig. 4.** Principal component analysis (PCA) of functional BRITE categories among *P. corrugata* subgroup genomes. (a) Biplot of the first two principal components (PC) explaining the 43% of the observed variance. The different groups are coloured according to the legend. Ellipses represent the core area added by the default confidence interval of 68% to facilitate the separation between observation groups. (b) Proportion of explained variance by each PC. Dashed horizontal line indicates 5% of explained variance. (c) Principal component variables, zero-centred and scaled to unit variance.

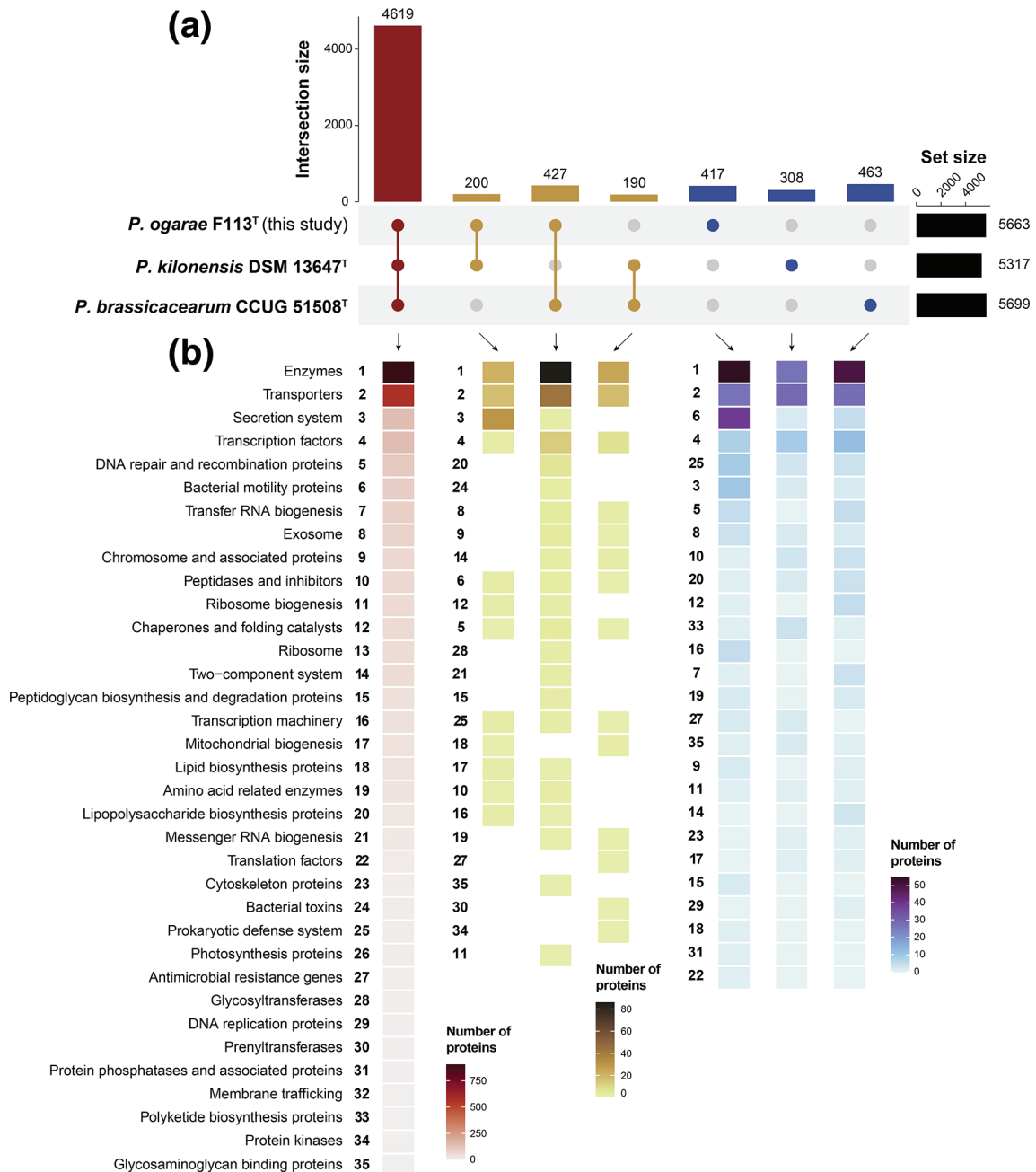
The core genome, pan genome and strain-specific genome were also determined based on the number of orthologous groups present in all genomes (core genome), the number of different orthologous groups (pan genome), and the novel orthologous groups (strain-specific genome) that appear with the addition of randomly sampled genomes (Fig. 3b–d). Regarding the core genome, with 121 genomes it is composed of 2208 orthologous groups. As expected, the core genome of the *P. corrugata* subgroup is smaller than the previously estimated (3438) based only on 12 genomes [15]. It is also considerably smaller than the core genome of only five *P. corrugata* and four *P. mediterranea* genomes, consisting of 4469 orthologous groups [14], probably due to a misrepresentation of the whole genetic diversity of this subgroup. On the other hand, the core-genome curve (Fig. 3b) shows a still decreasing slope, which indicates that it has not yet achieved a full representation of the subgroup core diversity. Similarly, the ‘open’ pan genome (Fig. 3c) of 16 530 orthologous groups suggests a higher genetic diversity yet to be discovered. Conversely, the strain-specific curve (Fig. 3d) shows that around 20 genomes are enough to cover most of the shared diversity and each genome adds

an average of 36 novel orthologous groups. This could indicate a high degree of genetic homogeneity among closely related clusters.

### Overall functional diversity of *P. corrugata* species groups

To further evaluate if the differences observed at the species groups based on intergenomic distances and the core-proteins also extend to the functional content of the groups, we performed a principal component analysis (PCA) using the BRITE categories assigned to the orthologous groups, previously calculated, as variables. To avoid misrepresentation of categories due to incomplete genomes, those with more than 75 contigs were removed from the analysis. The results show a clear distinction of genomes belonging to the same species-level cluster (Fig. 4a). The first two principal components explain the 43% of observed variance (32.7% PC1 and 10.3% PC2, respectively) and clearly distinguish *P. corrugata* and *P. mediterranea* pathogenic species [35, 36] from known PGPR, including *P. brassicacearum* (group 29 [17]) and *P. ogarae* (group 26 [32]). Interestingly, certain





**Fig. 5.** Comparative analysis of closely related *P. brassicacearum*, *P. kilonensis* and *P. ogarae* type strains genomes based on orthologous sequences. (a) Upset plot showing the intersections of shared orthologous proteins among the three genomes (red), the shared proteins between each two genomes (black) and specific proteins to each genome (blue). (b) Abundance of BRITE categories among the different intersections.

strains that have been shown to be plant pathogens, including N2C3 (group 22 [37]), are functionally more closely related to other PGPR groups than to *P. corrugata* and *P. mediterranea* (Fig. 4). This finding agrees with the observation of plant pathogens within the *P. corrugata* subgroup also harbour multiple PGPR traits, including hydrogen cyanide and auxin biosynthesis [27], and could indicate that traits relevant for pathogenesis are not the only driving feature

of certain groups. In fact, it is known that the pathogenicity of *P. corrugata*, *P. mediterranea* and other strains from the *P. corrugata* subgroup is associated with a genetic island [37] and, more interestingly, this pathogenic ability only occurs towards certain plant species, while in others these bacteria act as PGPR/biocontrol agents [27, 39]. Another interesting group (group 7), composed of eight unclassified isolates, clearly diverges in its functional content from the

**Table 1.** Differential phenotypic characteristics of strain F113<sup>T</sup> and their closest species type strains

Strains: 1, *P. ogarae* F113<sup>T</sup>; 2, *P. ogarae* RDP1; 3, *P. kilonensis* DSM 13647<sup>T</sup>; 4, *P. brassicacearum* DSM 13227<sup>T</sup>; 5, *Pseudomonas* sp. WCS365, 6, *P. brassicacearum* NFM421; 7, *P. fluorescens* Q8r1-96; 8, *P. thivervalensis* DSM 13194<sup>T</sup>.

Characteristic	1	2	3	4	5	6	7	8
<b>Assimilation of (Biolog):</b>								
N-acetyl-glucosamine	-	-	-	+	-	+	+	-
D-fructose	-	-	+	+	-	+	+	+
Inosine	+	+	+	+	-	+	+	-
D-arabitol	+	+	+	+	-	+	-	-
D-aspartic acid	+	+	-	+	+	+	+	+
L-arginine	+	+	+	+	+	+	+	-
L-histidine	-	+	+	+	+	+	-	-
L-serine	-	-	-	+	-	+	-	-
Pectin	+	+	+	-	-	+	-	-
D-galacturonic acid	-	+	-	+	+	+	+	-
L-galacturonic acid lactone	-	-	-	+	+	+	+	-
D-glucuronic acid	-	+	-	+	+	+	+	-
Tween 40	+	+	-	+	-	+	+	+
Acetic acid	+	+	+	+	-	-	-	-
<b>Sensitivity to (Biolog):</b>								
pH 5	-	-	-	-	-	+	+	-
1% Sodium lactate	-	-	-	+	-	+	+	+
Troleandomycin	-	-	-	-	-	-	-	+
Guanidine HCl	-	-	+	-	-	-	-	+
Niaproof 4	+	-	-	-	-	+	+	-
Potassium tellurite	+	+	+	+	-	+	+	+
Aztreonam	-	-	+	-	-	+	-	+
<b>Oxidation of (API 50 CH):</b>								
Glycerol	+	+	w <sup>a</sup>	+ <sup>a</sup>	+	+	+	+
D-arabinose	w	w	-	-	-	-	-	-
D-ribose	+	w	w <sup>a</sup>	w <sup>a</sup>	+	+	w	w <sup>a</sup>
L-xylose	w	-	-	-	-	-	-	-
Inositol	+	+	w <sup>*</sup>	w <sup>*</sup>	+	+	w	+
D-mannitol	+	+	w <sup>*</sup>	+	+	+	+	+
D-sorbitol	+	+	w <sup>*</sup>	+	+	+	+	+
D-melibiose	+	+	w <sup>*</sup>	-	+	+	w	+ <sup>*</sup>
D-saccharose (sucrose)	+	+	w <sup>*</sup>	- <sup>*</sup>	+	+	+	+
D-trehalose	+	+	+	w <sup>*</sup>	+	+	+	+
Gentiobiose	w	w	-	-	-	-	-	w
D-lyxose	+	-	-	-	+	-	-	-

Continued

Table 1. Continued

Characteristic	1	2	3	4	5	6	7	8
D-arabitol	+	+	-	w	+	w	w	+
<b>API 20 NE tests:</b>								
Reduction of nitrates	+	+	w*	-	+	w	w	-
Arginine dihydrolase	-	-	-	+	+	+	+	+
Gelatin hydrolysis (protease)	+	+	+	+	-	+	+	+
N-acetyl-glucosamine assimilation	-	-	-	+	+	+	+	+

For assimilation and oxidation of carbon compounds: positive (+), negative (-), weakly positive (w). For sensitivity: insensitive (+, can grow in its presence), sensitive (-, cannot grow in its presence).

Result differs from the published by Sikorsky et al. [67].

rest of the observation groups (Fig. 4), which suggests the presence of specialized and distinctive traits.

Statistical analysis of the first two PCs' centroids demonstrates that most PCA observation groups are significantly different (File S6). Specifically, *P. ogarae* (group 26), is significantly different at the functional level compared to their closest species: *P. thivervalensis*, *P. kilonensis* and *P. brassicacearum* (groups 19, 27 and 29, respectively), also significantly different among them, which substantiate the differences found at the genomic level and further support their status as different species. Interestingly, the closely related species *P. mediterranea* and *P. corrugata* (groups 1 and 5, respectively) were not found to be statistically different at the functional level among the first two PCs (File S6). This could be related to subtle differences that are not represented by the functional categories assessed. It is also important to notice that certain genomes were discarded from the functional analysis given that they contained more than 75 contigs. Therefore, the groups represented in the PCA only account for a fraction of genomes that might not be enough to reflect their total diversity. In addition, only the first two PCs were used for the statistical analysis, which accounts for 43% of the explained variance.

Regarding the specific BRITE functional categories that explain the variance of the observation groups, in the case of PGPR species, transporters and transcription factors are the variables that contribute the most to the observed variance, followed by protein kinases and secretion systems (Fig. 4c). In contrast, the variance of *P. corrugata* and *P. mediterranea* species are driven by two-component systems, bacterial toxins, and amino acid-related enzymes variables, among others. In the specific case of *P. ogarae* (group 26), prokaryotic defence systems, protein kinases and glycosyltransferases positively correlate with this group, while antimicrobial resistance genes and secretion systems contribute to *P. kilonensis* species group.

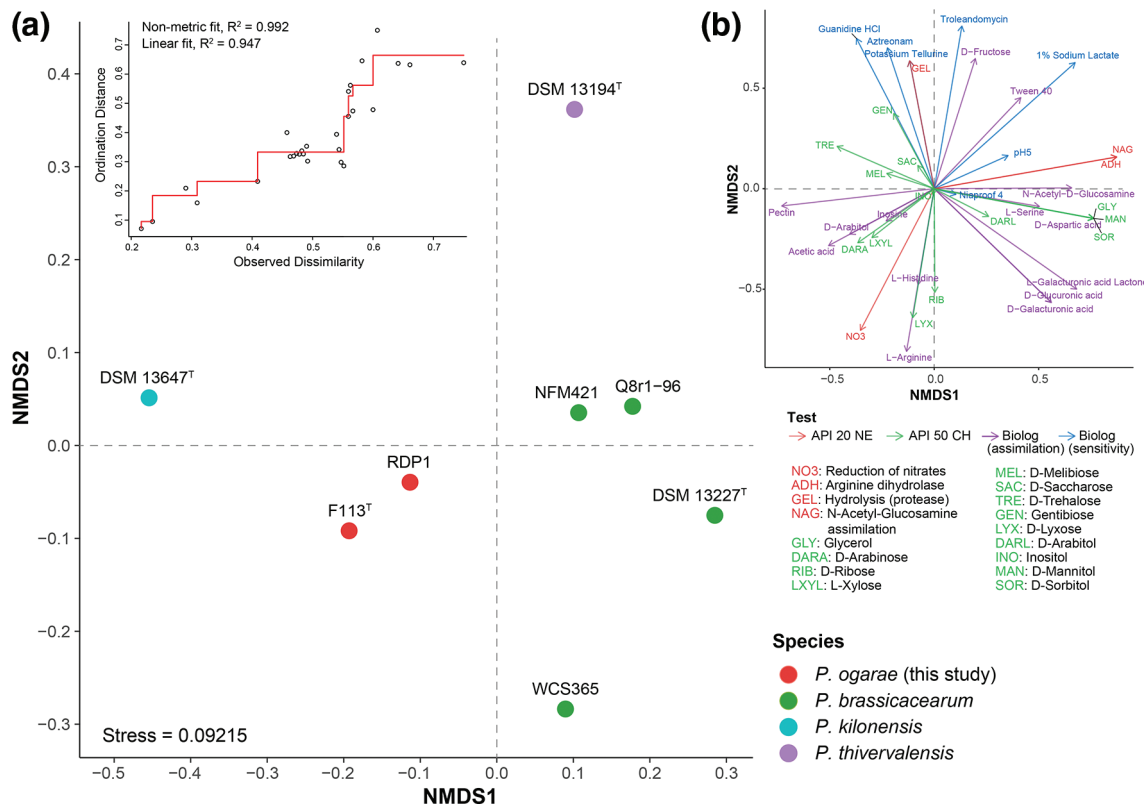
In addition, the search for genes involved in PGPR traits and the LPQ genetic island responsible for the pathogenicity of certain species and strains from the *P. corrugata* subgroup

[37] shows an interesting distribution pattern. As shown in File S7 (Fig. S1), 49 genomes harbour the LPQ island. These include all genomes from *P. mediterranea*, *P. corrugata* and species clusters numbers 2, 3, 4, 6, 7, 14, 15, 16, 20, 21, 22, 25 and 28, and it is partly present in genomes from the species-cluster 23 (clusters numbers according to Fig. 1). Conversely, this LPQ island is absent from *P. brassicacearum*, *P. kilonensis*, *P. ogarae* and *P. thivervalensis* species clusters (File S7). However, certain PGPR traits, such as the biosynthesis of the antibiotic DAPG, ACC deaminase, biosynthesis of hydrogen cyanide and genes for denitrification, among others, are found in genomes harbouring the LPQ island and those in which this island is absent. This observation was not unexpected, as it has been previously reported [27].

### Functional differences of *P. brassicacearum*, *P. kilonensis* and *P. ogarae* type strains

Type strains are defined as the nomenclatural type of a new taxon used for description, as defined by the Bacteriological Code [73], that 'exhibit all of the relevant phenotypic and genotypic properties cited in the original published taxonomic circumscriptions' [74]. Therefore, we used the type strain genomes of *P. brassicacearum* CCUG 51508<sup>T</sup> [66] and *P. kilonensis* DSM 13647<sup>T</sup> [67], together with *P. ogarae* F113<sup>T</sup> [68], the most studied strain of this novel species, to identify differences in their functional content based on the identification of orthologous sequences and BRITE classification of proteins. The results are summarized in Fig. 5 and discussed below.

The three genomes analysed, type strains of *P. kilonensis*, *P. brassicacearum* and *P. ogarae* (this study), share 4629 proteins, representing between 81–86.9% of their respective total protein content. Most of these shared proteins are classified in enzymes and transporters BRITE categories. However, there are more proteins common to *P. ogarae* and *P. brassicacearum* (427) than any other combination (200 between *P. ogarae* and *P. kilonensis* and 190 between *P. kilonensis* and *P. brassicacearum*). This higher functional similarity between *P. ogarae* and *P. brassicacearum* was previously observed in the



**Fig. 6.** Non-metric multidimensional scaling (NMDS) ordination analysis of the distinctive phenotypic characters. (a) Scattered NMDS plot across NMDS1 and NMDS2 axes. Colours according to the species groups indicated in the legend. Shepard plot (upper left) indicating the statistics for goodness of fit between ordination distances and observed dissimilarity. (b) Intrinsic vectors driving the distribution pattern. Colour according to the test they belong to.

PCA (Fig. 4a) and contrasts with the phylogenomic analysis, in which *P. kilonensis* and *P. ogarae* species groups are more closely related to one another than to *P. brassicacearum* (Fig. 1). The 427 proteins shared between *P. ogarae* and *P. brassicacearum* are classified into 22 distinct categories, and are mostly enzymes, transporters and transcription factors and, to a less extent, lipopolysaccharide biosynthetic proteins, bacterial toxins, glycosyl-transferases, peptidoglycan biosynthesis proteins and glycosaminoglycans binding proteins, which are not shared by the other genome combinations (Fig. 4b). This suggests that both *P. ogarae* and *P. brassicacearum* might form similar membrane and extracellular matrix components, which are crucial to form biofilms and to colonize the rhizosphere [75, 76]. Similarly, *P. brassicacearum* and *P. kilonensis* share 190 proteins classified into 14 different categories, mostly enzymes, transporters and transcription factors, but both also share certain categories not found in any other combination of genomes, including lipid biosynthesis proteins, antimicrobial resistance genes, phrenyltransferases and protein kinases (Fig. 4b). Finally, *P. ogarae* and *P. kilonensis* share 200 proteins classified into 12 different categories, being enzymes, transporters and secretion systems mainly represented. The secretion-system category is the most abundant one within the shared fraction of both *P. ogarae* and *P. kilonensis* genomes, being clearly absent or very reduced

within the other two genome comparisons (Fig. 4b). Secretion systems are complex and diverse key elements used by bacteria in a variety of processes [77]. In plant-associated bacteria, type-III and type-VI secretion systems (T3SS and T6SS, respectively) can function in interbacterial competition to suppress phytopathogens [78, 79]. Specifically, *P. ogarae* and *P. kilonensis* share the T3SS-I and the T6SS-III secretion systems (File S8) previously reported in F113 [80]. The co-occurrence of secretion systems in both *P. ogarae* and *P. kilonensis* might indicate a similar niche exploitation, which differs from that of *P. brassicacearum*.

Regarding the protein fraction specific to each genome (i.e. not found in the other two genomes), *P. ogarae* contains 417 proteins, 306 in *P. kilonensis* and 463 in *P. brassicacearum*, representing 7.36, 5.8 and 8.1% of their total protein content, respectively. Among the functional categories of the proteins specific to *P. ogarae*, bacterial motility is the predominant trait of this genome and is supported by the presence of an additional flagellar system (File S8) previously reported to be involved in effective plant root colonization [81]. In the case of *P. kilonensis*, among the different specific proteins, we identified proteins for cellulose biosynthesis (BcsABCGF), several polyketide biosynthesis proteins and a type-II toxin-antitoxin system (File S8).

Finally, among the specific proteins of *P. brassicacearum*, enzymes, transporters and transcriptional factors are the main distinctive categories.

### Phenotypic differences between *P. ogarae*, *P. brassicacearum*, *P. kilonensis* and *P. thivervalensis*

Phenotypic differences between *P. ogarae* F113<sup>T</sup> and RDP1, *P. kilonensis* DSM 13647<sup>T</sup>, *P. brassicacearum* DSM 13227<sup>T</sup> together with isolates WCS365, NFM421 and Q8r1-96, and *P. thivervalensis* DSM 13194<sup>T</sup>, were evaluated using Biolog GEN III, API 20 NE and API 50 CH tests. The results are summarized in Table 1 and extended information can be found in File S9. Among the total of 163 tests performed (94 Biolog, 20 API 20 NE and 49 API 50 CH), only 38 (23.3%) showed differences among the strains tested. This was not unexpected given that all strains belong to closely related species.

Among the different phenotypic features, *P. ogarae* F113<sup>T</sup> and RDP1 are the only ones that cannot use D-fructose as the carbon source while all the remaining type strains can. Similarly, they both can weakly use D-arabinose, while none of the other strains tested can. There are also phenotypic features that distinguish *P. ogarae* from the rest of the type strains. For instance, there are six phenotypic traits in which *P. ogarae* and *P. kilonensis* differ: assimilation of D-aspartic acid, Tween 40, sensitivity to Guanidine HCl and Aztreonam and oxidation of gentiobiose and D-arabitol (Table 1). Compared to *P. brassicacearum* type strain, there are also six differential phenotypic features: assimilation of N-acetyl-glucosamine, pectin, L-galacturonic acid lactone, sensitivity to 1% sodium lactate, oxidation of D-melibiose and gentiobiose. Finally, compared to *P. thivervalensis* type strain, there are 11 phenotypic traits in which they differ: assimilation of inosine, D-arabitol, L-arginine, pectin, acetic acid, sensitivity to 1% sodium lactate, troleandomycin, guanidine HCl and aztreonam and activity of arginine dihydrolase. Most of the results obtained in our study agree with the ones previously published for *P. kilonensis*, *P. brassicacearum* and *P. thivervalensis* [67] based on the oxidation of carbon compounds and the enzymatic tests of API strips. However, we found that *P. brassicacearum* type strain can oxidate glycerol and D-sucrose, and *P. thivervalensis* type strain can oxidize D-melibiose, conversely to the results previously reported [67].

In addition, a non-metric multidimensional scaling (NMDS) ordination analysis was performed using the differential phenotypic traits among the strains analysed to visualize if the observed differences correspond to the species groups. As shown in Fig. 6, phenotypic traits clearly distinguish *P. kilonensis*, *P. thivervalensis*, *P. brassicacearum* and *P. ogarae* strains in separated observation groups. The variables that have the most effect in the distinction of *P. ogarae* are their ability to use acetic acid, D-arabitol and inosine as the carbon source, oxidation of D-arabinose and L-xylose and reduction of nitrates.

### Description of *Pseudomonas ogarae* sp. nov.

*Pseudomonas ogarae* (o.ga.rae. N.L. gen. n. *ogarae*, after Fergal O’Gara, Irish microbiologist who isolated the strain F113<sup>T</sup> and defined its first plant growth-promoting features).

The cells of this species are aerobic, Gram-negative, non-spore-forming, motile by polar flagella and rod-shaped, being 0.5–1 µm wide and 1.5–5 µm in length. In minimal sucrose asparagine (SA) medium [82], colonies are yellowish with smooth and regular margins. After 24 h at 28 °C, they produce a yellow diffusible pigment giving a light yellow-green fluorescence when irradiated with UV light (254 nm). Results obtained with Biolog GEN III microplates indicate that cells can utilize as carbon and energy source the following substrates: D-trehalose, sucrose, α-D-glucose, D-mannose, D-galactose, inosine, D-sorbitol, D-mannitol, D-arabitol, myo-inositol, glycerol, D-aspartic acid, L-alanine, L-arginine, L-aspartic acid, L-glutamic acid, L-pyroglytamic acid, pectin, D-gluconic acid, mucic acid, quinic acid, D-saccharic acid, p-hydroxy-phenylacetic acid, methyl pyruvate, L-lactic acid, citric acid, α-keto-glutaric acid, L-malic acid, bromo-succinic acid, Tween 40, γ-amino-butyric acid, β-hydroxy-D,L-butyric acid and acetic acid. All other substrates included in Biolog GEN III were not utilized. Based on Biolog GEN III sensitivity assays, cells of this species can grow at pH 6 and 1% NaCl and resist rifamycin SV, lincomycin, niaproof 4, vancomycin and potassium tellurite. All other compounds included in Biolog GEN III sensitivity assays inhibited cell growth. Results obtained with API 50 CH strips indicate that cells oxidize the following substrates: glycerol, L-arabinose, D-ribose, D-xylose, D-galactose, D-glucose, D-fructose, D-mannose, inositol, D-mannitol, D-sorbitol, aesculin ferric citrate, D-melibiose, D-sucrose, D-trehalose, D-lyxose, D-fucose and D-arabitol, and weakly D-arabinose, D-xylose and gentiobiose. All other substrates included in the API 50 CH strips were not utilized. Results obtained with API 20 NE strips indicate that cells can utilize the following substrates as the carbon and energy source: glucose, arabinose, mannose, mannitol, potassium gluconate, capric acid, malate and trisodium citrate. All other substrates included in the API 20 NE strips were not utilized. In addition, the following enzyme activities were present in cells: hydrolysis (beta-glucosidase), hydrolysis (protease) and reduction of nitrates.

The type strain is strain F113<sup>T</sup> (=DSM 112162<sup>T</sup>=CECT 30235<sup>T</sup>), isolated in 1992 from the rhizosphere of sugar beet.

### CONCLUSIONS

The results presented in this work substantiate the description of *P. ogarae* as a novel species, composed of strains F113<sup>T</sup> (formerly *P. fluorescens* F113), RDP1, FR1, Ea\_RS28, Ep\_R1, ACN4, ACN7 and 1855–344. The differences compared to their closest species exist at the genomic, functional and phenotypic level. In addition, aside from the five previously described species of the *P. corrugata* subgroup,

together with the proposal of *P. ogarae* presented here, there still remain 23 clusters of genomes that likely correspond to novel species that should be properly validated. Our results also evidence that the LPQ island of pathogenicity follows a clear distribution in species clusters and is present in almost half of the genomes analysed. Nevertheless, none of the closely related *P. ogarae*, *P. kilonensis* and *P. brassicacearum* genomes harbour this island and therefore could be the preferred agents within this subgroup for biotechnological applications.

#### Funding information

This research was funded by Ministerio de Ciencia, Innovación y Universidades FEDER/EU, grant number RTI2018-093991-B-I00.

#### Author contributions

Conceptualization: D.G.-S., M.M. and R.R. Data curation: D.G.-S. Formal analysis: D.G.-S. Funding acquisition: M.M. and R.R. Investigation: D.G.-S. Methodology: D.G.-S. Project administration: M.M. and R.R. Resources: D.G.-S. Software: D.G.-S. Supervision: M.M. and R.R. Validation: D.G.-S. and M.R.-N. Visualization: D.G.-S. Writing—original draft preparation: D.G.-S. Writing—review and editing: All authors

#### Conflicts of interest

The authors declare that there are no conflicts of interest.

#### References

- Lessie T, Phibbs Jr P. Alternative pathways of carbohydrate utilization in *Pseudomonads*. *Annu Rev Microbiol* 1984;38:359–388.
- Lopes LD, Davis EW 2nd, Pereira E Silva M de C, Weisberg AJ, Bresciani L, et al. Tropical soils are a reservoir for fluorescent *Pseudomonas* spp. Biodiversity. *Environ Microbiol* 2018;20:62–74.
- Reddy GS, Matsumoto GI, Schumann P, Stackebrandt E, Shivaji S. Psychrophilic pseudomonads from Antarctica: *Pseudomonas antarctica* sp. nov., *Pseudomonas meridiana* sp. nov. and *Pseudomonas proteolytica* sp. nov. *Int J Syst Evol Microbiol* 2004;54:713–719.
- Janek T, Łukaszewicz M, Rezana T, Krasowska A. Isolation and characterization of two new lipopeptide biosurfactants produced by *Pseudomonas fluorescens* BD5 isolated from water from the Arctic Archipelago of Svalbard. *Bioresour Technol* 2010;101:6118–6123.
- Campos VL, Valenzuela C, Yarza P, Kämpfer P, Vidal R. *Pseudomonas arsenicoxydans* sp nov., an arsenite-oxidizing strain isolated from the Atacama desert. *Syst Appl Microbiol* 2010;33:193–197.
- Kristoffersen V, Rämä T, Isaksson J, Andersen JH, Gerwick WH. Characterization of rhamnolipids produced by an Arctic marine bacterium from the *Pseudomonas fluorescence* group. *Mar Drugs* 2018;16:163.
- Butaitė E, Baumgartner M, Wyder S, Kümmerli R. Siderophore cheating and cheating resistance shape competition for iron in soil and freshwater *Pseudomonas* communities. *Nat Commun* 2017;8:1–12.
- Cho JC, Tiedje JM. Biogeography and degree of endemism of fluorescent *Pseudomonas* strains in soil. *Appl Environ Microbiol* 2000;66:5448–5456.
- Cornforth DM, Dees JL, Ibberson CB, Huse HK, Mathiesen IH. *Pseudomonas aeruginosa* transcriptome during human infection. *Proceedings of the National Academy of Sciences* 2018;115:E5125–E5134.
- Hoshino T, Ishizaki K, Sakamoto T, Kumeta H, Yumoto I. Isolation of a *Pseudomonas* species from fish intestine that produces a protease active at low temperature. *Lett Appl Microbiol* 1997;25:70–72.
- Vela AI, Gutiérrez MC, Falsen E, Rollan E, Simarro I. *Pseudomonas simiae* sp. nov., isolated from clinical specimens from monkeys (*Callithrix geoffroyi*). *Int J Syst Evol Microbiol* 2006;56:2671–2676.
- Brown SD, Utturkar SM, Klingeman DM, Johnson CM, Martin SL. Twenty-one genome sequences from *Pseudomonas* species and 19 genome sequences from diverse bacteria isolated from the rhizosphere and endosphere of *Populus deltoides*. *Am Soc Microbiol* 2012.
- Xin XF, Kvitko B, SY H. *Pseudomonas syringae*: what it takes to be a pathogen. *Nat Rev Microbiol* 2018;16:316.
- Trantas EA, Licciardello G, Almeida NF, Witek K, Strano CP. Comparative genomic analysis of multiple strains of two unusual plant pathogens: *Pseudomonas corrugata* and *Pseudomonas mediterranea*. *Front Microbiol* 2015;6:811.
- Garrido-Sanz D, Meier-Kolthoff JP, Göker M, Martín M, Rivilla R, et al. Genomic and genetic diversity within the *Pseudomonas fluorescens* complex. *PLoS one* 2016;11:e0153733.
- David BV, Chandrasehar G, Selvam PN. *Pseudomonas fluorescens*: a plant-growth-promoting rhizobacterium (PGPR) with potential role in biocontrol of pests of crops. *Crop improvement through microbial biotechnology*. Elsevier, 2018, pp. 221–243.
- Yang M, Mavrodi DV, Thomashow LS, Weller DM. Differential response of wheat cultivars to *Pseudomonas brassicacearum* and take-all decline soil. *Phytopathology* 2018;108:1363–1372.
- Chen Y, Shen X, Peng H, Hu H, Wang W. Comparative genomic analysis and phenazine production of *Pseudomonas chlororaphis*, a plant growth-promoting rhizobacterium. *Genom Data* 2015;4:33–42.
- Andreolli M, Zapparoli G, Angelini E, Lucchetta G, Lampis S. *Pseudomonas protegens* MP12: A plant growth-promoting endophytic bacterium with broad-spectrum antifungal activity against grapevine phytopathogens. *Microbiol Res* 2019;219:123–131.
- Mulet M, Lalucat J, García-Valdés E. DNA sequence-based analysis of the *Pseudomonas* species. *Environ Microbiol* 2010;12:1513–1530.
- Marcelletti S, Scortichini M. Definition of plant-pathogenic *Pseudomonas* genomospecies of the *Pseudomonas syringae* complex through multiple comparative approaches. *Phytopathology* 2014;104:1274–1282.
- Peña A, Busquets A, Gomila M, Mulet M, Gomila RM. High quality draft genome sequences of *Pseudomonas fulva* DSM 17717<sup>T</sup>, *Pseudomonas parafulva* DSM 17004<sup>T</sup> and *Pseudomonas cremoricolorata* DSM 17059<sup>T</sup> type strains. *Stand Genomic Sci* 2016;11:55.
- Lalucat J, Mulet M, Gomila M, García-Valdés E. Genomics in bacterial taxonomy: impact on the genus *Pseudomonas*. *Genes (Basel)* 2020;11:139.
- Richter M, Rosselló-Móra R. Shifting the genomic gold standard for the prokaryotic species definition. *Proceedings of the National Academy of Sciences* 2009;106:19126–19131.
- Meier-Kolthoff JP, Auch AF, Klenk HP, Göker M. Genome sequence-based species delimitation with confidence intervals and improved distance functions. *BMC Bioinformatics* 2013;14:60.
- Garrido-Sanz D, Arrebola E, Martínez-Granero F, García-Mendez S, Muriel C. Classification of isolates from the *Pseudomonas fluorescens* complex into phylogenomic groups based in group-specific markers. *Front Microbiol* 2017;8:413.
- Gislason AS, de Kievit TR. Friend or foe? Exploring the fine line between *Pseudomonas brassicacearum* and phytopathogens. *J Med Microbiol* 2020;69:347–360.
- Nelkner J, Torres Tejerizo G, Hassa J, Lin TW, Witte J. Genetic potential of the biocontrol agent *Pseudomonas brassicacearum* (formerly *P. trivialis*) 3Re2-7 unraveled by genome sequencing and mining, comparative genomics and transcriptomics. *Genes (Basel)* 2019;10:601.
- Bangera MG, Thomashow LS. Identification and characterization of a gene cluster for synthesis of the polyketide antibiotic 2, 4-diacetylphloroglucinol from *Pseudomonas fluorescens* Q2-87. *J Bacteriol* 1999;181:3155–3163.
- Weller DM, Mavrodi DV, van Pelt JA, Pieterse CM, van Loon LC. Induced systemic resistance in *Arabidopsis thaliana* against *Pseudomonas syringae* pv. tomato by 2, 4-diacetylphloroglucinol-producing *Pseudomonas fluorescens*. *Phytopathology* 2012;102:403–412.

31. Brazelton JN, Pfeufer EE, Sweat TA, Gardener BBM, Coenen C. 2, 4-Diacetylphloroglucinol alters plant root development. *Mol Plant Microbe Interact* 2008;21:1349–1358.
32. Vacheron J, Combes-Meynet E, Walker V, Gouesnard B, Muller D. Expression on roots and contribution to maize phytostimulation of 1-aminocyclopropane-1-decarboxylate deaminase gene *acdS* in *Pseudomonas fluorescens* F113. *Plant and Soil* 2016;407:187–202.
33. Saravanakumar D, Samiyappan R. ACC deaminase from *Pseudomonas fluorescens* mediated saline resistance in groundnut (*Arachis hypogea*) plants. *J Appl Microbiol* 2007;102:1283–1292.
34. Sapre S, Gontia-Mishra I, Tiwari S. ACC Deaminase-producing Bacteria: a Key Player in Alleviating Abiotic Stresses in Plants. *Plant Growth Promoting Rhizobacteria for Agricultural Sustainability*. Springer, 2019, pp. 267–291.
35. Scarlett CM, Fletcher J, Roberts P, Lelliott R. Tomato pith necrosis caused by *Pseudomonas corrugata* n. sp. *Ann Appl Biol* 1978;88:105–114.
36. Catara V, Sutra L, Morineau A, Achouak W, Christen R. Phenotypic and genomic evidence for the revision of *Pseudomonas corrugata* and proposal of *Pseudomonas mediterranea* sp. nov. *Int J Syst Evol Microbiol* 2002;52:1749–1758.
37. Melnyk RA, Hossain SS, Haney CH. Convergent gain and loss of genomic islands drive lifestyle changes in plant-associated *Pseudomonas*. *ISME J* 2019;13:1575–1588.
38. Licciardello G, Bertani I, Steindler L, Bella P, Venturi V. *Pseudomonas corrugata* contains a conserved N-acyl homoserine lactone quorum sensing system; its role in tomato pathogenicity and tobacco hypersensitivity response. *FEMS Microbiol Ecol* 2007;61:222–234.
39. Catara V. *Pseudomonas corrugata*: plant pathogen and/or biological resource? *Mol Plant Pathol* 2007;8:233–244.
40. Almario J, Bruto M, Vacheron J, Prigent-Combaret C, Moënne-Loccoz Y. Distribution of 2, 4-diacetylphloroglucinol biosynthetic genes among the *Pseudomonas* spp. reveals unexpected polyphyletism. *Front Microbiol* 2017;8:1218.
41. Vesga P, Flury P, Vacheron J, Keel C, Croll D. Transcriptome plasticity underlying plant root colonization and insect invasion by *pseudomonas protegens*. *The ISME Journal* 2020;2020:1–17.
42. Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 2014;30:2114–2120.
43. Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *J Comput Biol* 2012;19:455–477.
44. Seemann T. Prokka: rapid prokaryotic genome annotation. *Bioinformatics* 2014;30:2068–2069.
45. Kumar S, Stecher G, Li M, Knyaz C, Tamura K. MEGA X: molecular evolutionary genetics analysis across computing platforms. *Mol Biol Evol* 2018;35:1547–1549.
46. Meier-Kolthoff JP, Göker M. TYGS is an automated high-throughput platform for state-of-the-art genome-based taxonomy. *Nat Commun* 2019;10:1–10.
47. Göker M, García-Blázquez G, Voglmayr H, Tellería MT, Martín MP. Molecular taxonomy of phytopathogenic fungi: A case study in *Peronospora*. *PLoS one* 2009;4:e6319.
48. Garrido-Sanz D, Redondo-Nieto M, Martín M, Rivilla R. Comparative genomics of the *Rhodococcus* genus shows wide distribution of biodegradation traits. *Microorganisms* 2020;8.
49. Emms DM, Kelly S. OrthoFinder: phylogenetic orthology inference for comparative genomics. *Genome Biol* 2019;20:1–14.
50. Buchfink B, Xie C, Huson DH. Fast and sensitive protein alignment using DIAMOND. *Nat Methods* 2015;12:59–60.
51. Enright AJ, Van Dongen S, Ouzounis CA. An efficient algorithm for large-scale detection of protein families. *Nucleic Acids Res* 2002;30:1575–1584.
52. Wickham H. ggplot2. *Wiley Interdisciplinary Reviews: Computational Statistics*. 2011;3:180–185.
53. Sievers F, Wilm A, Dineen D, Gibson TJ, Karplus K. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol Syst Biol* 2011;7:539.
54. Castresana J. Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. *Mol Biol Evol* 2000;17:540–552.
55. Talavera G, Castresana J. Improvement of phylogenies after removing divergent and ambiguously aligned blocks from protein sequence alignments. *Syst Biol* 2007;56:564–577.
56. Kozlov AM, Darriba D, Flouri T, Morel B, Stamatakis A. RAxML-NG: a fast, scalable and user-friendly tool for maximum likelihood phylogenetic inference. *Bioinformatics* 2019;35:4453–4455.
57. SQ L, Gascuel O. An improved general amino acid replacement matrix. *Mol Biol Evol* 2008;25:1307–1320.
58. Stamatakis A, Hoover P, Rougemont J. A rapid bootstrap algorithm for the RAxML web servers. *Syst Biol* 2008;57:758–771.
59. Pattengale ND, Alipour M, Bininda-Emonds OR, Moret BM, Stamatakis A. How many bootstrap replicates are necessary? *J Comput Biol* 2010;17:337–354.
60. Huerta-Cepas J, Forslund K, Coelho LP, Szklarczyk D, Jensen LJ. Fast genome-wide functional annotation through orthology assignment by eggNOG-Mapper. *Mol Biol Evol* 2017;34:2115–2122.
61. Huerta-Cepas J, Szklarczyk D, Heller D, Hernández-Plaza A, Forslund SK, et al. EggNOG 5.0: A hierarchical, functionally and phylogenetically annotated orthology resource based on 5090 organisms and 2502 viruses. *Nucleic Acids Res* 2019;47:D309–D314.
62. Tenenbaum D. KEGGREST: Client-side REST access to KEGG. R package version 1.24.0. 2019. 2019.
63. Conway JR, Lex A, Gehlenborg N. UpSetR: an R package for the visualization of intersecting sets and their properties. *Bioinformatics* 2017;33:2938–2940.
64. Goodpaster AM, Kennedy MA. Quantification and statistical significance analysis of group separation in NMR-based metabolomics studies. *Chemometr Intell Lab Syst* 2011;109:162–170.
65. Mahalanobis PC. *On the Generalized Distance in Statistics*. Calcutta: National Institute of Science of India; 1936.
66. Achouak W, Sutra L, Heulin T, Meyer JM, Fromin N, et al. *Pseudomonas brassicacearum* sp. nov. and *Pseudomonas thivervalensis* sp. nov., two root-associated bacteria isolated from *Brassica napus* and *Arabidopsis thaliana*. *Int J Syst Evol Microbiol* 2000;50 Pt 1:9–18.
67. Sikorski J, Stackebrandt E, Wackernagel W. *Pseudomonas kilonensis* sp. nov., a bacterium isolated from agricultural soil. *Int J Syst Evol Microbiol* 2001;51:1549–1555.
68. Shanahan P, O'Sullivan DJ, Simpson P, Glennon JD, O'Gara F. Isolation of 2, 4-diacetylphloroglucinol from a fluorescent pseudomonad and investigation of physiological parameters influencing its production. *Appl Environ Microbiol* 1992;58:353–358.
69. Redondo-Nieto M, Barret M, Morrissey JP, Germaine K, Martínez-Granero F. Genome sequence of the biocontrol strain *Pseudomonas fluorescens* f113. *Am Soc Microbiol* 2012.
70. Brazil G, Kenefick L, Callanan M, Haro A, Lorenzo D. Construction of a rhizosphere pseudomonad with potential to degrade polychlorinated biphenyls and detection of *bph* gene expression in the rhizosphere. *Appl Environ Microbiol* 1995;61:1946–1952.
71. Dixon P. VEGAN, a package of R functions for community ecology. *J Veg Sci* 2003;14:927–930.
72. Gomila M, Peña A, Mulet M, Lalucat J, García-Valdés E. Phylogenomics and systematics in *Pseudomonas*. *Front Microbiol* 2015;6:214.
73. Lapage S, Sneath P, Lessel E, Skerman V, Seeliger H. *International Code of Nomenclature of Bacteria: Bacteriological code, 1990 revision*. ASM Press, 1992.
74. Kyrpides NC, Hugenholtz P, Eisen JA, Woyke T, Göker M, et al. Genomic encyclopedia of bacteria and archaea: sequencing a myriad of type strains. *PLoS Biol* 2014;12:e1001920.

75. Ueda A, Saneoka H. Characterization of the ability to form biofilms by plant-associated *Pseudomonas* species. *Curr Microbiol* 2015;70:506–513.
76. Blanco-Romero E, Garrido-Sanz D, Rivilla R, Redondo-Nieto M, Martín M. *In silico* characterization and phylogenetic distribution of extracellular matrix components in the model rhizobacteria *Pseudomonas fluorescens* F113 and other *Pseudomonads*. *Microorganisms* 2020;8:1740.
77. Green ER, Mecsas J. Bacterial secretion systems: An overview. *Virulence Mechanisms of Bacterial Pathogens* 2016;2016:213–239.
78. Almarío J, Gobbin D, Défago G, Moënné-Loccoz Y, Rezzonico F. Prevalence of type III secretion system in effective biocontrol pseudomonads. *Res Microbiol* 2014;165:300–304.
79. Bernal P, Llamas MA, Filloux A. Type VI secretion systems in plant-associated bacteria. *Environ Microbiol* 2018;20:1–15.
80. Redondo-Nieto M, Barret M, Morrissey J, Germaine K, Martínez-Granero F. Genome sequence reveals that *Pseudomonas fluorescens* F113 possesses a large and diverse array of systems for rhizosphere function and host interaction. *BMC genomics* 2013;14:1–17.
81. Barahona E, Navazo A, Garrido-Sanz D, Muriel C, Martínez-Granero F. *Pseudomonas fluorescens* F113 can produce a second flagellar apparatus, which is important for plant root colonization. *Front Microbiol* 2016;7:1471.
82. Scher FM, Baker R. Effect of *Pseudomonas putida* and a synthetic iron chelator on induction of soil suppressiveness to Fusarium wilt pathogens. *Phytopathology* 1982;72:1567

#### Five reasons to publish your next article with a Microbiology Society journal

1. The Microbiology Society is a not-for-profit organization.
2. We offer fast and rigorous peer review – average time to first decision is 4–6 weeks.
3. Our journals have a global readership with subscriptions held in research institutions around the world.
4. 80% of our authors rate our submission process as 'excellent' or 'very good'.
5. Your article will be published on an interactive journal platform with advanced metrics.

Find out more and submit your article at [microbiologyresearch.org](https://microbiologyresearch.org).