

Article

The Ever-Expanding *Pseudomonas* Genus: Description of 43 New Species and Partition of the *Pseudomonas Putida* Group

Léa Girard¹⁺, Cédric Lood^{1,2+}, Monica Höfte³, Peter Vandamme⁴, Hassan Rokni-Zadeh⁵, Vera van Noort^{1,6}, Rob Lavigne^{2*}, René De Mot^{1,*}

¹ Centre of Microbial and Plant Genetics, Faculty of Bioscience Engineering, KU Leuven, Kasteelpark Arenberg 20, 3001 Leuven, Belgium; lea.girard@kuleuven.be (L.G.), cedric.lood@kuleuven.be (C.L.), vera.vannoort@kuleuven.be (V.v.N.)

² Department of Biosystems, Laboratory of Gene Technology, KU Leuven, Kasteelpark Arenberg 21, 3001 Leuven, Belgium; rob.lavigne@kuleuven.be

³ Department of Plants and Crops, Laboratory of Phytopathology, Faculty of Bioscience Engineering, Ghent University, Ghent, Belgium

⁴ Laboratory of Microbiology, Department of Biochemistry and Microbiology, Faculty of Sciences, Ghent University, K. L. Ledeganckstraat 35, 9000 Ghent, Belgium; Peter.Vandamme@UGent.be

⁵ Zanjan Pharmaceutical Biotechnology Research Center, Zanjan University of Medical Sciences, 45139-56184 Zanjan, Iran; hassan.roknizadeh@gmail.com

⁶ Institute of Biology, Leiden University, Sylviusweg 72, 2333 Leiden, The Netherlands

+ The authors contributed equally to this work.

* Correspondence: rob.lavigne@kuleuven.be, +3216379524; rene.demot@kuleuven.be ; Tel.: +3216329681

Abstract: The genus *Pseudomonas* hosts an extensive genetic diversity and is one of the largest genera among Gram-negative bacteria. Type strains of *Pseudomonas* are well-known to represent only a small fraction of this diversity and the number of available *Pseudomonas* genome sequences is increasing rapidly. Consequently, new *Pseudomonas* species are regularly reported and the number of species within the genus is in constant evolution. In this study, whole genome sequencing enabled us to define 43 new *Pseudomonas* species and to provide an update of the *Pseudomonas* evolutionary and taxonomic relationships. Phylogenies based on the *rpoD* gene and whole genome sequences, including 316 and 313 type strains of *Pseudomonas*, respectively, revealed sixteen groups of *Pseudomonas* and justified the partitioning of the *P. putida* group into fifteen subgroups. Pairwise average nucleotide identities were calculated between type strains and a selection of 60 genomes of non-type strains of *Pseudomonas*. Forty-one strains were incorrectly assigned at the species level and among those, 19 strains were shown to represent an additional 13 new *Pseudomonas* species that remain to be formally classified. This work pinpoints the importance of correct taxonomic assignment and phylogenetic classification in order to perform integrative studies linking genetic diversity, lifestyle and metabolic potential of *Pseudomonas* spp.

Keywords: *Pseudomonadaceae*, long-read sequencing, hybrid assembly, Taxonomy, Phylogeny, Non-Ribosomal Peptide Synthetases (NRPS), Cyclic Lipopeptides (CLPs).

1. Introduction

Over the last decade, the landscape of bacterial systematics has changed drastically [1]. Once dominated by a polyphasic approach including phenotypic characterization, DNA-DNA hybridization and 16S rRNA gene sequencing, the age of microbial genomics and metagenomics have reshaped the foundation of prokaryotic species definition [2,3]. While 16S rRNA phylogeny remains the most common tool to evaluate the diversity of mixed prokaryotic populations, estimating inter- and intraspecies relatedness was traditionally facilitated by DNA-typing methods. Multi-Locus Sequence Analysis (MLSA) for several years represented the most widely adopted methodology for bacterial systematics, as well as for the exploration of evolutionary relationships within specific fami-

lies/genera [4–7]. The success of high throughput and affordable Whole Genome Sequencing (WGS) technologies has increased the number of publicly available genomes tremendously and therefore genome-to-genome comparisons, with the Average Nucleotide Identity (ANI) and digital DNA–DNA Hybridization (dDDH), have become today's standards for species definition [1,8–11]. This genome-based elucidation of relatedness at the inter- and intraspecific level is now encouraged and, on larger scale, the creation of a Genome Taxonomy Database (GTDB), has allowed to standardize bacterial taxonomy [12,13].

According to GTDB the *Pseudomonadaceae* family currently includes seven genera: *Azomonas*, *Azotobacter*, *Entomomonas*, *Oblitimonas*, *Pseudomonas*, *Thiopseudomonas* and *Ventosimonas* (https://gtdb.ecogenomic.org/tree?r=f__Pseudomonadaceae). The genus *Pseudomonas* is the most complex with 254 validly named species (List of Prokaryotic Names with Standing in Nomenclature; <https://lpsn.dsmz.de/genus/pseudomonas>), excluding subspecies and synonymous species. However, this number is constantly evolving, with over 30 new *Pseudomonas* species described between March 2020 and March 2021. Since the first descriptions of *Pseudomonas* species, which were based on morphological and phenotypical characteristics, several studies updated the taxonomy of *Pseudomonas* based on 16S rRNA gene sequence analysis [14]. It allowed differentiating the genus *Pseudomonas* from its sister genera but also defining the three main *Pseudomonas* lineages, *P. pertucinogena*, *P. aeruginosa* and *P. fluorescens* [6,15]. In a similar fashion, MLSA has guided the re-definition of prokaryotic species and has also impacted the phylogenomics and systematics of the genus *Pseudomonas* [4,6,16]. Indeed, the analysis based on four housekeeping genes (i.e. 16S rRNA, *gyrB*, *rpoB* and *rpoD*), enabled the clarification of the *Pseudomonas* phylogeny by enhancing species delineation. This approach also proved to be a reliable tool for strain identification at the species level [4,6]. We recently demonstrated that the *rpoD* gene sequence alone provides a strong and low-cost alternative, particularly in the case of taxonomic affiliation of large batches of environmental *Pseudomonas* isolates [17].

Pseudomonas are motile, non-spore forming, Gram-negative rods belonging to the *Gammaproteobacteria*. *Pseudomonas* species are able to colonize and thrive in a wide range of ecological niches (e.g. soil, water, plants, associated to higher organisms) [18]. Besides the well-known human pathogen *P. aeruginosa*, other *Pseudomonas* species induce diseases in plants, fish, insects or other animals [19–21]. On the other hand, a large majority of *Pseudomonas* species are commensals but can also be used as bioremediation, biostimulation and biocontrol agents [22,23]. *Pseudomonas* are ubiquitous bacteria, often identified as fundamental component of bacterial communities and carrying essential ecological functions [24–26]. Furthermore, *Pseudomonas* are outstanding producers of bioactive secondary metabolites that often support their eclectic lifestyle (e.g. iron-scavenging, swarming motility, biofilm formation, pathogenicity, cooperation or antagonism) [27,28]. The link between secondary metabolites and *Pseudomonas* taxonomy has already been made through pyoverdines, a class of pigments, used for a long time as a specific marker of classification [18]. *Pseudomonas* cyclic lipopeptides (CLPs), with their broad antimicrobial activity profile and anti-proliferative properties have caught the attention of researchers for their promising application potential [29]. CLP production is widespread within the genus *Pseudomonas* and relationships between CLP diversity and *Pseudomonas* taxonomy were recently highlighted. Indeed, CLP producers tend to be grouped by CLP family and confined to specific groups or subgroups of *Pseudomonas*. Yet, exceptions are occurring within the *P. putida* group which hosts a large diversity of CLP producers from diverse families (i.e. Xantholysin, Entolysin, Putisolvin and Viscosin families) [30,31].

In this study, we report 43 new *Pseudomonas* species and use a combination of Nanopore and Illumina sequencing to provide high quality genomes. Through the genome analysis of these new species together with type strains of *Pseudomonas*, we provide an update of the *Pseudomonas* phylogeny based on a set of 1,508 core orthogroup sequences and another based on the *rpoD* gene. We used nucleotide identities based on the *rpoD* gene and whole genome comparisons to reassign, respectively, 128 and 41 *Pseudomonas*

strains to known and newly described *Pseudomonas* species. A large majority of the new species were affiliated to the *P. putida* group, increasing the count of species from 35 to 51. We thus explored genetic diversity within the *P. putida* group in greater depth and mapped, on an expanded phylogeny of the group, the presence of Biosynthetic Gene Clusters (BGCs) for the production of CLPs.

2. Materials and Methods

We recently highlighted the high discriminative power of the *rpoD* gene as a reliable identification method for environmental *Pseudomonas* isolates [17]. In the same study, we performed Illumina sequencing on 55 environmental *Pseudomonas* isolates and the *rpoD* gene together with whole genome comparisons allowed us to highlight the presence of 30 new *Pseudomonas* species (*Pseudomonas* #5, with strains SWRI59, SWRI68 and SWRI77, was later identified as *P. capeferrum*). The same methodology was applied to additional *Pseudomonas* strains and an extra seventeen new species were detected. To provide high-quality genomes for the type strains of these 47 new species we combined Illumina and Nanopore sequencing and performed hybrid assemblies as previously described [32]. The *rpoD* sequences were then used as query to search for additional strains belonging to these 47 new species using BlastN in Genbank (Table S1 and S2). The 47 isolates were deposited in two culture collections (i.e. Belgian Co-ordinated Collections of Micro-organisms (BCCM/LMG) and Collection Française de Bactéries associées aux Plantes (CFBP)) and their phenotypic profiles were obtained using the Biolog GEN III MicroPlate (BIOLOG, Hayward, CA) according to manufacturer's instructions (Table S3).

A total of 275 *Pseudomonas* species, nine type strains belonging to other genera within the *Pseudomonadaceae* family and *Cellvibrio japonicus* were included in this study. The list of type strains together with the accession number of their *rpoD* and whole genome sequences can be found in Table S4. As previously described, a similarity matrix, based on a 650 bp fragment of the *rpoD* gene, including all type strains of *Pseudomonas* (known and newly described species) was generated (Table S5) and Average Nucleotide Identities (ANI) values were calculated using pyani v0.2.10 with default parameters and the ANIb method (Table S6) [17]. In the case of ANIb values considered as ambiguous (i.e. between 95 and 96.5%) we calculated digital DNA-DNA Hybridization (dDDH) using the online tool Genome-to-Genome Distance Calculator (GGDC). An overview of the different groups and subgroups is shown Table 1.

The evolutionary relationships between newly described and previously known *Pseudomonas* species, have been assessed using *rpoD* and whole genome phylogenies. The *rpoD*-based phylogenies were conducted as previously described (Figure 1, right) [17]. The phylogenetic tree based on whole genomes was inferred with IQ-TREE v1.6.12 (automatic model selection and 1,000 ultrafast bootstrap) using 1508 core orthogroup sequences alignment that were delineated with the SCARAP pipeline (Figure 1, left) [33].

The *P. putida* group was previously highlighted to include CLP producers from the Viscosin (WLIP producers), Putisolvin, Entolysin and Xantholysin families [30,31]. Among the 16 type strains of the newly described species within the *P. putida* group, 4 were already described as CLP producers (WLIP and Xantholysin producers) [30,31]. Consequently, all strains belonging to the *P. putida* group, with available genome sequences, were subjected to an antiSMASH analysis (antiSMASH 6.0) [34].

Table 1. Newly proposed and emended groups of *Pseudomonas*. Newly described *Pseudomonas* species are highlighted in bold.

Groups/Subgroups	Species/Subspecies	Total No. of species/subspecies	<i>rpoD</i> identity (%) ¹	ANI (%) ¹
Existing groups				
<i>P. aeruginosa</i> G	<i>P. aeruginosa</i> , <i>P. citronellolis</i> , <i>P. delhiensis</i> , <i>P. humi</i> , <i>P. jinjuensis</i> , <i>P. knackmussii</i> , <i>P. nicosulfuronedens</i> , <i>P. nitritireducens</i> , <i>P. nitroreducens</i> , <i>P. panipatensis</i>	10	77.85 - 95.98	80.57 - 94.48
<i>P. anguilliseptica</i> G	<i>P. anguilliseptica</i> , <i>P. benzenivorans</i> , <i>P. borbori</i> , <i>P. cutrocieneegasensis</i> , <i>P. guineae</i> , <i>P. leptonychotis</i> , <i>P. marincola</i> , <i>P. peli</i> , <i>P. segetis</i> , <i>P. taeanensis</i>	10	72.33 - 91.58	76.68 - 89.45
<i>P. fluorescens</i> G		134	-	-
<i>P. asplenii</i> SG	<i>P. agarici</i> , <i>P. asplenii</i> , <i>P. batumici</i> , <i>P. fuscovaginae</i> , <i>P. vanderleydeniana</i>	5	84.41 - 89.66	84.10 - 88.35
<i>P. chlororaphis</i> SG	<i>P. chlororaphis</i> subsp. <i>aurantiaca</i> , <i>P. chlororaphis</i> subsp. <i>aureofaciens</i> , <i>P. chlororaphis</i> subsp. <i>chlororaphis</i> , <i>P. chlororaphis</i> subsp. <i>piscium</i>	4	97.83 - 98.45	94.73 - 96.95
<i>P. corrugata</i> SG	<i>P. alvandrae</i> , <i>P. beijiensis</i> , <i>P. brassicacearum</i> , <i>P. corrugata</i> , <i>P. kilonensis</i> , <i>P. marodashtae</i> , <i>P. mediterranea</i> , <i>P. tehranensis</i> , <i>P. thivervalensis</i> , <i>P. viciae</i> , <i>P. zanjaniensis</i> , <i>P. zarinae</i>	12	89.75 - 97.36	85.55 - 95.75
<i>P. fluorescens</i> SG	<i>P. allii</i> , <i>P. antarctica</i> , <i>P. asgharzadehiana</i> , <i>P. aylmerense</i> , <i>P. azadiae</i> , <i>P. azotoformans</i> , <i>P. canadensis</i> , <i>P. carnis</i> , <i>P. cedrina</i> subsp. <i>cedrina</i> , <i>P. cedrina</i> subsp. <i>fulgida</i> , <i>P. constantinii</i> , <i>P. cremoris</i> , <i>P. cyclaminis</i> , <i>P. edaphica</i> , <i>P. extremaustralis</i> , <i>P. extremorientalis</i> , <i>P. fildesensis</i> , <i>P. fluorescens</i> , <i>P. grimontii</i> , <i>P. haemolytica</i> , <i>P. kairouanensis</i> , <i>P. karstica</i> , <i>P. khavaziana</i> , <i>P. kunguensis</i> , <i>P. lactis</i> , <i>P. libanensis</i> , <i>P. lurida</i> , <i>P. marginalis</i> , <i>P. nabeulensis</i> , <i>P. orientalis</i> , <i>P. palleroniana</i> , <i>P. panacis</i> , <i>P. paracarnis</i> , <i>P. paralactis</i> , <i>P. pisciculturae</i> , <i>P. poae</i> , <i>P. rhodesiae</i> , <i>P. salmasensis</i> , <i>P. salomonii</i> , <i>P. simiae</i> , <i>P. sivasensis</i> , <i>P. spelaei</i> , <i>P. synxantha</i> , <i>P. tolaasii</i> , <i>P. tritici</i> , <i>P. trivialis</i> , <i>P. veronii</i> , <i>P. yamanorum</i>	48	85.16 - 98.91	83.52 - 95.68
<i>P. fragi</i> SG	<i>P. bubulae</i> , <i>P. deceptionensis</i> , <i>P. endophytica</i> , <i>P. fragi</i> , <i>P. helleri</i> , <i>P. lundensis</i> , <i>P. psychrophila</i> , <i>P. saxonica</i> , <i>P. taetrolens</i> , <i>P. versuta</i> , <i>P. weihenstephanensis</i>	11	83.00 - 97.67	80.50 - 90.49
<i>P. gessardii</i> SG	<i>P. brennerii</i> , <i>P. gessardii</i> , <i>P. mucidolens</i> , <i>P. proteolytica</i> , <i>P. shahriarae</i>	5	90.57 - 97.53	85.57 - 92.54
<i>P. jessenii</i> SG	<i>P. asgharzadehiana</i> , <i>P. azerbaijanoccidens</i> , <i>P. izuensis</i> , <i>P. jessenii</i> , <i>P. laurylsulfatiphila</i> , <i>P. laurylsulfatovorans</i> , <i>P. mohnii</i> , <i>P. moorei</i> , <i>P. reinekei</i> , <i>P. umsongensis</i> , <i>P. vancouverensis</i>	11	90.37 - 100	84.91 - 95.51

<i>P. koreensis</i> SG	<i>P. atacamensis</i> , <i>P. atagosis</i> , <i>P. baetica</i> , <i>P. banana-midigenes</i> , <i>P. botevensis</i> , <i>P. crudilactis</i> , <i>P. ekonensis</i> , <i>P. glycinae</i> , <i>P. granadensis</i> , <i>P. hamedanensis</i> , <i>P. helmanticensis</i> , <i>P. iranensis</i> , <i>P. khorasanensis</i> , <i>P. koreensis</i> , <i>P. kribbensis</i> , <i>P. monsensis</i> , <i>P. moraviensis</i> , <i>P. neuropathica</i> , <i>P. siliginis</i> , <i>P. tensinigenes</i> , <i>P. triticicola</i> , <i>P. zeae</i>	22	85.40 - 99.53	82.48 - 96.09
<i>P. mandelii</i> SG	<i>P. arsenicoxydans</i> , <i>P. farris</i> , <i>P. frederiksbergensis</i> , <i>P. gregormendelii</i> , <i>P. lini</i> , <i>P. mandelii</i> , <i>P. migulae</i> , <i>P. mucoides</i> , <i>P. piscium</i> , <i>P. prosekii</i> , <i>P. silesiensis</i>	11	91.04 - 96.89	84.68 - 94.29
<i>P. protegens</i> SG	<i>P. aestus</i> , <i>P. protegens</i> , <i>P. saponiphila</i> , <i>P. sessilinigenes</i>	4	89.52 - 95.57	86.41 - 91.86
<i>P. kielensis</i> SG	<i>P. kielensis</i>	1	-	-
<i>P. linyingensis</i> G	<i>P. guangdongensis</i> , <i>P. linyingensis</i> , <i>P. oryzae</i> , <i>P. sagittaria</i>	4	79.94 - 93.85	85.19 - 92.01
<i>P. lutea</i> G	<i>P. abietaniphila</i> , <i>P. bohémica</i> , <i>P. graminis</i> , <i>P. lutea</i>	4	83.31 - 88.30	81.89 - 85.81
<i>P. oleovorans</i> G	<i>P. alcaliphila</i> , <i>P. chaetoceroeae</i> , <i>P. chengduensis</i> , <i>P. composti</i> , <i>P. guguanensis</i> , <i>P. hydrolytica</i> , <i>P. indoloxydans</i> , <i>P. khazarica</i> , <i>P. mendocina</i> , <i>P. oleovorans</i> , <i>P. pseudoalcaligenes</i> , <i>P. sediminis</i> , <i>P. sihuisensis</i> , <i>P. toyotomiensis</i>	14	88.51 - 98.76	86.06 - 95.79
<i>P. oryzihabitans</i> G	<i>P. asuensis</i> , <i>P. duriflava</i> , <i>P. luteola</i> , <i>P. oryzihabitans</i> , <i>P. psychrotolerans</i> , <i>P. rhizoryzae</i> , <i>P. zeshuii</i>	7	66.46 - 94.18	73.64 - 88.62
<i>P. pertucinogena</i> G	<i>P. abyssi</i> , <i>P. aestusnigri</i> , <i>P. bauzanensis</i> , <i>P. formosensis</i> , <i>P. gallaeciensis</i> , <i>P. jilinensis</i> , <i>P. litoralis</i> , <i>P. oceani</i> , <i>P. pachastrellae</i> , <i>P. pelagia</i> , <i>P. pertucinogena</i> , <i>P. phragmitis</i> , <i>P. populi</i> , <i>P. profundus</i> , <i>P. sabulinigri</i> , <i>P. salegens</i> , <i>P. salina</i> , <i>P. saliphila</i> , <i>P. saudimassiliensis</i> , <i>P. xiamenensis</i> , <i>P. xinjiangensis</i> , <i>P. yangmingensis</i>	22	64.53 - 92.98	74.65 - 89.65
<i>P. putida</i> G		51		
<i>P. akappagea</i> SG	<i>P. akappagea</i>	1	-	-
<i>P. japonica</i> SG	<i>P. brassicae</i> , <i>P. defluvii</i> , <i>P. huaxiensis</i> , <i>P. japonica</i> , <i>P. laurentiana</i> , <i>P. qingdaonensis</i>	6	82.69 - 95.05	80.96 - 91.58
<i>P. vranovensis</i> SG	<i>P. alkylphenolica</i> , <i>P. donghuensis</i> , <i>P. hutmensis</i> , <i>P. tructae</i> , <i>P. vranovensis</i> , <i>P. wadenswillerensis</i>	6	84.23 - 94.10	84.52 - 93.03
<i>P. cremoricolorata</i> SG	<i>P. cremoricolorata</i>	1	-	-
<i>P. reidholzensis</i> SG	<i>P. reidholzensis</i> , <i>P. shirazensis</i> , <i>P. urmiensis</i>	3	85.78 - 92.89	84.27 - 86.77
<i>P. guariconensis</i> SG	<i>P. guariconensis</i>	1	-	-
<i>P. wayambapalatensis</i> SG	<i>P. wayambapalatensis</i>	1	-	-
<i>P. farsensis</i> SG	<i>P. farsensis</i>	1	-	-
<i>P. xanthosomae</i> SG	<i>P. fakonensis</i> , <i>P. xanthosomae</i>	2	97.84	95.06

<i>P. mosselii</i> SG	<i>P. entomophila</i> , <i>P. maumuensis</i> , <i>P. mosselii</i> , <i>P. muyukensis</i> , <i>P. peradeniyensis</i> , <i>P. sichuanensis</i> , <i>P. soli</i> , <i>P. xantholysinigenes</i>	8	87.48 - 95.35	87.35 - 94.87
<i>P. taiwanensis</i> SG	<i>P. taiwanensis</i>	1	-	-
<i>P. plecoglossicida</i> SG	<i>P. plecoglossicida</i>	1	-	-
<i>P. vlassakiae</i> SG	<i>P. hunanensis</i> , <i>P. promysalinigenes</i> , <i>P. vlassakiae</i>	3	89.34 - 93.04	86.58
<i>P. capeferrum</i> SG	<i>P. capeferrum</i> , <i>P. kermanshahensis</i>	2	93.01	90.26
<i>P. putida</i> SG	<i>P. alloputida</i> , <i>P. anuradhapurensis</i> , <i>P. asiatica</i> , <i>P. fulva</i> , <i>P. inefficax</i> , <i>P. juntendi</i> , <i>P. kurunegalensis</i> , <i>P. monteili</i> , <i>P. oryzicola</i> , <i>P. parafulva</i> , <i>P. putida</i> , <i>P. pyomelaminifaciens</i> , <i>P. persica</i> , <i>P. shirazica</i>	14	85.32 - 97.99	82.66 - 95.79
<i>P. resinovorans</i> G	<i>P. furukawai</i> , <i>P. lalkuanensis</i> , <i>P. mangiferae</i> , <i>P. otitidis</i> , <i>P. resinovorans</i>	5	81.47 - 90.46	79.25 - 87.84
<i>P. straminea</i> G	<i>P. argentinensis</i> , <i>P. daroniae</i> , <i>P. dryadis</i> , <i>P. flavescens</i> , <i>P. punonensis</i> , <i>P. seleniipraecipitans</i> , <i>P. straminea</i>	7	86.02 - 93.63	82.83 - 88.54
<i>P. stutzeri</i> G	<i>P. azotofigens</i> , <i>P. balearica</i> , <i>P. chloritidismutans</i> , <i>P. kirkiae</i> , <i>P. kunmingensis</i> , <i>P. nitrititolerans</i> , <i>P. nosocomialis</i> , <i>P. perfectomarina</i> , <i>P. saudiphocaensis</i> , <i>P. songnenensis</i> , <i>P. stutzeri</i> , <i>P. urumqiensis</i> , <i>P. xanthomarina</i> , <i>P. zhaodongensis</i>	14	73.60 - 89.69	76.39 - 88.15
<i>P. syringae</i> G	<i>P. amygdali</i> , <i>P. asturieensis</i> , <i>P. avellanae</i> , <i>P. cannabina</i> , <i>P. caricapapayae</i> , <i>P. caspiana</i> , <i>P. cerasi</i> , <i>P. cichorii</i> , <i>P. congelans</i> , <i>P. coronafaciens</i> , <i>P. ficuserectae</i> , <i>P. floridensis</i> , <i>P. meliae</i> , <i>P. ovata</i> , <i>P. savastanoi</i> , <i>P. syringae</i> , <i>P. tremae</i> , <i>P. viridiflava</i>	18	78.36 - 99.54	78.20 - 94.57
Newly described groups				
<i>P. pohangensis</i> G	<i>P. mangrovi</i> , <i>P. pohangensis</i>	2	67.02	77.07
<i>P. massiliensis</i> G	<i>P. massiliensis</i> , <i>P. typographi</i>	2	80.53	76.86
<i>P. rhizosphaerae</i> G	<i>P. baltica</i> , <i>P. coleopterorum</i> , <i>P. rhizosphaerae</i>	3	91.89 - 94.70	88.28 - 90.55
Orphan groups				
<i>P. indica</i> G	<i>P. indica</i>	1	-	-
<i>P. kuykendallii</i> G	<i>P. kuykendallii</i>	1	-	-
<i>P. thermotolerans</i> G	<i>P. thermotolerans</i>	1	-	-
<i>P. flexibilis</i> G	<i>P. flexibilis</i> , <i>P. tuomuerensis</i>	2	-	-
<i>P. fluvialis</i> G	<i>P. fluvialis</i> , <i>P. pharmacofabricae</i>	2	-	-
<i>P. alcaligenes</i> G	<i>P. alcaligenes</i>	1	-	-
<i>P. matsuisoli</i> G	<i>P. matsuisoli</i>	1	-	-
<i>P. turukhanskensis</i> G	<i>P. turukhanskensis</i>	1	-	-

¹ Excluding synonymous species (Table 2).

Table 2. Synonymous species of *Pseudomonas* (Table S6). Species are considered synonymous when ANIb values are greater than or equal to 96.5% [18].

Groups/Subgroups	<i>Pseudomonas</i> species		ANIb	Earlier synonyms
<i>P. aeruginosa</i> group	<i>P. citronellolis</i>	<i>P. humi</i>	96.70	<i>P. citronellolis</i>
	<i>P. nitroreducens</i>	<i>P. nitritireducens</i>	98.85	<i>P. nitroreducens</i>
<i>P. oleovorans</i> group	<i>P. oleovorans</i>	<i>P. pseudoalcaligenes</i>	97.17	<i>P. oleovorans</i>
	<i>P. chengduensis</i>	<i>P. sihuiensis</i>	96.25	<i>P. chengduensis</i>
<i>P. oryzihabitans</i> group	<i>P. oryzihabitans</i>	<i>P. psychrotolerans</i>	98.22	<i>P. oryzihabitans</i>
	<i>P. luteola</i>	<i>P. zeshuii</i>	97.87	<i>P. luteola</i>
<i>P. pertucinogena</i> group	<i>P. phragmitis</i>	<i>P. jilinensis</i>	98.70	<i>P. phragmitis</i>
	<i>P. gallaeciensis</i>	<i>P. abyssi</i>	97.56	<i>P. gallaeciensis</i>
<i>P. putida</i> group	<i>P. asiatica</i>	<i>P. pyomelaninifaciens</i>	99.03	<i>P. asiatica</i>
		<i>P. shirazica</i>	99.17	
<i>P. stutzeri</i> group	<i>P. chloritidismutans</i>	<i>P. kunmingensis</i>	96.49	<i>P. chloritidismutans</i>
<i>P. syringae</i> group	<i>P. tremae</i>	<i>P. coronafaciens</i>	98.74	<i>P. tremae</i>
	<i>P. amygdali</i>	<i>P. ficuserectae</i>	97.42	<i>P. amygdali</i>
		<i>P. meliae</i>	98.27	
		<i>P. savastanoi</i>	98.75	
<i>P. fluorescens</i> group	<i>P. asplenii</i>	<i>P. fuscovaginae</i>	98.23	<i>P. asplenii</i>
	<i>P. veronii</i>	<i>P. panacis</i>	99.95	<i>P. veronii</i>
Orphan groups	<i>P. flexibilis</i>	<i>P. tuomuerensis</i>	98.69	<i>P. flexibilis</i>
	<i>P. fluvialis</i>	<i>P. pharmacofabricae</i>	98.61	<i>P. fluvialis</i>

Positive hits were then inspected manually to confirm the typical features of *Pseudomonas* CLP NRPS clusters (i.e. the presence of tandem TE-domains and the absence of epimerization domains) and synteny (i.e. number of modules and distribution along the NRPS genes), all based on previously described CLP NRPS gene clusters annotations [35,36]. All known and newly identified strains carrying CLP BGCs together with their affiliation to CLP families and the accession numbers of their NRPS genes are presented in Table S7. The phylogenetic relationship between known and newly identified CLP producers was assessed, by family, based on concatenated NRPS amino acid sequences (Figure S2 to S4).

To explore the genetic diversity within the *P. putida* group, we present expanded phylogenies (*rpoD* and genome-based) of the group including type strains, type strains of newly described species and affiliated strains (Table S1), known and newly identified CLP producers and a selection of 44 strains tentatively reassigned to known species (Table S7). The phylogenetic tree based on whole genomes was constructed as described above but using 2,570 core orthogroups.

Several new phylogenetic groups (G) and subgroups (SG) were delineated based on branch length, grouping and bootstrap values on both *rpoD* and whole genome phylogenies (Table 1, Figure 1, 2 and S5). The new groups and subgroups were named after the first species described in a group or subgroup.

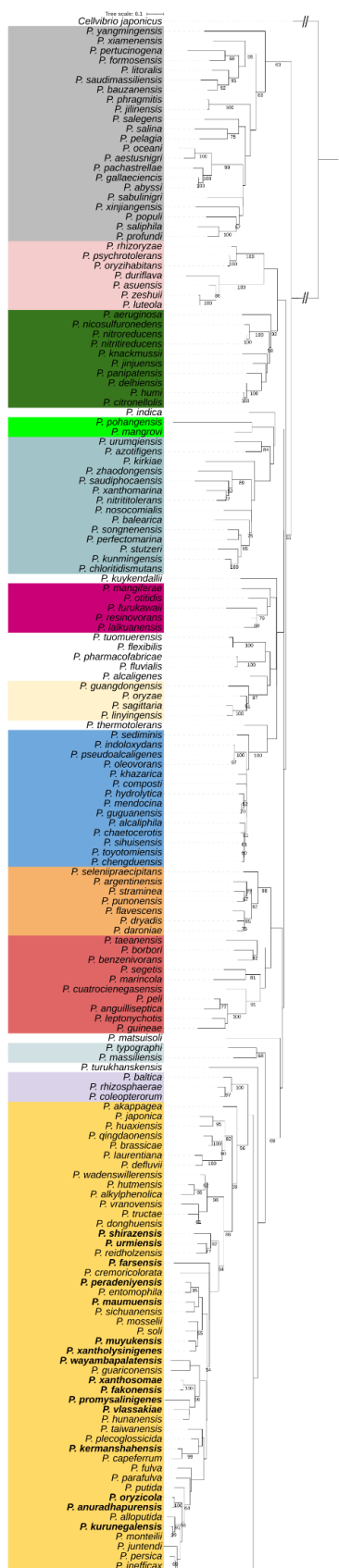
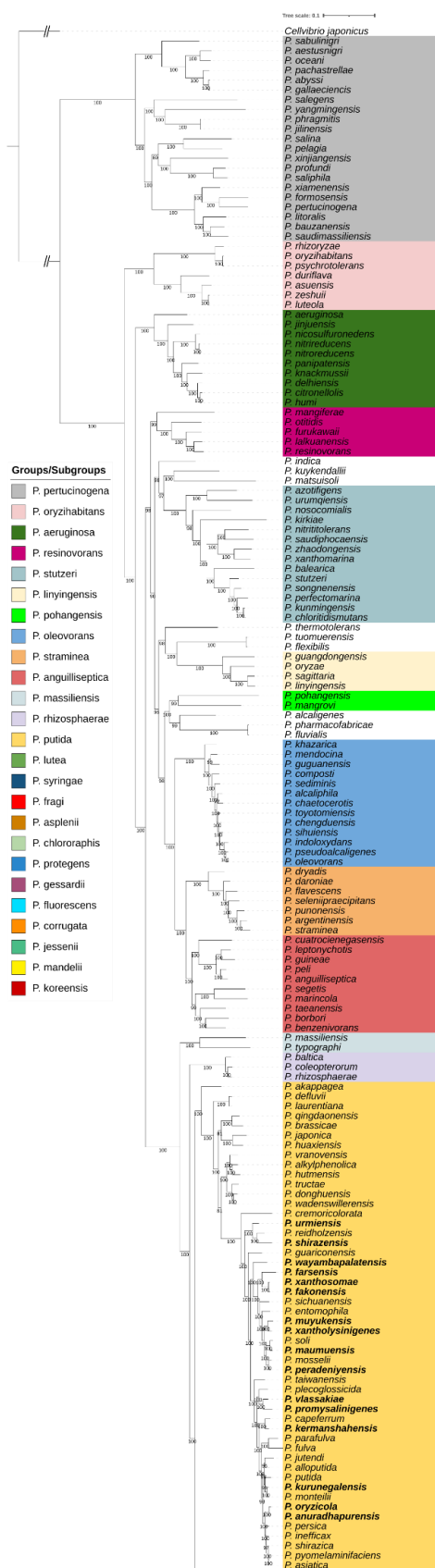
3. Results and Discussion

3.1 Defining new *Pseudomonas* species

In a recent study, we performed *rpoD*-based identifications which allowed us to identify 31 new *Pseudomonas* species [17]. In this study, three strains were incorrectly identified as representative strains of a new species (i.e. *Pseudomonas* #5, SWRI59, SWRI68 and SWRI77) and were later identified as *P. capeferrum* strains. Since the publication of this study, further *rpoD*-based identifications enabled us to identify 17 additional *Pseudomonas* species. Four strains, namely, SWRI22, OE 28.3, SWRI76 and CMR5c, were first assessed as new species but were later assigned to newly published *Pseudomonas* species (i.e. #29 *P. carnis*, #30 *P. edaphica*, #31 *P. atacamensis* and #45 *P. aestus*; Table S2 and S6). Finally, a total of 43 new *Pseudomonas* species could be defined (Appendix A) and the result of phenotypic profiling together with their two culture collection numbers are presented in Table S5. Hybrid assemblies of the genomes resulted in 23 closed genomes and 20 draft genomes with improved contiguity. Due to technical issues we have not been able to increase the quality of the draft genomes of strains BW11P2, COW3 and SWRI196. To avoid the proposal of new species based on single strains, the *rpoD* sequences of the 43 new species were used as query to search for additional strains in Genbank. We therefore reassigned 128 *Pseudomonas* strains, of which 29 with whole genome sequences, available through Genbank (Table S1 and S2). Finally, ANIb values were calculated between a total of 348 *Pseudomonas* species (272 type strains and 76 (47+29) *Pseudomonas* strains affiliated to new species) and allowed us to confirm these affiliations and the presence of 43 new *Pseudomonas* species (Table S6). The phylogenetic position of the 43 type strains is shown in Figure 1 and their distribution within the different groups of *Pseudomonas* is detailed in Table 1. All the new species are clustering within the *P. fluorescens* (n=27) and *P. putida* (n=16) groups. We amended the existing subgroups of *P. fluorescens* as follows: *P. asplenii* (inclusion of *P. vanderleydeniana*), *P. corrugata* (inclusion of *P. alvandrae*, *P. marvdashtae*, *P. tehranensis*, *P. zanzanensis* and *P. zarinae*), *P. fluorescens* (inclusion of *P. asgharzadehiana*, *P. azadiae*, *P. khavaziana*, *P. salmasensis* and *P. tritici*), *P. gessardii* (inclusion of *P. shahriarae*), *P. jessenii* (inclusion of *P. asgharzadehiana* and *P. azerbaijanoccidens*), *P. koreensis* (inclusion of *P. bananamidigenes*, *P. botevensis*, *P. ekonensis*, *P. hamedanensis*, *P. iranensis*, *P. khorasanensis*, *P. monsensis*, *P. siliginis*, *P. tensinigenes*, *P. trititicola* and *P. zae*), *P. mandelii* (inclusion of *P. farris*), *P. protegens* (inclusion of *P. sessiligenes*). The remaining sixteen new species allowed the partitioning of the *P. putida* group in fifteen subgroups as described in section 3.3.

3.2 Comparison of whole genome and *rpoD*-based phylogenies

The phylogenetic relationships between known and newly described *Pseudomonas* species are presented in Figure 1, respectively the whole genome, based on 1,502 core orthogroups, and the *rpoD*-based phylogenies. The phylogenies include 273 type strains of *Pseudomonas* species (270 for the whole genome phylogeny; the 3 type strains with no genome available are highlighted in Table S4) and 43 type strains of the newly described *Pseudomonas* species. Three type strains of *Pseudomonas* were excluded from the analysis, (1) *P. hydrolytica* with an abnormally long genome (10.4 Mbp) and (2) *P. hussainii* and *P. caeni* harboring short genomes (respectively 3.68 and 3.03 Mbp) and clustering with members of other genera within the *Pseudomonadaceae* [4,6,17]. We suspect that the latter two are not *Pseudomonas* species and need to be reclassified.



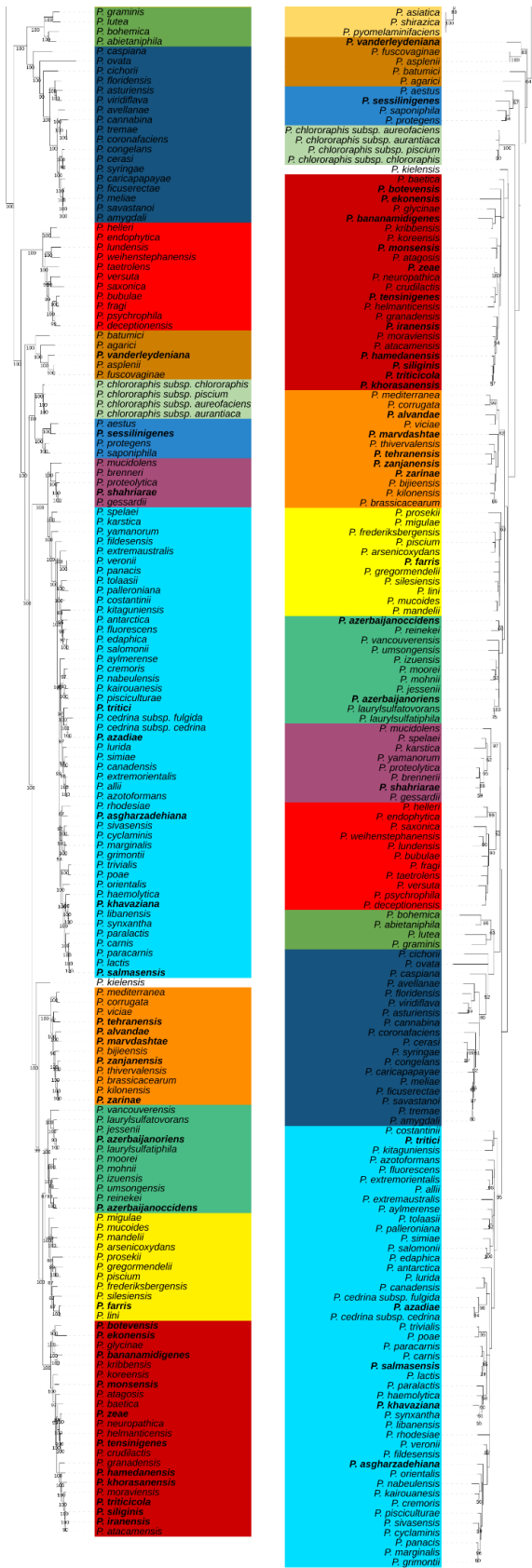


Figure 1. Phylogenetic tree based on 1,508 core orthogroups using IQ-TREE with automatic model selection and 1,000 ultrafast bootstraps (left) and the *rpoD* gene (right; Maximum likelihood tree, GTR+G+I model (MEGA-X)) from, respectively, 313 and 316, *Pseudomonas* species and subspecies. Bootstrap value were calculated based on 1,000 replications. Newly described species are highlighted in bold. *Cellvibrio japonicus* is used as outgroup.

Indeed, *P. caeni* caught the attention of Hesse and colleagues [18], for its unusual genomic features and is already displayed as *Thiopseudomonas caeni* in the GTDB (https://gtdb.ecogenomic.org/tree?r=f__Pseudomonadaceae). A tree of the *Pseudomonadaceae* family, including *P. hussainii* and *P. caeni* (*T. caeni*) but also all type strains of the *Azomonas*, *Azotobacter*, *Entomomonas*, *Oblitimonas*, *Pseudomonas*, *Thiopseudomonas* and *Ventosimonas* genera is shown in Figure S1.

The thirteen groups of *Pseudomonas* previously identified in several studies (i.e. *P. pertucinogena*, *P. oryzihabitans*, *P. aeruginosa*, *P. resinovorans*, *P. stutzeri*, *P. linyingensis*, *P. oleovorans*, *P. straminea*, *P. anguilliseptica*, *P. putida*, *P. lutea*, *P. syringae* and *P. fluorescens*) are all well supported in both trees [4,6,17,18]. In addition to these thirteen groups, three new groups, namely *P. pohangensis*, *P. massiliensis* and *P. rhizosphaerae*, were identified based on branch length and the strong bootstrap support values separating them from the neighboring groups (Figure 1). Furthermore, as previously observed, ten species are scattered across the tree and represent orphan groups formed by only one species (Figure 1). An overview of all known and newly proposed groups is summarized in Table 1. Overall, both trees are highly consistent in topology, although the tree inferred by whole genome analysis is supported by stronger bootstrap values. Two main differences can still be highlighted, (1) the position of the *P. syringae* and *P. lutea* group, clustering inside the *P. fluorescens* group in the *rpoD*-based tree and (2) the position of *P. karstica*, *P. spelaei* and *P. yamanorum*, clustering within the *P. gessardii* subgroup in *rpoD* or MLSA phylogenies [4,6,17], while in phylogenies based on whole genome analysis they cluster within the *P. fluorescens* subgroup ([18] and Figure 1).

3.3 Genomic diversity within the *P. putida* group

3.3.1 Identification and reassignment at the species level

Several studies have revealed inconsistencies within public databases in which genomes of *Pseudomonas* are not identified (*Pseudomonas* sp.) or incorrectly assigned at the species level [4,37,38]. Within the *P. putida* group, a huge number of strains are incorrectly assigned to *P. putida* [4,37]. Here, we propose to update the *P. putida* group with 16 new *Pseudomonas* species and tentatively reassigned 44 genomes of non-type strains (Table 3). A total of 25 strains are affiliated to known and newly described species (*P. shirazensis* n=1, *P. guariconensis* n=2, *P. wayambapalatensis* n=2, *P. farsensis* n=1, *P. peradeniyensis* n=1, *P. capeferrum* n=2, *P. kermanshahensis* n=4, *P. juntendi* n=2, *P. alloputida* n=6 and *P. kurunegalensis* n=4), while the remaining 19 strains represent an additional 13 new species. As previously observed for the genus *Pseudomonas*, these results confirm the fact that type strains still represent a small fraction of the genomic diversity within the *P. putida* group.

Table 3. Phylogenetic affiliation based on ANIb values for the 44 whole genome sequenced strains belonging to the *P. putida* group, previously not assigned, or incorrectly assigned at the species level. Accession numbers are shown in Table S7.

Subgroups	Strain	Closest type strain	ANIb %	Re-identified species
<i>P. japonica</i>	<i>P. putida</i> CSV86	<i>P. japonica</i>	86.94	<i>Pseudomonas</i> sp. #1
<i>P. reidholzensis</i>	<i>P. putida</i> 02C-26	<i>P. shirazensis</i>	97.25	<i>P. shirazensis</i>
<i>P. guariconensis</i>	<i>P. putida</i> IEC33019	<i>P. guariconensis</i>	91.47	<i>Pseudomonas</i> sp. #2
	<i>P. putida</i> WP4-W18-CRE-03	<i>P. guariconensis</i>	99.38	<i>P. guariconensis</i>
	<i>P. putida</i> WP8-W18-CRE-01		99.47	
<i>P. wayambapalatensis</i>	<i>P. putida</i> NX-1	<i>P. wayambapalatensis</i>	94.69	<i>Pseudomonas</i> sp. #3
	<i>P. putida</i> PC2		94.78	
	<i>Pseudomonas</i> sp. RW3S2	<i>P. wayambapalatensis</i>	99.21	<i>P. wayambapalatensis</i>
	<i>Pseudomonas</i> sp. RW10S2		99.26	
<i>P. farsensis</i>	<i>Pseudomonas</i> sp. SWRI51	<i>P. farsensis</i>	98.65	<i>P. farsensis</i>
<i>P. mosselii</i>	<i>Pseudomonas</i> sp. 250J	<i>P. peradeniyensis</i>	96.15*	<i>Pseudomonas</i> sp. #4
	<i>Pseudomonas</i> sp. BW16M2	<i>P. peradeniyensis</i>	96.59	<i>P. peradeniyensis</i>
<i>P. plecoglossicida</i>	<i>P. putida</i> GM84	<i>P. plecoglossicida</i>	91.13	<i>Pseudomonas</i> sp. #5
	<i>P. putida</i> GIMC5401-PPKH-115	<i>P. plecoglossicida</i>	87.01	<i>Pseudomonas</i> sp. #6
	<i>P. putida</i> BR-PH17	<i>P. plecoglossicida</i>	86.75	<i>Pseudomonas</i> sp. #7
	<i>P. putida</i> W619		85.75	
<i>P. vlassakiae</i>	<i>P. putida</i> AA7	<i>P. vlassakiae</i>	90.88	<i>Pseudomonas</i> sp. #8
	<i>P. putida</i> W5	<i>P. vlassakiae</i>	91.64	<i>Pseudomonas</i> sp. #9
<i>P. capeferrum</i>	<i>Pseudomonas</i> sp. SWRI68	<i>P. capeferrum</i>	98.66	<i>P. capeferrum</i>
	<i>Pseudomonas</i> sp. SWRI59		98.65	
	<i>P. putida</i> E41	<i>P. kermanshahensis</i>	97.48	<i>P. kermanshahensis</i>
	<i>P. putida</i> E46		97.55	
	<i>Pseudomonas</i> sp. SWRI50		99.39	
	<i>Pseudomonas</i> sp. SWRI67		99.99	
<i>P. putida</i>	<i>P. putida</i> SY153	<i>P. jutendi</i>	98.15	<i>P. jutendi</i>
	<i>P. putida</i> TIJ-51		97.77	
	<i>P. putida</i> GB-1	<i>P. alloputida</i>	90.49	<i>Pseudomonas</i> sp. #10
	<i>P. putida</i> PP112420		90.54	
	<i>P. putida</i> S13-1-2	<i>P. putida</i>	94.55	<i>Pseudomonas</i> sp. #11
	<i>P. putida</i> KF715		93.73	<i>Pseudomonas</i> sp. #12
	<i>P. putida</i> ZXPA-20		93.40	
	<i>P. putida</i> H8234		93.33	
	<i>P. putida</i> B1		93.40	
	<i>P. putida</i> R51	<i>P. alloputida</i>	95.00*	<i>Pseudomonas</i> sp. #13
	<i>P. putida</i> BS3701	<i>P. alloputida</i>	96.67	<i>P. alloputida</i>
	<i>P. putida</i> MX-2		96.49	
	<i>P. putida</i> LS46		96.44	
	<i>P. putida</i> 15420352		96.42	
	<i>P. putida</i> YC-AE1		96.40	
	<i>P. putida</i> T25-27		96.51	

<i>P. monteilii</i> 170620603RE	<i>P. kurunegalensis</i>	99.45	<i>P. kurunegalensis</i>
<i>P. monteilii</i> 170918607		99.44	
<i>P. monteilii</i> STW0522-72		99.64	
<i>P. monteilii</i> FDAARGOS171		99.77	

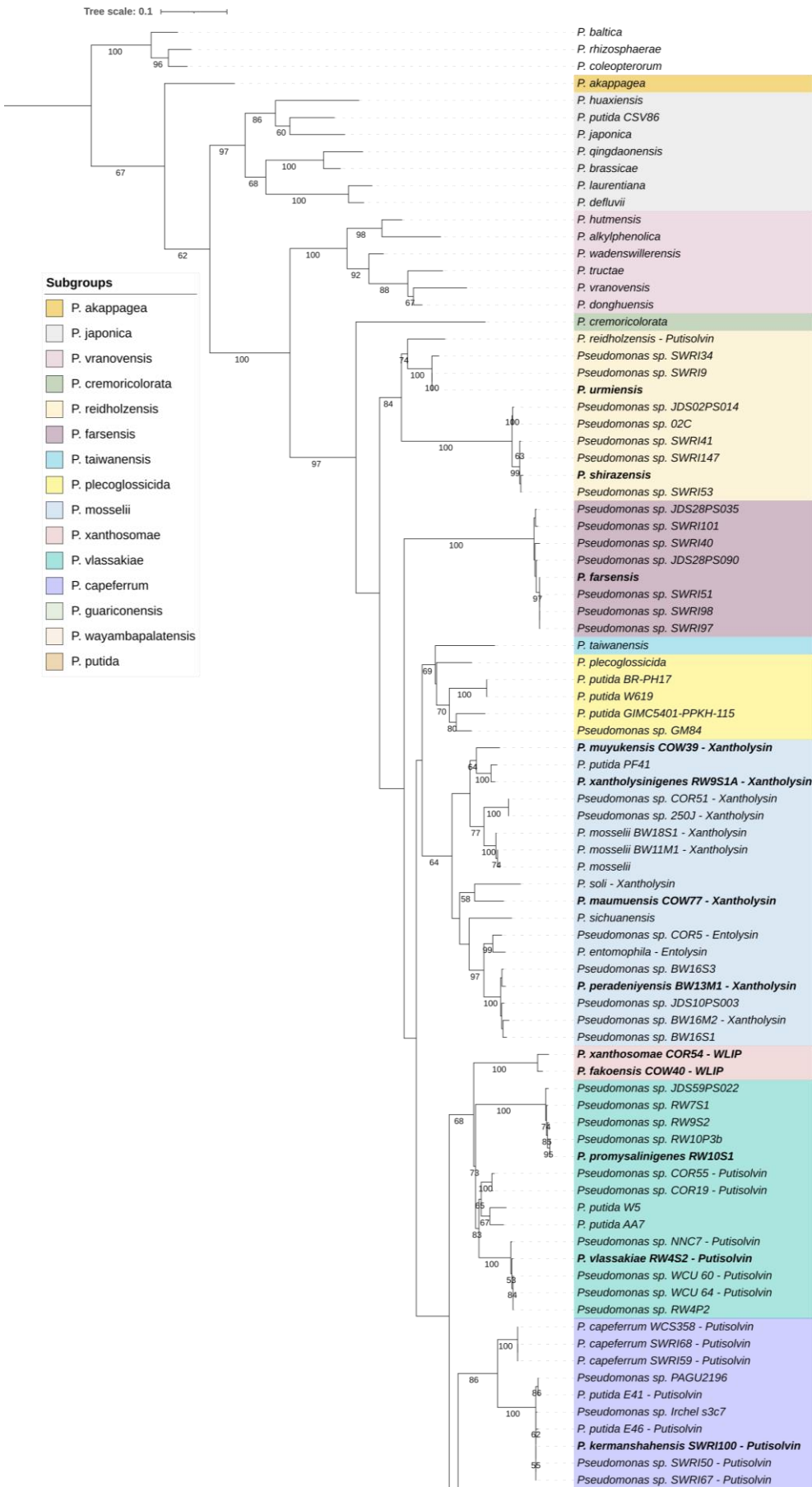
* dDDH < 70%.

3.3.2 Distribution of CLP biosynthesis gene clusters

CLPs are specialized metabolites that often support important ecological functions including cooperation, phytopathogenicity or antagonism [29,36,39]. CLPs consist of a fatty acid tail attached to a cyclized oligopeptide and are synthesized by Non-Ribosomal Peptide Synthetases (NRPSs) [29,35]. The modularity of these enzymes enable *Pseudomonas* strains to produce a wide diversity of CLPs, resulting in their classification in several families [28,29,35]. The relationship between CLP diversity and *Pseudomonas* taxonomy was recently highlighted and it was demonstrated that certain CLP families were exclusive to specific subgroups of *P. fluorescens* [30,31,36]. On the other hand, the *P. putida* group was demonstrated to host CLP producers from different families [30,31]. Indeed, CLP production is widespread within *P. putida* group and different type strains (i.e. *P. capeferrum*, *P. entomophila* and *P. soli*) and many non-type strains (e.g. RW10S2, PCL1445, BW11M1, 250J, COR5, COW10, COR19, COR51; Table S7) were formerly characterized as producers of CLPs from the Viscosin (WLIP producers), Putisolvin, Entolysin and Xantholysin families [30,31,39–49]. Among the 16 type strains of the newly described species, four were previously described as CLP producers, two WLIP producers, *P. fakonensis* COW40 and *P. xanthosomae* COR54 and two Xantholysin producers, *P. maumuensis* COW77 and *P. muyukensis* COW39 (Table S7) [30,31]. We therefore searched for CLP NRPSs in a selection of *Pseudomonas* genomes, including all the type strains belonging to the *P. putida* group (known and newly described, n=51) and the 44 genomes of non-type strains presented in Table 3. About 65% of the strains (i.e type strains 34/51; non-type strains (28/44)) did not carry CLP NRPSs in their genomes (Tables S7). Our analysis revealed the presence of NRPSs from the Viscosin family (WLIP-like NRPSs) in the genomes of two strains (*P. wayambapalatensis* RW3S1^T and RW3S2), from the Putisolvin family in five type strains (*P. fulva*, *P. kermanshahensis*, *P. parafulva*, *P. reidholzensis* and *P. vlassakiae*) and five non-type strains (*P. capeferrum* SWRI59 and SWRI68 and *P. kermanshahensis* SWRI67, SWRI50, E46) and from the Xantholysin family in two type strains (*P. peradeniyensis* and *P. xantholysinigenes*) and two non-type strains (*P. mosselii* BW18S1 and *P. peradeniyensis* BW16M2). The poor genome quality of two type strains, namely *P. brassicae* and *P. juntendi*, revealed the presence of NRPS gene fragments coding for tandem thioesterase (TE) domains. Tandem TE domains are specific of *Pseudomonas* CLP NRPS therefore these results indicate *P. brassicae* and *P. juntendi* carry CLP NRPS genes and most likely produce CLPs. Further analyses, chemical characterization and/or a hybrid assembly based on long read sequencing and Illumina sequencing, are needed to identify the CLPs. We previously highlighted, in the type strains of *P. asplenii* and *P. fucovaginae*, a NRPS system predicted to assemble a lipotridecapeptide (LP-13) but this metabolite still awaits chemical and functional characterization [36]. A putative LP-13 biosynthesis gene cluster is also present in the genome of *P. tractae*. Oni and colleagues also reported the presence of a new CLP (N8, 17:8, 17 amino acids of which 8 in the macrocycle) within the *P. putida* group [30,31]. Altogether, these results highlight a wide diversity of CLP producers from known, and yet to describe new, CLP families within the *P. putida* group.

3.3.3 Partitioning of the *P. putida* group

To present an integrated approach linking the genetic diversity and the metabolic potential of *Pseudomonas* species, we mapped the presence of CLP biosynthesis gene clusters on an extended phylogeny of the *P. putida* group (Figure 2 and S5). As shown in Figure 1, the *P. putida* group is composed of several subgroups (Figure 1). The extended phylogeny allowed us to define 15 subgroups, *P. japonica* (n=6), *P. vranovensis* (n=6), *P. reidholzensis* (n=3), *P. xanthosomae* (n=2), *P. mosselii* (n=8), *P. vlassakiae* (n=3), *P. capeferrum* (n=2), *P. putida* (n=14) including 7 orphan subgroups (*P. akappagea*, *P. cremoricolorata*, *P. guariconensis*, *P. wayambapalatensis*, *P. farsensis*, *P. taiwanensis* and *P. plecoglossicida*).



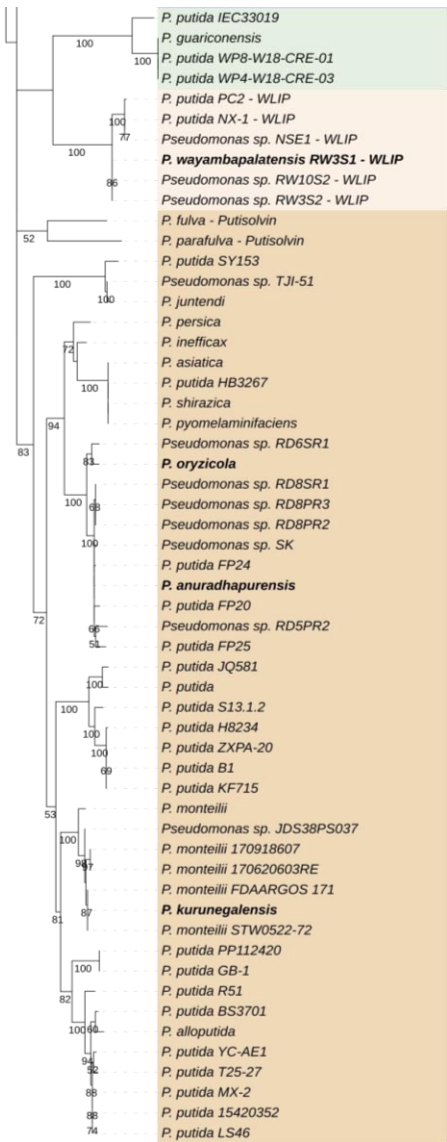


Figure 2. Phylogenetic tree of the *P. putida* group based on the *rpoD* gene of 139 *Pseudomonas* strains. All the strains included in this analysis, together with their accession numbers and the output of the prospection for CLP BGCs are detailed in Table S7. The maximum likelihood phylogenetic tree was constructed using the GTR+G+I model (MEGA-X). Bootstrap values were calculated based on 1,000 replications and only bootstrap values higher than 50% are indicated. Newly described species are highlighted in bold. The *P. rhizosphaerae* group is used as outgroup. The corresponding tree based on whole genome sequences is shown in Figure S5.

The distribution of all type strains in the 15 subgroups is detailed in Table 1. Among the 44 non-type strains used in section 3.3.1, 19 were highlighted to represent 13 new species distributed in 7 subgroups *P. japonica* (n=1), *P. guariconensis* (n=1), *P. wayambapalatensis* (n=1), *P. mosselii* (n=1), *P. plecoglossicida* (n=3), *P. vlassakiae* (n=2) and *P. putida* (n=4) (Table 3). These additions to the *P. putida* phylogenies allowed us to seize a small portion of the genomic diversity among environmental *Pseudomonas* strains, but also to pinpoint the immediate growing potential of the newly defined subgroups.

In addition, we observed that in both *rpoD* and whole genome phylogenies, the distribution of CLP biosynthesis gene clusters was associated with this phylogenetic groupings. All xantholysin and entolysin producers were grouped within the *P. mosselii* subgroup, putisolvin producers are clustering in four subgroups (i.e. *P. putida*, *P. reidholzensis*, *P. vlassakiae* and *P. capeferrum*) and WLIP producers are distributed over two subgroups (*P. xanthosomae* and *P. wayambapalatensis*). Moreover, the phylogenies based

on concatenated NRPS amino acid sequences (Xantholysin/Entolysin families Figure S2, Putisolvin family Figure S3 and WLIP producers Figure S4) revealed different clusters that perfectly match the distribution of CLP producers within the different subgroups. These results demonstrate that the *rpoD*-gene allows both the identification of *Pseudomonas* isolates and the construction of robust phylogenies providing information about the affiliation of producers to CLP families.

The strong congruence between the phylogenetic tree based on the NRPS sequences and the *rpoD*- and whole genome based phylogenies indicates that CLP biosynthesis genes have largely evolved in accordance to the evolutionary history of *Pseudomonas* species within the *P. putida* group. However, *P. reidholzensis* carries a putisolvin biosynthetic gene cluster which is absent from the genome of the closely related species. Furthermore, CLP producers from the Viscosin family, including WLIP producers, are predominantly found within the *P. fluorescens* group [30,31,50], the two subclusters of WLIP producers present in the *P. putida* group being an exception. Altogether, these observations indicate that *Pseudomonas* CLP NRPS clusters have a complex evolutionary history probably involving both vertical and horizontal gene transfer.

4. Conclusions

Our update of the genus with 43 new species together with our analysis of 313 genomes of type strains allowed us to propose a robust revised phylogeny of the *Pseudomonas* spp. This study aimed to fill the gap between the currently named species and the real genomic diversity within the genus *Pseudomonas*. Additional work is needed to complete this task and genome-based standards for species definition should be favored over highly variable phenotypic tests for publication. Our study validated the use of the *rpoD* gene for species identification but also to study evolutionary relationships within the genus *Pseudomonas*. Furthermore, *rpoD*-based phylogenies can as well be very useful to specifically prospect for CLP biosynthesis gene clusters and affiliation of producers to known CLP families. Finally, the use of genomic sequences appears as essential to reveal the ecological and metabolic potential of *Pseudomonas* spp.

Supplementary Materials: The following are available online at www.mdpi.com/xxx/s1, Figure S1: Genome-based phylogeny of the *Pseudomonadaceae*. Figure S2: Phylogenetic tree based on concatenated NRPS proteins from the Xantholysin family. Figure S3: Phylogenetic tree based on concatenated NRPS proteins from the Putisolvin family. Figure S4: Phylogenetic tree based on concatenated NRPS proteins of WLIP producers from the Viscosin family. Figure S5: Genome-based phylogeny of the *P. putida* group. Table S1: RpoD-based affiliation of strains to newly described *Pseudomonas* species. Table S2: Whole genome based affiliation of strains to newly described *Pseudomonas* species. Table S3: Phenotypic profiles of the 43 new *Pseudomonas* species. Table S4: List of type strains used in this study. Table S5: *rpoD* similarity matrix. Table S6: ANIb similarity matrix. Table S7: Prospection of CLP biosynthesis gene clusters within the *P. putida* group.

Author Contributions: Conceptualization, methodology, investigation, writing original draft and editing, L.G. and C.L.; methodology, investigation and writing-review and editing, M.H and P.V.; resources and formal analysis, H.R.-Z; writing- review and editing and funding acquisition, V.v.N, R.L. and R.D.M. All authors have read and agreed to the published version of the manuscript.

Funding: This research and the APC were funded by the EOS grant 30650620 (RHIZOCLIP). C.L. is supported by an SB PhD fellowship from FWO Vlaanderen (1S64720N).

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A : Descriptions of the 43 new *Pseudomonas* species. The phenotypic descriptions are presented in Table S3. *RpoD* and whole genome based assignment of additional *Pseudomonas* strains to the newly described species are shown in Table S1 and S2.

(#1) Description of *Pseudomonas anuradhapurensis* sp. nov.

Pseudomonas anuradhaturensis (a.nu.ra.dha.pur.en'sis. N.L. fem. adj. *anuradhaturensis*, from Anuradhapura, a city in Sri Lanka).

The type strain is RD8MR3^T (LMG 32021^T = CFBP 8837^T) and was isolated from the endorhizosphere of rice, Anuradhapura, Sri Lanka in 1990. Its G + C content is 63.43 mol% (calculated based on its genome sequence). The 16S rRNA gene, *rpoD* and whole-genome sequence of RD8MR3^T are publicly available through the accession numbers AM911640, MT621460 and CP077097, respectively.

(#2) Description of *Pseudomonas oryzicola* sp. nov.

Pseudomonas oryzicola (o.ry.zi'co.la. L. fem n. *Oryza* rice, L. suff. -cola (from L. n. *incola*) inhabitant dweller; N.L. n. *oryzicola*, rice-dweller).

The type strain is RD9SR1^T (LMG 32022^T = CFBP 8838^T) and was isolated from the exorhizosphere of rice, Anuradhapura, Sri Lanka in 1990. Its G + C content is 62.91 mol% (calculated based on its genome sequence). The 16S rRNA gene, *rpoD* and whole-genome sequence of RD9SR1^T are publicly available through the accession numbers AM911646, MT621461, and JABWRZ000000000, respectively.

(#3) Description of *Pseudomonas kurunegalensis* sp. nov.

Pseudomonas kurunegalensis (ku.ru.ne.gal.en'sis. N.L. fem. adj. *kurunegalensis*, from Kurunegala, a city in Sri Lanka).

The type strain is RW1P2^T (LMG 32023^T = CFBP 8839^T) and was isolated from the rhizoplane of rice, Kurunegala, Sri Lanka in 1990. Its G + C content is 62.09 mol% (calculated based on its genome sequence). The 16S rRNA gene, *rpoD* and whole-genome sequence of RW1P2^T are publicly available through the accession numbers AM911650, MT621449, and JABWSB000000000, respectively.

(#4) Description of *Pseudomonas kermanshahensis* sp. nov.

Pseudomonas kermanshahensis (ker.man.shah.en'sis. N.L. fem. adj. *kermanshahensis*, from Kermanshah, a city in Iran).

The type strain is SWRI100^T (LMG 32024^T = CFBP 8840^T) and was isolated from the rhizosphere of wheat (cultivar Marvdasht), Kermanshah, Iran in 2004. Its G + C content is 62.22 mol% (calculated based on its genome sequence). The *rpoD* and whole-genome sequence of SWRI100^T are publicly available through the accession numbers MT621423 and JABWRY000000000, respectively.

(#5) Description of *Pseudomonas wayambapalatensis* sp. nov.

Pseudomonas wayambapalatensis (wa.yam.ba.pa.lat.en'sis. N.L. fem. adj. *wayambapalatensis*, from Wayamba Palata, the name of the Northwestern province in Sri Lanka).

The type strain is RW3S1^T (LMG 32025^T = CFBP 8841^T) and was isolated from the exorhizosphere of rice, Kurunegala, Sri Lanka in 1990. Its G + C content is 63.24 mol% (calculated based on its genome sequence). The 16S rRNA gene, *rpoD* and whole-genome sequence of RW3S1^T are publicly available through the accession numbers AM911665, MT621434 and CP077096, respectively.

(#6) Description of *Pseudomonas xantholysinigenes* sp. nov.

Pseudomonas xantholysinigenes (xan.tho.ly.si.ni' ge.nes. N.L. neut. n. *xantholysinum*, *xantholysin*; Gr. v. *gennao* to produce; N.L. part. adj. *xantholysinigenes*, *xantholysin-producing*).

The type strain is RW9S1A^T (LMG 32026^T = CFBP 8842^T) and was isolated from the exorhizosphere of rice, Kurunegala, Sri Lanka in 1990. Its G + C content is 64.16 mol% (calculated based on its genome sequence). The 16S rRNA gene, *rpoD* and whole-genome sequence of RW9S1A^T are publicly available through the accession numbers AM911667, MT621442 and CP077095, respectively.

(#7) Description of *Pseudomonas peradeniyensis* sp. nov.

Pseudomonas peradeniyensis (pe.ra.de.niy.en'sis. N.L. fem. adj. *peradeniyensis*, from Peradeniya, a city in Sri Lanka).

The type strain is BW13M1^T (LMG 32027^T = CFBP 8887^T) and was isolated from banana plant endorhizosphere, Peradeniya, Sri Lanka in 1990. Its G + C content is 64.62 mol% (calculated based on its genome sequence). The *rpoD* and whole-genome sequence of BW13M1^T are publicly available through the accession numbers MT621446 and JABWRJ000000000, respectively.

(#8) Description of *Pseudomonas vlassakiae* sp. nov.

Pseudomonas vlassakiae (vlas.sak.i.a'e. N.L. gen. n. *vlassakiae*, from Katrien Vlassak, a Belgian microbiologist who isolated the strain RW4S2 but also RD3MR3, RD9SR1, RW1P2, RW3S1, RW9S1A, BW13M1, RW10S1, RW8P3 and BW11P2, which represent 10 new *Pseudomonas* species).

The type strain is RW4S2^T (LMG 32028^T = CFBP 8843^T) and was isolated from the exorhizosphere of rice, Kurunegala, Sri Lanka in 1990. Its G + C content is 62.98 mol% (calculated based on its genome sequence). The 16S rRNA gene, *rpoD* and whole-genome sequence of RW4S2^T are publicly available through the accession numbers AM911658, MT621428 and JABWRP000000000, respectively.

(#9) Description of *Pseudomonas promysalinigenes* sp. nov.

Pseudomonas promysalinigenes (pro.my.sa.li.ni'ge.nes. N.L. neut. n. *promysalinum*, *promysalin*; Gr. v. *gennao* to produce; N.L. part. adj. *promysalinigenes*, *promysalin-producing*).

The type strain is RW10S1^T (LMG 32029^T = CFBP 8844^T) and was isolated from the exorhizosphere of rice, Kurunegala, Sri Lanka in 1990. Its G + C content is 60.62 mol% (calculated based on its genome sequence). The 16S rRNA gene, *rpoD* and whole-genome sequence of RW10S1^T are publicly available through the accession numbers AM911668, MT621430 and CP077094, respectively.

(#10) Description of *Pseudomonas urmiensis* sp. nov.

Pseudomonas urmiensis (ur.mi.en'sis. N.L. fem. adj. *urmiensis*, from Urmia, a city in Iran).

The type strain is SWRI10^T (LMG 32030^T = CFBP 8845^T) and was isolated from the rhizosphere of wheat (cultivar Marvdasht), West Azerbaijan, Iran in 2004. Its G + C content is 61.81 mol% (calculated based on its genome sequence). The *rpoD* and whole-genome sequence of SWRI10^T are publicly available through the accession numbers MT621419 and JABWRE000000000, respectively.

(#11) Description of *Pseudomonas shirazensis* sp. nov.

Pseudomonas shirazensis (shi.raz.en'sis. N.L. fem. adj. *shirazensis*, from Shiraz, a city in Iran).

The type strain is SWRI56^T (LMG 32031^T = CFBP 8846^T) and was isolated from the rhizosphere of wheat (cultivar Shiraz), Shiraz, Iran in 2004. Its G + C content is 61.85 mol% (calculated based on its genome sequence). The *rpoD* and whole-genome sequence of SWRI56^T are publicly available through the accession numbers MT621418 and JABWRD000000000, respectively.

(#12) Description of *Pseudomonas farsensis* sp. nov.

Pseudomonas farsensis (fars.en'sis. N.L. fem. adj. *farsensis*, from Fars, a province in Iran).

The type strain is SWRI107^T (LMG 32032^T = CFBP 8847^T) and was isolated from the rhizosphere of wheat (cultivar Azadi), Shiraz, Iran in 2004. Its G + C content is 62.58 mol% (calculated based on its genome sequence). The *rpoD* and whole-genome sequence of SWRI107^T are publicly available through the accession numbers MT621411 and JABWRF000000000, respectively.

(#13) Description of *Pseudomonas vanderleydeniana* sp. nov.

Pseudomonas vanderleydeniana (van.der.ley.den.i.a'na. N.L. fem. adj. *vanderleydeniana*, from Jos Vanderleyden, a Belgian microbiologist who studied plant growth-promoting properties of root-associated alpha- and gammaproteobacteria, including nitrogen-fixing and fluorescent *Pseudomonas* isolates.

The type strain is RW8P3^T (LMG 32033^T = CFBP 8848^T) and was isolated from the rhizoplane of rice, Kurunegala, Sri Lanka in 1990. Its G + C content is 62.97 mol% (calculated based on its genome sequence). The *rpoD* and whole-genome sequence of RW8P3^T are publicly available through the accession numbers MT621472 and CP077093, respectively.

(#14) Description of *Pseudomonas bananamidigenes* sp. nov.

Pseudomonas bananamidigenes (ba.na.na.mi.di'ge.nes. N.L. neut. n. *bananamidum*, bananamide; Gr. v. *gennao* to produce; N.L. part. adj. *bananamidigenes*, bananamide-producing).

The type strain is BW11P2^T (LMG 32034^T = CFBP 8849^T) and was isolated from banana plant exorhizosphere, Galagedara, Sri Lanka in 1990. Its G + C content is 60.62 mol% (calculated based on its genome sequence). The *rpoD* and whole-genome sequence of BW11P2^T are publicly available through the accession numbers MT621496 and LRUN000000000, respectively.

(#15) Description of *Pseudomonas iranensis* sp. nov.

Pseudomonas iranensis (i.ran.en'sis. N.L. fem. adj. *iranensis*, from Iran).

The type strain is SWRI54^T (LMG 32035^T = CFBP 8850^T) and was isolated from the rhizosphere of wheat (cultivar Shiraz), Shiraz, Iran in 2004. Its G + C content is 59.89 mol% (calculated based on its genome sequence). The *rpoD* and whole-genome sequence of SWRI54^T are publicly available through the accession numbers MT621504 and CP077092, respectively.

(#16) Description of *Pseudomonas khorasanensis* sp. nov.

Pseudomonas khorasanensis (kho.ra.san.en'sis. N.L. fem. adj. *khorasanensis*, from Khorasan, a province in Iran).

The type strain is SWRI153^T (LMG 32036^T = CFBP 8851^T) and was isolated from the rhizosphere of wheat (cultivar Kaasparoo), Khorasan, Iran in 2004. Its G + C content is 59.71 mol% (calculated based on its genome sequence). The *rpoD* and whole-genome sequence of SWRI153^T are publicly available through the accession numbers MT621508 and JABWQP000000000, respectively.

(#17) Description of *Pseudomonas hamedanensis* sp. nov.

Pseudomonas hamedanensis (ha.me.dan.en'sis. N.L. fem. adj. *hamedanensis*, from Hamedan, a city in Iran).

The type strain is SWRI65^T (LMG 32037^T = CFBP 8852^T) and was isolated from the rhizosphere of wheat, Hamedan, Iran in 2004. Its G + C content is 59.99 mol% (calculated based on its genome sequence). The *rpoD* and whole-genome sequence of SWRI65^T are publicly available through the accession numbers MT621514 and CP077091, respectively.

(#18) Description of *Pseudomonas zeae* sp. nov.

Pseudomonas zeae (ze'ae. L. gen. n. *zeae*, from *Zea mays*, corn).

The type strain is OE 48.2^T (LMG 32038^T = CFBP 8853^T) and was isolated from the rhizosphere of maize, in Belgium ~ 1984-1985. Its G + C content is 58.99 mol% (calculated based on its genome sequence). The *rpoD* and whole-genome sequence of OE 48.2^T are publicly available through the accession numbers MT621498 and CP077090, respectively.

(#19) Description of *Pseudomonas tensinigenes* sp. nov.

Pseudomonas tensinigenes (ten.si.ni' ge.nes. N.L. neut. n. *tensinum*, tensin; Gr. v. *gennao* to produce; N.L. part. adj. *tensinigenes*, tensin-producing).

The type strain is ZA 5.3^T (LMG 32039^T = CFBP 8882^T) and was isolated from the rhizosphere of wheat, in Belgium ~ 1984-1985. Its G + C content is 59.17 mol% (calculated based on its genome sequence). The *rpoD* and whole-genome sequence of ZA 5.3^T are publicly available through the accession numbers MT621501 and CP077089, respectively.

(#20) Description of *Pseudomonas mونسensis* sp. nov.

Pseudomonas mونسensis (mons.en'sis. N.L. fem. adj. *mونسensis*, from Mons, a city in Belgium).

The type strain is PGSB 8459^T (LMG 32040^T = CFBP 8854^T) and was isolated from the rhizosphere of maize, Mons, Belgium ~ 1984-1985. Its G + C content is 60.05 mol% (calculated based on its genome sequence). The *rpoD* and whole-genome sequence of PGSB 8459^T are publicly available through the accession numbers MT621495 and CP077087, respectively.

(#21) Description of *Pseudomonas zانجانensis* sp. nov.

Pseudomonas zانجانensis (zan.jan.en'sis. N.L. fem. adj. *zانجانensis*, from Zanjan, a city in Iran).

The type strain is SWRI12^T (LMG 32041^T = CFBP 8855^T) and was isolated from the rhizosphere of wheat (cultivar Alvand), Zanjan, Iran in 2004. Its G + C content is 61.21

mol% (calculated based on its genome sequence). The *rpoD* and whole-genome sequence of SWRI12^T are publicly available through the accession numbers MT621484 and JABWRB000000000, respectively.

(#22) Description of *Pseudomonas zarinae* sp. nov.

Pseudomonas zarinae (za.ri'nae. N.L. gen. n. *zarinae*, from Zarin, a wheat cultivar).

The type strain is SWRI108^T (LMG 32042^T = CFBP 8856^T) and was isolated from the rhizosphere of wheat (cultivar Zarin), Kermanshah, Iran in 2004. Its G + C content is 60.86 mol% (calculated based on its genome sequence). The *rpoD* and whole-genome sequence of SWRI108^T are publicly available through the accession numbers MT621493 and CP077086, respectively.

(#23) Description of *Pseudomonas tehranensis* sp. nov.

Pseudomonas tehranensis (teh.ran.en'sis. N.L. fem. adj. *tehranensis*, from Tehran, a city in Iran).

The type strain is SWRI196^T (LMG 32043^T = CFBP 8857^T) and was isolated from the rhizosphere of wheat, Tehran, Iran in 2004. Its G + C content is 60.46 mol% (calculated based on its genome sequence). The *rpoD* and whole-genome sequence of SWRI196^T are publicly available through the accession numbers MT621473 and JABWQV000000000, respectively.

(#24) Description of *Pseudomonas marvdashtae* sp. nov.

Pseudomonas marvdashtae (marv.dash'tae. N.L. gen. n. *marvdashtae*, from Marvdasht, a wheat cultivar).

The type strain is SWRI102^T (LMG 32044^T = CFBP 8858^T) and was isolated from the rhizosphere of wheat (cultivar Marvdasht), Kermanshah, Iran in 2004. Its G + C content is 60.64 mol% (calculated based on its genome sequence). The *rpoD* and whole-genome sequence of SWRI102^T are publicly available through the accession numbers MT621490 and JABWQX000000000, respectively.

(#25) Description of *Pseudomonas shahriarae* sp. nov.

Pseudomonas shahriarae (shah.ri.a'rae. N.L. gen. n. *shahriarae*, from Shahriar, a wheat cultivar).

The type strain is SWRI52^T (LMG 32045^T = CFBP 8859^T) and was isolated from the rhizosphere of wheat (cultivar Shahriar), Zanjan, Iran in 2004. Its G + C content is 60.59 mol% (calculated based on its genome sequence). The *rpoD* and whole-genome sequence of SWRI52^T are publicly available through the accession numbers MT621521 and CP077085, respectively.

(#26) Description of *Pseudomonas azadiae* sp. nov.

Pseudomonas azadiae (a.za'di.ae. N.L. gen. n. *azadiae*, from Azadi, a wheat cultivar).

The type strain is SWRI103^T (LMG 32046^T = CFBP 8860^T) and was isolated from the rhizosphere of wheat (cultivar Azadi), Shiraz, Iran in 2004. Its G + C content is 60.69 mol% (calculated based on its genome sequence). The *rpoD* and whole-genome sequence of SWRI103^T are publicly available through the accession numbers MT621536 and JAHSTY000000000, respectively.

(#27) Description of *Pseudomonas tritici* sp. nov.

Pseudomonas tritici (tri'ti.ci. L. gen. n. tritici, of Triticum, wheat).

The type strain is SWRI145^T (LMG 32047^T = CFBP 8883^T) and was isolated from the rhizosphere of wheat, Zanjan, Iran in 2004. Its G + C content is 59.87 mol% (calculated based on its genome sequence). The *rpoD* and whole-genome sequence of SWRI145^T are publicly available through the accession numbers MT621537 and CP077084, respectively.

(#28) Description of *Pseudomonas salmasensis* sp. nov.

Pseudomonas salmasensis (sal.mas.en'sis. N.L. fem. adj. *salmasensis*, from Salmas, a city in Iran).

The type strain is SWRI126^T (LMG 32048^T = CFBP 8861^T) and was isolated from the rhizosphere of wheat (cultivar Zarin), Salmas, Iran in 2004. Its G + C content is 60.16 mol% (calculated based on its genome sequence). The *rpoD* and whole-genome sequence of SWRI126^T are publicly available through the accession numbers MT621526 and CP077083, respectively.

(#29) SWRI22 (LMG 32049 = CFBP 8862): *Pseudomonas carnis*

(#30) OE 28.3 (LMG 32050 = CFBP 8863): *Pseudomonas edaphica*

(#31) SWRI76 (LMG 32051 = CFBP 8864): *Pseudomonas atacamensis*

(#32) Description of *Pseudomonas triticicola* sp. nov.

Pseudomonas triticicola (tri.ti.ci.co'la. L. fem. n. *Triticum*, wheat; L. suff. -cola (from L. n. *incola*), inhabitant dweller; N.L. n. *oryzicola*, wheat-dweller).

The type strain is SWRI88^T (LMG 32052^T = CFBP 8865^T) and was isolated from the rhizosphere of wheat (cultivar Marvdasht), Kermanshah, Iran in 2004. Its G + C content is 59.99 mol% (calculated based on its genome sequence). The whole-genome sequence of SWRI88^T is publicly available through the accession JAHSTX000000000.

(#33) Description of *Pseudomonas siliginis* sp. nov.

Pseudomonas siliginis (si.li'gi.nis. L. gen. n. *siliginis*, of siligo, winter wheat).

The type strain is SWRI31^T (LMG 32053^T = CFBP 8866^T) and was isolated from the rhizosphere of wheat (cultivar Zarin), Kermanshah, Iran in 2004. Its G + C content is 59.97 mol% (calculated based on its genome sequence). The whole-genome sequence of SWRI31^T is publicly available through the accession JAHSTW000000000.

(#34) Description of *Pseudomonas farris* sp. nov.

Pseudomonas farris (far'ris. L. gen. n. *farris*, of husked wheat, of a grain).

The type strain is SWRI79^T (LMG 32054^T = CFBP 8867^T) and was isolated from the rhizosphere of wheat (cultivar Local), Zanjan, Iran in 2004. Its G + C content is 58.74 mol% (calculated based on its genome sequence). The whole-genome sequence of SWRI79^T is publicly available through the accession JAHSTV000000000.

(#35) Description of *Pseudomonas azerbaijanoccidens* sp. nov.

Pseudomonas azerbaijanoccidens (a.zer.bai.jan.oc'ci.dens. Azerbaijan, geographic name; L. fem. adj. *occidens*, western; N.L. fem. adj. *azerbaijanoccidens*, from West Azerbaijan, a province in Iran).

The type strain is SWRI74^T (LMG 32055^T = CFBP 8868^T) and was isolated from the rhizosphere of wheat (cultivar Zarin), Salmas, Iran in 2004. Its G + C content is 59.30 mol% (calculated based on its genome sequence). The whole-genome sequence of SWRI74^T is publicly available through the accession JAHSTU000000000.

(#36) Description of *Pseudomonas alvandae* sp. nov.

Pseudomonas alvandae (al.van'dae. N.L. gen. n. *alvandae*, from Alvand, a wheat cultivar).

The type strain is SWRI17^T (LMG 32056^T = CFBP 8869^T) and was isolated from the rhizosphere of wheat (cultivar Alvand), Zanjan, Iran in 2004. Its G + C content is 60.86 mol% (calculated based on its genome sequence). The whole-genome sequence of SWRI17^T is publicly available through the accession CP077080.

(#37) Description of *Pseudomonas asgharzadehiana* sp. nov.

Pseudomonas asgharzadehiana (as.ghar.za.deh.i.a'na. N.L. fem. adj. *asgharzadehiana*, from Ahmad Asgharzadeh, an Iranian microbiologist who, together with Kazem Khavazi, isolated the strain SWRI132 but also SWRI10, SWRI12, SWRI17, SWRI31, SWRI52, SWRI54, SWRI56, SWRI65, SWRI74, SWRI79, SWRI88, SWRI100, SWRI102, SWRI103, SWRI107, SWRI108, SWRI123, SWRI124, SWRI126, SWRI145, SWRI153 and SWRI196, which represent 23 new *Pseudomonas* species).

The type strain is SWRI132^T (LMG 32057^T = CFBP 8870^T) and was isolated from the rhizosphere of wheat (cultivar Marvdasht), Kermanshah, Iran in 2004. Its G + C content is 60.59 mol% (calculated based on its genome sequence). The whole-genome sequence of SWRI132^T is publicly available through the accession CP077079.

(#38) Description of *Pseudomonas azerbaijanoriens* sp. nov.

Pseudomonas azerbaijanoriens (a.zer.bai.jan.o'ri.ens. Azerbaijan, geographic name; L. fem. adj. *oriens*, eastern; N.L. fem. adj. *azerbaijanoriens*, from East Azerbaijan, a province in Iran).

The type strain is SWRI123^T (LMG 32058^T = CFBP 8871^T) and was isolated from the rhizosphere of wheat (cultivar Zarin), East Azerbaijan, Iran in 2004. Its G + C content is 60.11 mol% (calculated based on its genome sequence). The whole-genome sequence of SWRI123^T is publicly available through the accession CP077078.

(#39) Description of *Pseudomonas khavaziana* sp. nov.

Pseudomonas khavaziana (kha.va.zi.a'na. N.L. fem. adj. *khavaziana*, from Kazem Khavazi, an Iranian microbiologist who, together with Ahmad Asgharzadeh, isolated the strain SWRI124 but also SWRI10, SWRI12, SWRI17, SWRI31, SWRI52, SWRI54, SWRI56, SWRI65, SWRI74, SWRI79, SWRI88, SWRI100, SWRI102, SWRI103, SWRI107, SWRI108, SWRI123, SWRI126, SWRI132, SWRI145, SWRI153 and SWRI196, which represent 23 new *Pseudomonas* species).

The type strain is SWRI124^T (LMG 32059^T = CFBP 8872^T) and was isolated from the rhizosphere of wheat (cultivar Zarin), East Azerbaijan, Iran in 2004. Its G + C content is

59.61 mol% (calculated based on its genome sequence). The whole-genome sequence of SWRI124^T is publicly available through the accession JAHSTT000000000.

(#40) Description of *Pseudomonas botevensis* sp. nov.

Pseudomonas botevensis (bo.tev.en'sis. N.L. fem. adj. *botevensis*, from Boteva, a city in Cameroon).

The type strain is COW3^T (LMG 32176^T = CFBP 8873^T) and was isolated from the roots of white cocoyam (*Xanthosoma sagittifolium*), Boteva, Cameroon in 2008. Its G + C content is 61.21 mol% (calculated based on its genome sequence). The 16S rRNA gene, *rpoB*, *gyrB*, *rpoD* and whole-genome sequence of COW3^T are publicly available through the accession numbers MT507065, MT506178, MT506955, MT506158 and JAHTKI000000000 respectively.

(#41) Description of *Pseudomonas ekonensis* sp. nov.

Pseudomonas ekonensis (e.kon.en'sis. N.L. fem. adj. *ekonensis*, from Ekona, a city in Cameroon).

The type strain is COR58^T (LMG 32175^T = CFBP 8874^T) and was isolated from the roots of red cocoyam (*Xanthosoma sagittifolium*), Ekona, Cameroon in 2008. Its G + C content is 64.87 mol% (calculated based on its genome sequence). The 16S rRNA gene, *rpoB*, *gyrB*, *rpoD* and whole-genome sequence of COR58^T are publicly available through the accession numbers MT507072, MT506185, MT506962, MT506165 and JAHSTS000000000, respectively.

(#42) Description of *Pseudomonas maumuensis* sp. nov.

Pseudomonas maumuensis (mau.mu.en'sis. N.L. fem. adj. *maumuensis*, from Maumu, a city in Cameroon).

The type strain is COW77^T (LMG 32179^T = CFBP 8888^T) and was isolated from the roots of white cocoyam (*Xanthosoma sagittifolium*), Maumu, Cameroon in 2008. Its G + C content is 64.12 mol% (calculated based on its genome sequence). The *rpoB*, *rpoD* and whole-genome sequence of COW77^T are publicly available through the accession numbers MH594167, MK251918 and CP077077, respectively.

(#43) Description of *Pseudomonas fakonensis* sp. nov.

Pseudomonas fakonensis (fa.kon.en'sis. N.L. fem. adj. *fakonensis*, from Fako, a county in Cameroon).

The type strain is COW40^T (LMG 32178^T = CFBP 8875^T) and was isolated from the roots of white cocoyam (*Xanthosoma sagittifolium*), Ekona (Fako county), Cameroon in 2008. Its G + C content is 64.27 mol% (calculated based on its genome sequence). The *rpoB*, *rpoD* and whole-genome sequence of COW40^T are publicly available through the accession numbers MH594146, MK251899 and CP077076, respectively.

(#44) Description of *Pseudomonas xanthosomae* sp. nov.

Pseudomonas xanthosomae (xan.tho.so'mae. L. gen. n. *xanthosomae*, of *Xanthosoma*, cocoyam).

The type strain is COR54^T (LMG 32174^T = CFBP 8876^T) and was isolated from the roots of red cocoyam (*Xanthosoma sagittifolium*), Ekona, Cameroon in 2008. Its G + C content is

64.19 mol% (calculated based on its genome sequence). The *rpoB*, *rpoD* and whole-genome sequence of COR54^T are publicly available through the accession numbers MH594196, MK251872 and CP077075, respectively.

(#45) CMR5c (LMG 32172 = CFBP 8889): *Pseudomonas aestus*

(#46) Description of *Pseudomonas sessilinigenes* sp. nov.

Pseudomonas sessilinigenes (ses.si.li.ni'ge.nes. N.L. neut. n. *sessilinum*, sessilin; Gr. v. *genao* to produce; N.L. part. adj. *sessilinigenes*, sessilin-producing).

The type strain is CMR12a^T (LMG 32173^T = CFBP 8877^T) and was isolated from the roots of red cocoyam (*Xanthosoma sagittifolium*), Bokwai, Cameroon in 2001. Its G + C content is 62.80 mol% (calculated based on its genome sequence). The 16S rRNA gene, *rpoB*, *gyrB* and whole-genome sequence of CMR12a^T are publicly available through the accession numbers FJ652622, FJ652703, FJ652730 and CP077074, respectively.

(#47) Description of *Pseudomonas muyukensis* sp. nov.

Pseudomonas muyukensis (mu.yuk.en'sis. N.L. fem. adj. *muyukensis*, from Muyuka, a district in Cameroon).

The type strain is COW39^T (LMG 32177^T = CFBP 8890^T) and was isolated from the roots of white cocoyam (*Xanthosoma sagittifolium*), Ekona, Cameroon in 2008. Its G + C content is 65.11 mol% (calculated based on its genome sequence). The *rpoD* and whole-genome sequence of COW39^T are publicly available through the accession numbers MK329212 and CP077073, respectively.

References

1. Stackebrandt, E. Report of the Ad Hoc Committee for the Re-Evaluation of the Species Definition in Bacteriology. *Int. J. Syst. Evol. Microbiol.* **2002**, *52*, 1043–1047, doi:10.1099/ijs.0.02360-0.
2. Pallen, M.J.; Telatin, A.; Oren, A. The Next Million Names for Archaea and Bacteria. *Trends Microbiol.* **2021**, *29*, 289–298, doi:10.1016/j.tim.2020.10.009.
3. Gevers, D.; Cohan, F.M.; Lawrence, J.G.; Spratt, B.G.; Coenye, T.; Feil, E.J.; Stackebrandt, E.; de Peer, Y.V.; Vandamme, P.; Thompson, F.L.; et al. Re-Evaluating Prokaryotic Species. *Nat. Rev. Microbiol.* **2005**, *3*, 733–739, doi:10.1038/nrmicro1236.
4. Gomila, M.; Peña, A.; Mulet, M.; Lalucat, J.; García-Valdés, E. Phylogenomics and Systematics in *Pseudomonas*. *Front. Microbiol.* **2015**, *6*, doi:10.3389/fmicb.2015.00214.
5. Mulet, M.; Lalucat, J.; García-Valdés, E. DNA Sequence-Based Analysis of the *Pseudomonas* Species. *Environ. Microbiol.* **2010**, doi:10.1111/j.1462-2920.2010.02181.x.
6. Lalucat, J.; Mulet, M.; Gomila, M.; García-Valdés, E. Genomics in Bacterial Taxonomy: Impact on the Genus *Pseudomonas*. *Genes* **2020**, *11*, 139, doi:10.3390/genes11020139.
7. Sawabe, T.; Ogura, Y.; Matsumura, Y.; Feng, G.; Amin, A.R.; Mino, S.; Nakagawa, S.; Sawabe, T.; Kumar, R.; Fukui, Y.; et al. Updating the *Vibrio* Clades Defined by Multilocus Sequence Phylogeny: Proposal of Eight New Clades, and the Description of *Vibrio tritonius* sp. nov. *Front. Microbiol.* **2013**, *4*, doi:10.3389/fmicb.2013.00414.
8. Kim, M.; Oh, H.-S.; Park, S.-C.; Chun, J. Towards a Taxonomic Coherence between Average Nucleotide Identity and 16S rRNA Gene Sequence Similarity for Species Demarcation of Prokaryotes. *Int. J. Syst. Evol. Microbiol.* **2014**, *64*, 346–351, doi:10.1099/ijs.0.059774-0.
9. Meier-Kolthoff, J.P.; Auch, A.F.; Klenk, H.-P.; Göker, M. Genome Sequence-Based Species Delimitation with Confidence Intervals and Improved Distance Functions. *BMC Bioinformatics* **2013**, *14*, 60, doi:10.1186/1471-2105-14-60.
10. Meier-Kolthoff, J.P.; Klenk, H.-P.; Göker, M. Taxonomic Use of DNA G+C Content and DNA–DNA Hybridization in the Genomic Age. *Int. J. Syst. Evol. Microbiol.* **2014**, *64*, 352–356, doi:10.1099/ijs.0.056994-0.
11. Auch, A.F.; von Jan, M.; Klenk, H.-P.; Göker, M. Digital DNA–DNA Hybridization for Microbial Species Delineation by Means of Genome-to-Genome Sequence Comparison. *Stand. Genomic Sci.* **2010**, *2*, 117–134, doi:10.4056/sigs.531120.

12. Chaumeil, P.-A.; Mussig, A.J.; Hugenholtz, P.; Parks, D.H. GTDB-Tk: A Toolkit to Classify Genomes with the Genome Taxonomy Database. *Bioinformatics* **2019**, btz848, doi:10.1093/bioinformatics/btz848.
13. Parks, D.H.; Chuvochina, M.; Waite, D.W.; Rinke, C.; Skarshewski, A.; Chaumeil, P.-A.; Hugenholtz, P. A Standardized Bacterial Taxonomy Based on Genome Phylogeny Substantially Revises the Tree of Life. *Nat. Biotechnol.* **2018**, *36*, 996–1004, doi:10.1038/nbt.4229.
14. Moore, E.R.B.; Mau, M.; Arnscheidt, A.; Böttger, E.C.; Hutson, R.A.; Collins, M.D.; Van De Peer, Y.; De Wachter, R.; Timmis, K.N. The Determination and Comparison of the 16S rRNA Gene Sequences of Species of the Genus *Pseudomonas* (Sensu Stricto and Estimation of the Natural Intrageneric Relationships. *Syst. Appl. Microbiol.* **1996**, *19*, 478–492, doi:10.1016/S0723-2020(96)80021-X.
15. Brosch, R.; Lefèvre, M.; Grimont, F.; Grimont, P.A.D. Taxonomic Diversity of Pseudomonads Revealed by Computer-Interpretation of Ribotyping Data. *Syst. Appl. Microbiol.* **1996**, *19*, 541–555, doi:10.1016/S0723-2020(96)80026-9.
16. Mulet, M.; García-Valdés, E.; Lalucat, J. Phylogenetic Affiliation of *Pseudomonas putida* Biovar A and B Strains. *Res. Microbiol.* **2013**, *164*, 351–359, doi:10.1016/j.resmic.2013.01.009.
17. Girard, L.; Lood, C.; Rokni-Zadeh, H.; van Noort, V.; Lavigne, R.; De Mot, R. Reliable Identification of Environmental *Pseudomonas* Isolates Using the *rpoD* Gene. *Microorganisms* **2020**, *8*, 1166, doi:10.3390/microorganisms8081166.
18. Hesse, C.; Schulz, F.; Bull, C.T.; Shaffer, B.T.; Yan, Q.; Shapiro, N.; Hassan, K.A.; Varghese, N.; Elbourne, L.D.H.; Paulsen, I.T.; et al. Genome-Based Evolutionary History of *Pseudomonas* spp. *Environ. Microbiol.* **2018**, *20*, 2142–2159, doi:10.1111/1462-2920.14130.
19. Höfte, M.; De Vos, P. Plant pathogenic *Pseudomonas* species. In *Plant-Associated Bacteria*; Gnanamanickam, S.S., Ed.; Springer Netherlands: Dordrecht, 2007; pp. 507–533 ISBN 978-1-4020-4536-3.
20. Wiklund, T. *Pseudomonas anguilliseptica* Infection as a Threat to Wild and Farmed Fish in the Baltic Sea. *Microbiol. Aust.* **2016**, *37*, 135, doi:10.1071/MA16046.
21. Beaton, A.; Lood, C.; Cunningham-Oakes, E.; MacFadyen, A.; Mullins, A.J.; Bestawy, W.E.; Botelho, J.; Chevalier, S.; Coleman, S.; Dalzell, C.; et al. Community-Led Comparative Genomic and Phenotypic Analysis of the Aquaculture Pathogen *Pseudomonas baetica* A390T Sequenced by Ion Semiconductor and Nanopore Technologies. *FEMS Microbiol. Lett.* **2018**, *365*, doi:10.1093/femsle/fny069.
22. Wasi, S.; Tabrez, S.; Ahmad, M. Use of *Pseudomonas* spp. for the Bioremediation of Environmental Pollutants: A Review. *Environ. Monit. Assess.* **2013**, *185*, 8147–8155, doi:10.1007/s10661-013-3163-x.
23. Weller, D.M. *Pseudomonas* Biocontrol Agents of Soilborne Pathogens: Looking Back Over 30 Years. *Phytopathology* **2007**, *97*, 250–256, doi:10.1094/PHTO-97-2-0250.
24. Viggør, S.; Jøesaar, M.; Vedler, E.; Kiiker, R.; Pärnpuu, L.; Heinaru, A. Occurrence of Diverse Alkane Hydroxylase *alkB* Genes in Indigenous Oil-Degrading Bacteria of Baltic Sea Surface Water. *Mar. Pollut. Bull.* **2015**, *101*, 507–516, doi:10.1016/j.marpolbul.2015.10.064.
25. Gwon, H.-J.; Teruhiko, I.; Shigeaki, H.; Baik, S.-H. Identification of Novel Non-Metal Haloperoxidases from the Marine Metagenome. *J. Microbiol. Biotechnol.* **2014**, *24*, 835–842, doi:10.4014/jmb.1310.10070.
26. Mendes, R.; Kruijt, M.; de Bruijn, I.; Dekkers, E.; van der Voort, M.; Schneider, J.H.M.; Piceno, Y.M.; DeSantis, T.Z.; Andersen, G.L.; Bakker, P.A.H.M.; et al. Deciphering the Rhizosphere Microbiome for Disease-Suppressive Bacteria. *Science* **2011**, *332*, 1097–1100, doi:10.1126/science.1203980.
27. Gross, H.; Loper, J.E. Genomics of Secondary Metabolite Production by *Pseudomonas* spp. *Nat. Prod. Rep.* **2009**, *26*, 1408–1446, doi:10.1039/b817075b.
28. Götze, S.; Stallforth, P. Structure Elucidation of Bacterial Nonribosomal Lipopeptides. *Org. Biomol. Chem.* **2020**, *18*, 1710–1727, doi:10.1039/C9OB02539A.
29. Geudens, N.; Martins, J.C. Cyclic Lipopeptides From *Pseudomonas* spp. – Biological Swiss-Army Knives. *Front. Microbiol.* **2018**, *9*, doi:10.3389/fmicb.2018.01867.
30. Oni, F.E.; Geudens, N.; Omoboye, O.O.; Bertier, L.; Hua, H.G.K.; Adiobo, A.; Sinnaeve, D.; Martins, J.C.; Höfte, M. Fluorescent *Pseudomonas* and Cyclic Lipopeptide Diversity in the Rhizosphere of Cocoyam (*Xanthosoma sagittifolium*). *Environ. Microbiol.* **2019**, *21*, 1019–1034, doi:10.1111/1462-2920.14520.
31. Oni, F.E.; Geudens, N.; Onyeka, J.T.; Olorunleke, O.F.; Salami, A.E.; Omoboye, O.O.; Arias, A.A.; Adiobo, A.; De Neve, S.; Ongena, M.; et al. Cyclic Lipopeptide-producing *Pseudomonas koreensis* Group Strains Dominate the Cocoyam Rhizosphere of a *Pythium* Root Rot Suppressive Soil Contrasting with *P. putida* Prominence in Conductive Soils. *Environ. Microbiol.* **2020**, *1462-2920.15127*, doi:10.1111/1462-2920.15127.
32. Lood, C.; Peeters, C.; Lamy-Besnier, Q.; Wagemans, J.; De Vos, D.; Proesmans, M.; Pirnay, J.-P.; Echahidi, F.; Piérard, D.; Thimmesch, M.; et al. Genomics of an Endemic Cystic Fibrosis *Burkholderia multivorans* Strain Reveals Low Within-Patient Evolution but High between-Patient Diversity. *PLOS Pathog.* **2021**, *17*, e1009418, doi:10.1371/journal.ppat.1009418.
33. Wittouck, S.; Wuyts, S.; Meehan, C.J.; van Noort, V.; Lebeer, S. A Genome-Based Species Taxonomy of the *Lactobacillus* Genus Complex. *mSystems* **2019**, *4*, doi:10.1128/mSystems.00264-19.
34. Blin, K.; Shaw, S.; Kloosterman, A.M.; Charlop-Powers, Z.; van Wezel, G.P.; Medema, M.H.; Weber, T. AntiSMASH 6.0: Improving Cluster Detection and Comparison Capabilities. *Nucleic Acids Res.* **2021**, *49*, W29–W35, doi:10.1093/nar/gkab335.

35. Götze, S.; Stallforth, P. Structure, Properties, and Biological Functions of Nonribosomal Lipopeptides from *Pseudomonads*. *Nat. Prod. Rep.* **2019**, 10.1039/C9NP00022D, doi:10.1039/C9NP00022D.
36. Girard, L.; Höfte, M.; Mot, R.D. Lipopeptide Families at the Interface between Pathogenic and Beneficial *Pseudomonas*-Plant Interactions. *Crit. Rev. Microbiol.* **2020**, 1–23, doi:10.1080/1040841X.2020.1794790.
37. Morimoto, Y.; Tohya, M.; Aibibula, Z.; Baba, T.; Daida, H.; Kirikae, T. Re-Identification of Strains Deposited as *Pseudomonas aeruginosa*, *Pseudomonas fluorescens* and *Pseudomonas putida* in GenBank Based on Whole Genome Sequences. *Int. J. Syst. Evol. Microbiol.* **2020**, 70, 5958–5963, doi:10.1099/ijsem.0.004468.
38. Tohya, M.; Watanabe, S.; Tada, T.; Tin, H.H.; Kirikae, T. Genome Analysis-Based Reclassification of *Pseudomonas fuscovaginae* and *Pseudomonas shirazica* as Later Heterotypic Synonyms of *Pseudomonas asplenii* and *Pseudomonas asiatica*, Respectively. *Int. J. Syst. Evol. Microbiol.* **2020**, doi:10.1099/ijsem.0.004199.
39. Omoboye, O.O.; Oni, F.E.; Batool, H.; Yimer, H.Z.; De Mot, R.; Höfte, M. *Pseudomonas* Cyclic Lipopeptides Suppress the Rice Blast Fungus *Magnaporthe oryzae* by Induced Resistance and Direct Antagonism. *Front. Plant Sci.* **2019**, 10, 901, doi:10.3389/fpls.2019.00901.
40. Nguyen, D.D.; Melnik, A.V.; Koyama, N.; Lu, X.; Schorn, M.; Fang, J.; Aguinaldo, K.; Lincecum, T.L.; Ghequire, M.G.K.; Carrión, V.J.; et al. Indexing the *Pseudomonas* Specialized Metabolome Enabled the Discovery of Poaeamide B and the Bananamides. *Nat. Microbiol.* **2016**, 2, 16197, doi:10.1038/nmicrobiol.2016.197.
41. Vallet-Gely, I.; Novikov, A.; Augusto, L.; Liehl, P.; Bolbach, G.; Péchy-Tarr, M.; Cosson, P.; Keel, C.; Caroff, M.; Lemaitre, B. Association of Hemolytic Activity of *Pseudomonas entomophila*, a Versatile Soil Bacterium, with Cyclic Lipopeptide Production. *Appl. Environ. Microbiol.* **2010**, 76, 910–921, doi:10.1128/AEM.02112-09.
42. Pascual, J.; García-López, M.; Carmona, C.; Sousa, T. da S.; de Pedro, N.; Cautain, B.; Martín, J.; Vicente, F.; Reyes, F.; Bills, G.F.; et al. *Pseudomonas soli* sp. nov., a Novel Producer of Xantholysin Congeners. *Syst. Appl. Microbiol.* **2014**, 37, 412–416, doi:10.1016/j.syapm.2014.07.003.
43. Aiman, S.; Shehroz, M.; Munir, M.; Gul, S.; Shah, M.; Khan, A. Species-Wide Genome Mining of *Pseudomonas putida* for Potential Secondary Metabolites and Drug-Like Natural Products Characterization. *J. Proteomics Bioinform.* **2018**, 11, doi:10.4172/jpb.1000460.
44. Rokni-Zadeh, H.; Li, W.; Sanchez-Rodriguez, A.; Sinnaeve, D.; Rozenski, J.; Martins, J.C.; De Mot, R. Genetic and Functional Characterization of Cyclic Lipopeptide White-Line-Inducing Principle (WLIP) Production by Rice Rhizosphere Isolate *Pseudomonas putida* RW10S2. *Appl. Environ. Microbiol.* **2012**, 78, 4826–4834, doi:10.1128/AEM.00335-12.
45. Bernat, P.; Nesme, J.; Paraszkiwicz, K.; Schloter, M.; Płaza, G. Characterization of Extracellular Biosurfactants Expressed by a *Pseudomonas putida* Strain Isolated from the Interior of Healthy Roots from *Sida hermaphrodita* Grown in a Heavy Metal Contaminated Soil. *Curr. Microbiol.* **2019**, 76, 1320–1329, doi:10.1007/s00284-019-01757-x.
46. Kuiper, I.; Lagendijk, E.L.; Pickford, R.; Derrick, J.P.; Lamers, G.E.M.; Thomas-Oates, J.E.; Lugtenberg, B.J.J.; Bloemberg, G.V. Characterization of Two *Pseudomonas putida* Lipopeptide Biosurfactants, Putisolvin I and II, Which Inhibit Biofilm Formation and Break down Existing Biofilms. *Mol. Microbiol.* **2004**, 51, 97–113.
47. Dubern, J.-F.; Coppoolse, E.R.; Stiekema, W.J.; Bloemberg, G.V. Genetic and Functional Characterization of the Gene Cluster Directing the Biosynthesis of Putisolvin I and II in *Pseudomonas putida* Strain PCL1445. *Microbiol. Read. Engl.* **2008**, 154, 2070–2083, doi:10.1099/mic.0.2008/016444-0.
48. Li, W.; Rokni-Zadeh, H.; De Vleeschouwer, M.; Ghequire, M.G.K.; Sinnaeve, D.; Xie, G.-L.; Rozenski, J.; Madder, A.; Martins, J.C.; De Mot, R. The Antimicrobial Compound Xantholysin Defines a New Group of *Pseudomonas* Cyclic Lipopeptides. *PLoS One* **2013**, 8, e62946, doi:10.1371/journal.pone.0062946.
49. Molina-Santiago, C.; Udaondo, Z.; Ramos, J.-L. Draft Whole-Genome Sequence of the Antibiotic-Producing Soil Isolate *Pseudomonas* sp. Strain 250J. *Environ. Microbiol. Rep.* **2015**, 7, 288–292, doi:10.1111/1758-2229.12245.
50. Biessy, A.; Novinscak, A.; Blom, J.; Léger, G.; Thomashow, L.S.; Cazorla, F.M.; Josic, D.; Fillion, M. Diversity of Phytobeneficial Traits Revealed by Whole-Genome Analysis of Worldwide-Isolated Phenazine-Producing *Pseudomonas* spp. *Environ. Microbiol.* **2019**, 21, 437–455, doi:10.1111/1462-2920.14476.