RAFAELA DA SILVA RODRIGUES

Pseudomonas spp. CAUSING SPOILAGE IN BRAZILIAN CHEESES BY PIGMENT PRODUCTION

Thesis submitted to the Food Science and Technology Graduate Program of the Universidade Federal de Viçosa in partial fulfillment of the requirements for the degree of *Doctor Scientiae*.

Adviser: Luís Augusto Nero

Co-advisers: Antônio Fernandes de Carvalho
Solimar Gonçalves Machado

VIÇOSA – MINAS GERAIS
2022

Ficha catalográfica elaborada pela Biblioteca Central da Universidade Federal de Viçosa - Campus Viçosa

T
R696p
2022

Rodrigues, Rafaela da Silva, 1995-
Pseudomonas spp. causing spoilage in Brazilian cheeses by pigment production / Rafaela da Silva Rodrigues. – Viçosa, MG, 2022.
1 tese eletrônica (178 f.): il. (algumas color.).

Texto em inglês.
Inclui anexo.
Orientador: Luis Augusto Nero.
Tese (doutorado) - Universidade Federal de Viçosa, Departamento de Tecnologia de Alimentos, 2022.
Inclui bibliografia.
DOI: https://doi.org/10.47328/ufvbbt.2023.049
Modo de acesso: World Wide Web.

1. Alimentos - Microbiologia. 2. Queijos - Deterioração. 3. Pseudomonas. 4. Marcadores genéticos. 5. Pigmentos. I. Nero, Luis Augusto, 1975-. II. Universidade Federal de Viçosa. Departamento de Tecnologia de Alimentos. Programa de Pós-Graduação em Ciência e Tecnologia de Alimentos. III. Título.

CDD 22. ed. 664.0001579

Bibliotecário(a) responsável: Bruna Silva CRB-6/2552

RAFAELA DA SILVA RODRIGUES

Pseudomonas spp. CAUSING SPOILAGE IN BRAZILIAN CHEESES BY PIGMENT PRODUCTION

Thesis submitted to the Food Science and Technology Graduate Program of the Universidade Federal de Viçosa in partial fulfillment of the requirements for the degree of *Doctor Scientiae*.

APPROVED: December 19, 2022.

Assent:

Documento assinado digitalmente
gov.br RAFAELA DA SILVA RODRIGUES
Data: 28/02/2023 10:54:10-0300
Verifique em https://verificador.iti.br

Rafaela da Silva Rodrigues
Author

Documento assinado digitalmente
gov.br LUIS AUGUSTO NERO
Data: 28/02/2023 10:26:58-0300
Verifique em https://verificador.iti.br

Luís Augusto Nero
Adviser

ACKNOWLEDGMENT

A Deus por me conceder o Dom da Vida, sabedoria e oportunidade de chegar até aqui.

Aos meus familiares, por todo apoio e torcida ao longo dessa caminhada.

Ao meu orientador prof. Luís Augusto Nero, e meus coorientadores prof. Antônio Fernandes de Carvalho e profa. Solimar Gonçalves Machado, pelos ensinamentos, paciência e confiança. Essa pesquisa não teria sido possível sem a excelente orientação que tive ao longa dessa jornada!

Ao grupo InovaLeite, por me receber e permitir executar experimentos junto a equipe.

Aos colegas e amigos do InsPOA e Inovaleite, pela amizade, apoio e incentivo nesses anos de convívio.

À Universidade Federal de Viçosa (UFV) e ao Departamento de Tecnologia de Alimentos, pela oportunidade de realização do treinamento no Programa de Pós Graduação em Ciência e Tecnologia de Alimentos.

A todos os professores, funcionários e técnicos que direta ou indiretamente contribuíram para a realização dos trabalhos. Em especial aos técnicos da Diretoria de Tecnologia de Informação (DTI) da UFV, que sempre estiveram disponíveis para instalação de softwares no *cluster* da UFV, necessários para execução da pesquisa, ou auxílio durante problemas de acesso ao sistema.

À Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), pela concessão da bolsa de estudos.

This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) – Finance Code 001.

ABSTRACT

RODRIGUES, Rafaela da Silva, D.Sc., Universidade Federal de Viçosa, December, 2022. ***Pseudomonas* spp. causing spoilage in Brazilian cheeses by pigment production**. Adviser: Luís Augusto Nero. Co-advisers: Antônio Fernandes de Carvalho and Solimar Gonçalves Machado.

Cheeses are foods prone to the growth of spoilage microorganisms, such as *Pseudomonas*: one of the most complex and diverse genera of Gram-negative bacteria, which makes it difficult to correctly assign strains to the species. Reports in recent years have shown that some strains belonging to this genus are capable of producing a blue pigment that is diffusible in dairy products with a high moisture content, in cold storage. Therefore, research has tried to elucidate aspects related to the production of this pigment, as well as the genomic region related to this phenotype, however there are still contradictions between their results. Here, we identified blue pigment-producing *Pseudomonas* strains isolated from spoiled samples of Minas Frescal cheese and processed cheese "requeijão em barra", and used whole-genome sequencing (WGS) and bioinformatics tools to aid in the studies. For the first time, *Pseudomonas* strains that cause blue discoloration in Brazilian cheeses were investigated and identified. We found *Pseudomonas carnis* and *Pseudomonas paracarnis* as pigment producers, with rapid production at refrigerated temperatures and at 25 °C. Strains obtained here are phylogenetically related to others already investigated in other countries, and bioinformatics analysis helped to identify a genomic region of 74 genes, contained only in pigment producers, which includes tryptophan biosynthesis genes (*trpABCDF*) and genes for mobile genetic elements (MGE) such as genes involved in horizontal gene transfer (e. g., *pilM*, *pilV*, *pilX*, *tadA*, *traD*). This region has also been identified by bioinformatics tools as an integrative and conjugative element (ICE), due to components related to conjugative-transposon-like mobile genetic elements and it has plasmid-like conjugative properties. Furthermore, a phylogenomic investigation of strains confirmed as *P. carnis*, deposited in GenBank, showed that there is no evolutionary relationship between those that are pigmenting and carriers of duplicated *trpABCDF* genes, all of which are isolated only from food. With this, it appears that the recent spoilage in food reported in different countries may not have been caused by the spread of a single strain. Therefore, studies like this help to understand an issue that is affecting different nationalities, as well as providing a basis for the development of markers that will allow the tracking and control of these spoilers in food.

In addition, it shows that a correct bacterial identification at the species level is essential for a reliable execution of investigations, when it comes to *Pseudomonas*.

Keywords: *Pseudomonas*. Spoilage. Blue discoloration. Genomic analysis.

RESUMO

RODRIGUES, Rafaela da Silva, D.Sc., Universidade Federal de Viçosa, dezembro de 2022. ***Pseudomonas* spp. causadoras de deterioração em queijos brasileiros por produção de pigmento**. Orientador: Luís Augusto Nero. Coorientadores: Antônio Fernandes de Carvalho e Solimar Gonçalves Machado.

Queijos são alimentos propensos a multiplicação de microrganismos deterioradores, como *Pseudomonas*: um dos gêneros mais complexos e diversos de bactérias Gram-negativas, o que dificulta a correta atribuição de cepas à espécie. Relatos nos últimos anos tem mostrado que algumas cepas pertencentes a esse gênero são capazes de produzir um pigmento azul difusível em produtos lácteos de elevado teor de água, sob refrigeração. Portanto, pesquisas têm buscado elucidar aspectos relacionados a produção desse pigmento, bem como a região genômica relacionada a esse fenótipo, entretanto ainda há contradições entre os resultados. Aqui identificamos cepas de *Pseudomonas* produtoras de pigmento azul isoladas de amostras deterioradas de queijo Minas Frescal e requeijão em barra, e utilizamos sequenciamento completo de genoma (WGS) e ferramentas de bioinformática para auxiliar nos estudos. Pela primeira vez, cepas de *Pseudomonas* causadoras de pigmentação azul em queijos no Brasil foram investigadas e identificadas. Constatamos *Pseudomonas carnis* e *Pseudomonas paracarnis* como produtores do pigmento, com rápida produção em temperaturas de refrigeração e a 25 °C. Cepas obtidas aqui são filogeneticamente relacionadas com outras já investigadas em outros países, e análises de bioinformática auxiliaram na identificação de uma região genômica de 74 genes, contida apenas em produtores de pigmento, que inclui genes de biossíntese de triptofano (*trpABCDF*) e genes de elementos genéticos móveis (MGE), como os envolvidos na transferência horizontal de genes (por exemplo, *pilM*, *pilV*, *pilX*, *tadA*, *traD*). Essa região também foi identificada por ferramentas de bioinformática como elemento integrativo conjugativo (ICE), devido aos componentes relacionados a transposon conjugativo e de propriedades conjugativas do tipo plasmídeo. Além disso, uma investigação filogenômica de cepas confirmadas como *P. carnis* depositadas no GenBank mostrou que não há relação evolutiva entre aquelas que são pigmentantes, e portadoras dos genes duplicados de *trpABCDF*, sendo todas isoladas apenas de alimentos. Com isso, se verifica que as recentes deteriorações em alimentos constatadas em diferentes países podem não ter sido causadas por disseminação de uma única cepa. Portanto, estudos como esses auxiliam na compressão de um problema que está afetando diversas nacionalidades, assim como fornece fundamento para o desenvolvimento

de marcadores que permitirão o rastreamento e controle desses deterioradores em alimentos. Além disso, mostra que uma correta identificação bacteriana a nível de espécie é essencial para uma execução confiável de investigações, quando se trata de *Pseudomonas*.

Palavras-chave: *Pseudomonas*. Deterioração. Pigmentação azul. Análise genômica.

SUMMARY

GENERAL INTRODUCTION

Bacteria belonging to the genus *Pseudomonas* are often associated with food spoilage. In dairy products, this spoilage is generally attributed to the production of proteolytic and lipolytic enzymes, however, another type of spoilage has also been linked to contamination by *Pseudomonas*: the appearance of a blue discoloration. Since 2010, many cases of blue discoloration in some cheeses, mainly produced in Europe and the United States, have been reported and official analyzes have linked this to contamination of the cheese with strains of *Pseudomonas fluorescens*. The frequently reported episodes were a serious problem for producers in the dairy sector and, consequently, studies began to investigate and detect the causative agents of blue discoloration in cheeses.

The emergence of this blue pigment in cheese is a phenomenon usually recorded in fresh, high-moisture cheeses stored under refrigeration, but it is also reported in meat products. However, until the present moment, there are divergences regarding the results of investigations regarding the unique genes of these bacteria that could, therefore, be involved in the expression of this phenotype. Distinct genes have been associated, such as: genes of the tryptophan biosynthesis pathway; genes involved in iron uptake; and, defective prophage structure genes. It is worth mentioning that the use of whole-genome sequencing (WGS) has helped in these investigations, in addition to the fact that these sequences can also be used in the general genomic characterization of these microorganisms and help in correct bacterial taxonomic classifications, especially of genus *Pseudomonas*, one of the most diverse and complex genera.

Here we isolated, phenotypically analyzed, and sequenced the genome of *Pseudomonas* causing blue discoloration in Brazilian cheeses (Minas Frescal cheese and processed cheese "requeijão em barra"). The present work showed that two species, *P. carnis* and *P. paracarnis*, are involved with this spoilage in cheeses, as well as trying to evaluate the evolutionary

relationship between *Pseudomonas* producing food discoloration. The characterization of genomic regions of interest to the food industry makes it possible to stimulate advances in detection methods, in addition to inducing the improvement of manufacturing practices, when strains constantly found in foods belong to species that generally carry genomic regions related to spoilage.

OBJECTIVES

The aim of this study was to investigate the cause of a blue discoloration in Brazilian cheeses. Considering the main goal, specific objectives were:

- Isolate and phenotypically identify blue pigment-producing *Pseudomonas* from Minas Frescal cheese and processed cheese “requeijão em barra”.
- Sequence and characterize the whole genome of the isolates.
- Identify to species level *Pseudomonas* causing the blue discoloration.
- Compare genomic sequences with a database of strains previously involved in cases of blue discoloration.
- Identify the existence of a conserved genomic region.
- Verify the existence of a branch, by phylogenomic analysis, of strains that cause blue discoloration within the species most related to this spoilage.

LITERATURE REVIEW

1. Blue discolored foods

1.1 Reports

In recent years, there has been an increase in research to identify the causes of an anomalous blue discoloration in different foods, but mainly in dairy products. Strains of the genus *Pseudomonas*, in particular, those belonging to the *Pseudomonas fluorescens* group have been associated with the occurrence of blue pigments in cheeses (Andreani et al., 2014; Caputo et al., 2015; Carrascosa et al., 2015; Chierici et al., 2016; Martin et al., 2011). Investigations into the production of these pigments in dairy products increased after a large number of reports about blue discoloration on Mozzarella cheese in Europe, in 2010. This event was reported by the Rapid Alert System for Food and Feed (RASFF), in the 2010 annual report, and became known as "blue mozzarella" (RASFF, 2010). Italian authorities have reported a problem with Mozzarella cheese from Germany because consumers complained that the product turned blue after the packaging was opened. After investigations, they identified *P. fluorescens* as the causative agent and the water used in processing environments as a source of contamination.

The occurrence of this pigment was more observed in fresh cheese, such as Mozzarella cheese (Caputo et al., 2015; Carminati et al., 2019; Sechi et al., 2011), Latin-style fresh cheese (Martin et al., 2011), and fresh cheeses in Gran Canaria (Carrascosa et al., 2015). Reports on other foods have also emerged, such as: soymilk and tofu (De León et al., 2021), minced beef (Lick et al., 2021), and rabbit carcasses (Circella et al., 2020). Overall, most isolates were identified as belonging to the *P. fluorescens* group. More details about these investigations can be seen in **Table 1**.

Table 1. General information of the main investigations of *Pseudomonas* related with blue discolored foods.

Identification	Isolation			Identification method	References
	Date	Source	Country		
P. fluorescens biovar IV	2006	Latin-style fresh cheese	USA	16S rRNA, API20NE, Biolog GN2 system	(Martin et al., 2011)
P. fluorescens	2010	Mozzarella cheeses and whey cheese products	Italy and Germany	—	(Nogarol et al., 2013)
P. fluorescens group	2010-2012	Blue mozzarella cheese and spoiled meat	Germany, Italy, England	API20E, API20NE, 16S rRNA	(Andreani et al., 2014)
P. fluorescens biovar IV	—	Fresh cheeses	Spain	16S rRNA	(Carrascosa et al., 2015)
P. fluorescens group	2010-2014	Spoiled fresh cheese	Italy and Germany	16S rRNA	(Chierici et al., 2016)
P. fluorescens group	—	Cheese factory	Italy	16S rRNA	(Carminati et al., 2019)
*P.*cf. *lactis**	2006, 2015, 2017, 2018	Spoiled dairy products	USA	WGS	(Reichler et al., 2019)
P. azotoformans	—	Pigmented rabbit carcasses	Italy	16S rRNA	(Circella et al., 2020)
P. carnis	2007 and 2014	Raw turkey meat and spoiled pork	Germany	16S rRNA, *rpoB*, WGS, phenotypic and chemotaxonomic characterizations	(Lick et al., 2020)
P. paracarnis	2017	Minced beef	—	16S rRNA, *rpoB*, WGS, phenotypic and chemotaxonomic characterizations	(Lick et al., 2021)
P. carnis	2020	Spoiled soymilk and tofu	USA	16S rRNA and WGS	(León et al., 2021)
P. fluorescens group	—	Cheese farm	Spain	16S rRNA and MALDI-TOF	(Carrascosa et al., 2021)

The symbol "—" indicated uninformed data.

* The authors used the designation cf. (confer) because these isolates are closely related to the *P. lactis* type strain but border the currently accepted cutoff for speciation by whole genome average nucleotide identity (Reichler et al., 2019).

1.2 Conditions of appearance in cheeses

Most foods that are observed to develop blue discoloration had been stored under refrigeration. This observation has been exploited by researchers to assist in phenotypic identification of isolates, and it is generally observed that the pigment is related to low temperatures. Chierici et al. (2016) studied the pigment production in Mozzarella preserving fluid (a brine with a pH 6.3; water, salt and cheese minerals) at three different incubation temperatures (4, 14 and 30 °C). The blue pigment did not appear in the brine incubated at 30 °C, but at 4 °C it appeared after 10 days of incubation, and at 14 °C after 72 h. In addition to the relationship between pigment production and temperature, Chierici et al. (2016) also showed that there is an association with carbon sources. The researchers observed the growth of the strains in M9 minimal medium (pH 6.3) with different carbon sources: D-glucose, D-galactose, sodium succinate, sodium citrate, and DL-lactic acid sodium salt. Blue pigment production was observed only when glucose was present in the medium (at 4 and 14 °C, but not at 30 °C). For other carbon sources, the strains generated other colors (usually green or gray).

Chierici (2016) demonstrated that there are other nutritional requirements for pigment production. The tests were carried out in M9 medium, with 0.2% (w/v) glucose, evaluating nine metals (Mo, Cu, Zn, Ca, Mg, Bo, Co, Mn, Fe) and 18 amino acids (tyrosine, isoleucine, glycine, valine, histidine, lysine, glutamine, proline, alanine, phenylalanine, glutamic acid, asparagine, arginine, serine, tryptophan, leucine, cysteine, threonine), at 4 °C. The research demonstrated that the presence of cobalt (Co) or lysine inhibited the synthesis of the blue pigment, while the most intense blue color was obtained when proline was added. However, there was not a single amino acid responsible for pigment production, and each strain studied had a different intensity of the pigment.

Another important factor that has been evidenced for the production of pigments is the concentration of pigmenting bacteria in the milk matrix studied. Rossi et al. (2018), when artificially contaminating Mozzarella cheese with strains of *Pseudomonas*, observed that the blue pigment appeared after the count reached values between 6.5 and 7.5 CFU/g (10 °C after 3 d). Andreani et al. (2014) also artificially contaminated Mozzarella cheese, with the preserving fluid, and observed that at the end of the test the concentration of *Pseudomonas* reached at least 10^8 CFU/g. The fresh cheeses produced in Gran Canaria with blue discoloration studied by Carrascosa et al. (2015) had a count above 10^5 CFU/g. In addition, there are milk contamination tests with these strains, in which gray discoloration occurs only after the concentration of cells reaches 10^8 CFU/mL (Reichler et al., 2019).

The appearance of blue pigment in Mozzarella cheese has also been associated with oxygen concentration, mainly after the packaging is opened by the consumer and the cheese comes into contact with the air (RASFF, 2010). Cenci-Goga et al. (2014) evaluated the appearance of blue pigment in intentionally contaminated Mozzarella cheese, under different storage conditions. They showed that for packages stored at 8 °C and opened after 8d of storage, the blue color appeared less than 1h after opening the package. The researchers also noted that the longer the cheese is stored after production, the faster the blue color appears after opening the package. For packages stored in temperature abuse (20 °C), the blue color was also apparent in the closed package, but only after 8d of storage.

Reichler et al. (2019) similarly showed a strong relationship between the presence of oxygen and the production and/or intensity of pigments, when studying two strains already classified as producing blue and/or gray pigments in dairy products. They did tests regarding the production of pigments in PDA (potato dextrose agar) and Mozzarella cheese and also in pasteurized milk, in filled and half-filled containers. The two strains studied produced different pigments in PDA (5 d incubation at 6 °C): one produced reddish-brown pigment and the other

blue pigment. However, they showed the same phenotype when incubated in cheese and milk, producing blue and gray pigments, respectively. In milk, discoloration was observed only on the upper part of half-filled containers, showing a possible correlation between the appearance of the pigment and the presence of oxygen.

1.3 Blue pigment biosynthesis pathway

In the research carried out to date, the results found demonstrate that the production of the blue pigment by *Pseudomonas* is mainly related to the tryptophan biosynthesis pathway (*trp* genes) and also metabolic pathways of other pigments. Andreani et al. (2019) used transposon mutagenesis and characterized the mutants to identify genes involved in the production of blue pigment by *P. fluorescens* ps_77. Genetic analyzes based on the mapping of the random insertions allowed to identify genes involved in pigment production, including genes strictly related to *trp* genes. In this study, the researchers identified a total of 14 co-expressed genes, included in a large operon, involved in the production of pigment, which are possibly genes for encoding enzymes for pigment biosynthesis and related functions, such as secretion and transport.

In a previous work, there is a contradiction regarding the genes involved in regulating the production of this pigment. Andreani et al. (2015) concluded that, in strains that produce blue pigment, the genes involved in iron uptake and in the biosynthesis of the siderophore pyoverdine were more expressed. These authors raised the hypothesis that iron would therefore be a cofactor for enzymes involved in pigment production, but it was not clear whether the main requirement for iron was due to the need to produce the blue molecule or whether the blue pigment produced by the strains could be useful in uptake iron.

In this study by Andreani et al. (2015), the production of blue pigment by *P. fluorescens* was investigated using four phylogenetically related strains, belonging to the so-called "blue branch", but with different phenotypes: two producers of blue pigment (ps_22 and ps_77) and two non-producers (ps_20 and ps_40). This distinction was used mainly to identify, using transcriptomic analysis, the differential expression of genes that could be related to pigment production. The global expression profiling of pigment-producing strains versus non-producing strains showed a general up-regulation of genes involved in iron uptake.

Andreani et al. (2015) also used pangenome analysis among the four strains studied and identified in the pigmenting strains the presence of two copies of genes involved in the tryptophan biosynthesis pathway. The presence of double copies of the *trp* genes demonstrated a possible relationship with pigment production. Thus, this biosynthesis pathway could be used to form an indole-derivative. However, another genomic analysis showed the absence of the tryptophanase (*tnaA*) gene that encodes the enzyme responsible for converting tryptophan to indole (observed in indole-producing bacteria). Comparing the results of Andreani et al. (2015) and Andreani et al. (2019), it is possible to notice that the most recent study carried out by the group of Italian researchers contradicts the results of the work published in 2015, confirming that the pigment production is related to the *trp* genes.

Reichler et al. (2019) investigated the genomes of nine strains of *Pseudomonas* isolated in the United States, producing blue and gray pigment in cheese and milk, carrying out a study with whole genome sequencing. In addition to the strains in question, the authors also evaluated the genomic sequences of the four strains studied by the Italian group, two of which are pigmented (ps_20, ps_40, ps_22 and ps_77; Andreani et al. (2015)). The comparison between the genomes showed that there is an accessory locus, present in all blue and gray pigment-producing isolates. These strains had two copies of the *trp* operon genes, while the non-producing strains had a single copy of each gene. Thus, there may be conditions in which this

additional operon is more easily regulated in order to enable the synthesis of tryptophan, which does not lead to the expression of the broadly distributed operon.

Another evidence of this study is that the exclusive locus in the pigmented isolates includes the *rhtB* gene, which encodes for an L-homoserine and L-homoserine lactone exporter. It is possible that these molecules serve as a signal for pigment biosynthesis through quorum sensing, and there is no other gene within the locus that appears to be involved in gene regulation (Reichler et al., 2019).

Thus, the results found by Reichler et al. (2019) and Andreani et al. (2015, 2019) indicate molecular targets, such as the accessory genes of tryptophan biosynthesis, which can be used to develop initial molecular tests for the identification of blue or gray pigment-producing *Pseudomonas* strains. Reichler et al. (2019) supports the idea that pigment-producing strains are part of a new species that includes several distinct strains that can cause blue and/or gray color defects in dairy products. However, Andreani et al. (2014) showed that not only strains that cause pigment in dairy products, but also pigment strains from beef and pork, were genetically related, by MLST (Multilocus Sequence Typing) analysis. And, in this same analysis, it was shown that pigment strains in cheese, but with a different blue phenotype, were included in another cluster. Therefore, there is a need for further studies that explore both the classification of these isolates and the real pathway of pigment biosynthesis.

There are also other divergences between researches that have studied pigment-producing *Pseudomonas* isolates. Martin et al. (2011) isolated a producer of blue pigment in cheese and was identified with *P. fluorescens* biovar IV. Andreani et al. (2014) included this biovar IV (already named *Pseudomonas lemonnieri*, DSM 50415) in the development of the *P. fluorescens* MLST. This strain was not grouped in the cluster of strains producing blue pigment. However, Reichler et al. (2019) included this strain in the study, and identified that, genetically,

it also had the exclusive locus and identified a second *trp* operon, therefore it also has double copies of genes involved in tryptophan synthesis.

Reichler et al. (2019) investigated using whole-genome sequencing other pathways that could be associated with pigment production. They did a genome analysis for families or domains of target proteins that could be involved in pigment biosynthesis (pyoverdine, pyocianin, tryptophan, indigoidine, violacein and blue copper-binding proteins). However, the tests used failed to detect proteins present only in pigmented strains and absent in non-pigmented strains that could be associated with pigment production.

Chierici (2016) also investigated the genome of blue pigment producing strains. Initially, the complete genome of three strains producing blue pigment was sequenced and then a comparative genomic analysis was performed with the genome of two strains with the same phenotype (again ps_77 and ps_22 of Andreani et al., (2014)) and with other non-pigmenting, to identify the region of DNA encoding the blue phenotype. A total of 24 genes were recognized as specific for the five pigment-producing strains. And, among these genes, the researcher evidenced the presence of genes from bacteriophage structures. Thus, a hypothesis was raised: the sequence shared only among *Pseudomonas* producing blue pigment could originate from a bacteriophage integrated into the bacterial genome of an ancestral strain, which lost some of its functional genes becoming a defective prophage. After identifying this region, induction of phage production was investigated and performed. Although the researcher was not successful in isolating the phages, it was possible to visualize two bacteriophage morphologies by transmission electron microscopy.

1.4 Identification techniques of blue producers

Researches have been developed with the objective of creating methodologies or tests that allow classifying or facilitating the identification of blue pigment-producing strains. Andreani et al., (2014) developed a Multilocus Sequence Typing (MLST) scheme to reveal genetic relationships and identify a possible connection with specific phenotypic characteristics between strains belonging to the *P. fluorescens* group isolated from different food products. Among the strains selected for the study were strains related to blue discoloration in Mozzarella cheese. The study was successful in developing the online MLST database, specific to the *P. fluorescens* group (https://pubmlst.org/pfluorescens/), thus facilitating the application of this scheme and data sharing. In addition, from the association of this scheme with data from phenotypic tests, which evaluated the production of blue pigments from the strains, it revealed that the pigment producers were grouped into a single phylogenetic group, entitled "blue branch", which also contained some non-pigmented strains. Therefore, the MLST genetic analysis revealed the presence of a connection between the production of blue pigment in a specific phylogenetic cluster (Andreani et al., 2014).

Chierici et al. (2016) also used this MLST scheme in strains of the *P. fluorescens* group that had blue pigment production in dairy products. Four new allelic profiles were discovered, and the new sequences were grouped within the cluster formed by the blue pigment-producing strains. Carminati et al. (2019) used this MLST protocol in a *Pseudomonas* strain isolated from a dairy in Italy previously involved in cases of Mozzarella cheese with blue discoloration. The researchers found a new allelic profile, and the strain was also included in the phylogenetic cluster "blue branch". Corroborating the hypothesis that there is a connection between the production of this pigment and a specific cluster within the *P. fluorescens* group.

Despite this, a blue pigment-producing strain that was isolated and identified by Martin et al. (2011) and inserted in study of Andreani et al. (2014) for the development of the MLST scheme, was not grouped in the “blue branch”. However, this difference shown by the MLST analysis is also observed when comparing the phenotypic characteristics of the strains regarding the production of blue pigments. This distinct strain, identified in the initial study as *P. fluorescens* biovar IV, produced a light blue pigment that is not diffusible on PDA (Potato Dextrose Agar) (Martin et al., 2011), and can be produced in this same medium at different temperatures (6, 22 and 31 °C) (Andreani et al., 2014). Unlike this strain, the isolates belonging to the "blue branch" produce a blue pigment diffusible at 6 and 22 °C, but not at 30 or 31 °C (Andreani et al., 2014; Rossi et al., 2018). This shows that there are genetic differences in strains related to the production of blue pigment in cheese.

Regarding species-level identification, most work to date has mainly employed only the 16s gene sequencing (**Table 1**). This method has allowed researchers to associate many of the isolates with the *P. fluorescens* group. Differently, Circella et al., (2020) using only the 16s rRNA sequence stated that blue pigment in rabbit meat was caused by *P. azotoformans*. However, the most reliable identification regarding the species has been using whole genomic sequencing, and subsequent use of adequate tools for analysis. This is how Lick et al. in 2020 and 2021 noticed that strains involved in the production of pigments in meat were part of new species that had not yet been described, *P. carnis* and *P. paracarnis*, respectively.

2. The genus *Pseudomonas*

Pseudomonas are bacteria naturally widespread in the environment. They are rod-shaped, aerobic, Gram-negative, non-spore forming, catalase positive, and motile by one or

several polar flagella – rarely nonmotile. Some species are pathogenic for humans, animals, or plants, and several members produce pigments such as phenazines (pyocyanin, pyorubin, chlororaphin, oxiphenazin), blue compounds, or pyoverdins, yellow-green fluorescent siderophore. Furthermore, the best growth temperature of most strains is approximately 28 °C. However, some species can grow at 4 °C, and so considered psychrotrophic (Palleroni, 2005).

Pseudomonas is one of the most diverse and complex genera, as the genus of Gram-negative bacteria with the highest number of validated species. As of January 2022, the List of Prokaryotic Names with Standing in Nomenclature (LPSN) recognized as validly published under the ICNP (International Code of Nomenclature of Prokaryotes) and correct name 309 species and 10 subspecies in the *Pseudomonas* genus (https://lpsn.dsmz.de/genus/pseudomonas). The numbers of species have increased considerably. In 2020 and 2021 the number of new species described were 29 and 34, respectively, and so far in 2022 there were 68 new species.

It is constantly observed that there is great diversity within this genus and, therefore, it has been subdivided into groups and subgroups, according to evolutionary relationships between species. This classification has been established by researches that try to infer the phylogeny of the genus based on gene sequences or complete genome sequences. Garrido-Sanz et al., (2016) divided the genus *Pseudomonas* into 14 groups, based on MLSA (Multilocus Sequence Analysis): *P. fluorescens* "complex", *P. asplenii*, *P. lutea*, *P. syringae*, *P. putida*, *P. anguilliseptica*, *P. straminea*, *P. oleovorans*, *P. stutzeri*, *P. pertucinogena*, *P. oryzihabitans*, *P. resinovorans*, *P. nitroreducens*, *P. aeruginosa*. The *P. fluorescens* "complex" is one of the most diverse groups within the genus. Garrido-Sanz et al., (2016) subsequently, based on MLSA and phylogenomic methods, subdivided the complex into 8 subgroups: *P. protegens*, *P. chlororaphis*, *P. corrugata*, *P. koreensis*, *P. jessenii*, *P. mandelii*, *P. gessardiie*, and *P.*

fluorescens. The *P. fluorescens* subgroup was recently described as containing 72 species and 3 subspecies (Lalucat et al., 2020).

The separation of these subgroups has been important in the design of classification methods using specific markers, in order to select strains that belong to the subgroups that contain more species associated with some characteristic of interest (Garrido-Sanz et al., 2017). As subgroups containing plant-associated bacteria with potential biotechnological applications in agriculture and environmental protection. Garrido-Sanz et al., (2016) show that strains belonging to *P. protegens*, *P. chlororaphis* and *P. corrugata* subgroups of *P. fluorescens* "complex" would be particularly suited for biocontrol applications, because they harbor biosynthetic clusters for compounds related to it.

However, inferring the phylogeny of the genus *Pseudomonas* has some obstacles, such as strains that had been isolated for several years and that have not been updated taxonomically by modern techniques. Gomila et al., (2015) confirmed that 30% of the whole-genome sequences of non-type strains available in the databases were not correctly assigned at the species level. Strains misidentified at the species level may affect, for example, subsequent studies that will use genic or genomic sequences, and may therefore propagate misinformation.

3. Genomic analyzes of *Pseudomonas*

Whole genomic sequences have often been indispensable for correctly assigning strains to *Pseudomonas* species. However, the approaches to delineate bacterial species need to be correctly selected. In overall, whole genomic sequences of the species-type strains are needed to apply a phylogenomic analysis in bacterial taxonomy, as they are the representatives of the species (Lalucat et al., 2020). With these sequences, it is possible to employ digital whole

genome comparisons. The best approaches to clearly separated strains of different *Pseudomonas* species based on genome sequences are: average nucleotide identities based on Blast (ANIb), genome-to-genome distance calculation (GGDC), and digital DNA-DNA hybridization (dDDH) (Garrido-Sanz et al., 2016; Gomila et al., 2015; Lalucat et al., 2020). However, there are approaches that cannot delimit *Pseudomonas* species even using the complete genome, such as tetranucleotide usage patterns (TETRA), which can only discriminate *Pseudomonas* from other genera (Gomila et al., 2015).

In addition to inferring the phylogeny of the genus *Pseudomonas*, the use of complete genomic sequences has aided in tracking several metabolic pathways and regions of interest in members of the genus. Garrido-Sanz et al. (2016) inferred the phylogenetic relationship between genomes of the *P. fluorescens* group, and also performed a screening for distinct characteristics revealed by the analysis of genomic sequences. The study identified genetic clusters related to biocontrol, bioremediation, siderophores, and others, that could allow applications of strains of species more related to these characteristics. Some *Pseudomonas* strains are related with biocontrol, based on plant protection by antagonizing plant-pathogenic microorganisms, and others are important to bioremediation by degrade or detoxify hazardous environmental contaminants (Silby, Winstanley, Godfrey, Levy, & Jackson, 2011). In addition, species containing biosynthetic clusters for siderophores, which sequester iron from the environment through high-affinity interactions, could have a competitive advantage in iron-limited environments (Hider & Kong, 2010).

Comparative genomic approaches also help in identifying and highlighting genetic evolutions, such as pathogenicity genes and resistance genes in *Pseudomonas*. The *in-silico* prediction of antibiotic resistance and virulence potential of *Pseudomonas aeruginosa*, a commonly multidrug-resistant human opportunistic pathogen, can be useful for clinical cases

that fail or challenge the limits of traditional laboratory tests, since this species is related to infections that are difficult to eradicate (Jeukens et al., 2019).

Studies of mobile genetic elements of this genus are also benefiting from the increasing accessibility of whole genome sequencing. Botelho et al., (2018) performed genomic analysis of carbapenem-resistant *P. aeruginosa* clinical isolates, from six geographically distant hospitals within Portugal, isolated over a two-decade period. They concluded that the spread of the carbapenem-resistance gene may had been by an integron associated with integrative conjugative elements. Mobile genetic elements are genetic material with the ability to move through the genome and from one genome to another (Singh et al., 2014). Therefore, whole-genome sequencing has become a tool for understanding the mechanisms involved in this process, not only within the *Pseudomonas* genus but also in others.

Investigations of members of this genus also aid food science in terms of spoilage ability. *Pseudomonas* are recognized as food spoilage bacteria mainly because many secrete hydrolytic enzymes, however the production of proteases is the most studied in terms of food spoilage. These enzymes are great concern in the dairy industry as it causes hydrolysis of milk proteins leading to defects like textural flaws by decomposition of milk components and off-flavors (Stoeckel et al., 2016; Zhang et al., 2020). In addition, some *Pseudomonas* produce heat-resistant enzymes that remain in fluid milk after heat treatment, like the metallopeptidase AprX (Zhang, Bijl, & Hettinga, 2018). The use of genomic analysis is allowing us to understand the distribution of the operon of this enzyme in *Pseudomonas* species, encoded by the first gene of the *aprX-lipA2* operon (Woods et al., 2001). Maier et al. (2020) studied the genetic variety of *aprX-lipA2* genetic organizations in the genus *Pseudomonas* using genome sequences and correlated it with proteolytic activity. A total of 22 different types of the operon genetic organizations were identified in the genus, and isolates harboring one specific type of operon

(Type 1; *aprXIDEF prtAB lipA2*) were highly proteolytic compared to strains with other operon types.

References

Andreani, N. A., Carraro, L., Martino, M. E., Fondi, M., Fasolato, L., Miotto, G., … Cardazzo, B. (2015). A genomic and transcriptomic approach to investigate the blue pigment phenotype in *Pseudomonas fluorescens*. *International Journal of Food Microbiology*, *213*, 88–98. https://doi.org/10.1016/j.ijfoodmicro.2015.05.024

Andreani, N. A., Carraro, L., Zhang, L., Vos, M., & Cardazzo, B. (2019). Transposon mutagenesis in *Pseudomonas fluorescens* reveals genes involved in blue pigment production and antioxidant protection. *Food Microbiology*, *82*, 497–503. https://doi.org/10.1016/j.fm.2019.03.028

Andreani, N. A., Martino, M. E., Fasolato, L., Carraro, L., Montemurro, F., Mioni, R., … Cardazzo, B. (2014). Tracking the blue: A MLST approach to characterise the *Pseudomonas fluorescens* group. *Food Microbiology*, *39*, 116–126. https://doi.org/10.1016/j.fm.2014.11.009

Botelho, J., Grosso, F., Quinteira, S., Brilhante, M., Ramos, H., & Peixe, L. (2018). Two decades of blaVIM-2-producing *Pseudomonas aeruginosa* dissemination: an interplay between mobile genetic elements and successful clones. *Journal of Antimicrobial Chemotherapy*, *73*(4), 873–882. https://doi.org/10.1093/jac/dkx517

Caputo, L., Quintieri, L., Bianchi, D. M., Decastelli, L., Monaci, L., Visconti, A., & Baruzzi, F. (2015). Pepsin-digested bovine lactoferrin prevents Mozzarella cheese blue

discoloration caused by *Pseudomonas fluorescens*. *Food Microbiology*, *46*, 15–24. https://doi.org/10.1016/j.fm.2014.06.021

Carminati, D., Bonvini, B., Rossetti, L., Zago, M., Tidona, F., & Giraffa, G. (2019). Investigation on the presence of blue pigment-producing *Pseudomonas* strains along a production line of fresh mozzarella cheese. *Food Control*, *100*, 321–328. https://doi.org/10.1016/j.foodcont.2019.02.009

Carrascosa, C., Martínez, R., Sanjuán, E., Millán, R., del Rosario-Quintana, C., Acosta, F., … Jaber, J. R. (2021). Identification of the *Pseudomonas fluorescens* group as being responsible for blue pigment on fresh cheese. *Journal of Dairy Science*, *104*(6). https://doi.org/10.3168/jds.2020-19517

Carrascosa, C., Millán, R., Jaber, J. R., Lupiola, P., del Rosario-Quintana, C., Mauricio, C., & Sanjuán, E. (2015). Blue pigment in fresh cheese produced by *Pseudomonas fluorescens*. *Food Control*, *54*, 95–102. https://doi.org/10.1016/j.foodcont.2014.12.039

Cenci-Goga, B. T., Karama, M., Sechi, P., Iulietto, M. F., Novelli, S., & Mattei, S. (2014). Evolution under different storage conditions of anomalous blue coloration of Mozzarella cheese intentionally contaminated with a pigment-producing strain of *Pseudomonas fluorescens*. *Journal of Dairy Science*, *97*, 6708–6718. https://doi.org/10.3168/jds.2014-8611

Chierici, M. (2016). *Investigation on the blue phenotype in Pseudomonas species involved in blue discoloration defect of fresh cheese*. Università degli studi di Milano.

Chierici, M., Picozzi, C., La Spina, M. G., Orsi, C., Vigentini, I., Zambrini, V., & Foschino, R. (2016). Strain diversity of *Pseudomonas fluorescens* group with potential blue pigment phenotype isolated from dairy products. *Journal of Food Protection*, *79*(8), 1430–1435. https://doi.org/10.4315/0362-028X.JFP-15-589

Circella, E., Schiavone, A., Barrasso, R., Camarda, A., Pugliese, N., & Bozzo, G. (2020). *Pseudomonas azotoformans* Belonging to *Pseudomonas fluorescens* Group as Causative Agent of Blue Coloration in Carcasses of Slaughterhouse Rabbits. *Animals*, *10*(2), 256. https://doi.org/10.3390/ani10020256

De León, M. E., Wilson, H. S., Jospin, G., & Eisen, J. A. (2021). Draft genome sequences and genomic analysis for pigment production in bacteria isolated from blue discolored Soymilk and Tofu. *Journal of Genomics*, *9*, 55–67. https://doi.org/10.7150/jgen.65500

Garrido-Sanz, D., Arrebola, E., Martínez-Granero, F., García-Méndez, S., Muriel, C., Blanco-Romero, E., … Redondo-Nieto, M. (2017). Classification of isolates from the *Pseudomonas fluorescens* complex into phylogenomic groups based in group-specific markers. *Frontiers in Microbiology*, *8*, Article 413. https://doi.org/10.3389/fmicb.2017.00413

Garrido-Sanz, D., Meier-Kolthoff, J. P., Göker, M., Martín, M., Rivilla, R., & Redondo-Nieto, M. (2016). Genomic and genetic diversity within the *Pseudomonas fluorescens* complex. *PLoS ONE*, *11*(2). https://doi.org/10.1371/journal.pone.0150183

Gomila, M., Peña, A., Mulet, M., Lalucat, J., & García-Valdés, E. (2015). Phylogenomics and systematics in *Pseudomonas*. *Frontiers in Microbiology*, *6*(Article 214). https://doi.org/10.3389/fmicb.2015.00214

Hider, R. C., & Kong, X. (2010). Chemistry and biology of siderophores. *Natural Product Reports*, *27*(5), 637–657. https://doi.org/10.1039/b906679a

Jeukens, J., Freschi, L., Kukavica-Ibrulj, I., Emond-Rheault, J.-G., Tucker, N. P., & Levesque, R. C. (2019). Genomics of antibiotic-resistance prediction in *Pseudomonas aeruginosa*. *Annals of the New York Academy of Sciences*, *1435*, 5–17. https://doi.org/10.1111/nyas.13358

Lalucat, J., Mulet, M., Gomila, M., & García-Valdés, E. (2020). Genomics in Bacterial Taxonomy: Impact on the Genus *Pseudomonas*. *Genes*, *11*(2), 139. https://doi.org/10.3390/genes11020139

Lick, S., Kröckel, L., Wibberg, D., Winkler, A., Blom, J., Bantleon, A., … Kalinowski, J. (2020). *Pseudomonas carnis* sp. nov., isolated from meat. *International Journal of Systematic and Evolutionary Microbiology*. https://doi.org/10.1099/ijsem.0.003928

Lick, S., Wibberg, D., Winkler, A., Blom, J., Grimmler, C., Goesmann, A., … Kröckel, L. (2021). *Pseudomonas paracarnis* sp. nov., isolated from refrigerated beef. *International Journal of Systematic and Evolutionary Microbiology*, *71*, 004652. https://doi.org/10.1099/ijsem.0.004652

Maier, C., Huptas, C., von Neubeck, M., Scherer, S., Wenning, M., & Lücking, G. (2020). Genetic Organization of the *aprX-lipA2* Operon Affects the Proteolytic Potential of *Pseudomonas* Species in Milk. *Frontiers in Microbiology*, *11*(June), 1–13. https://doi.org/10.3389/fmicb.2020.01190

Martin, N. H., Murphy, S. C., Ralyea, R. D., Wiedmann, M., & Boor, K. J. (2011). When cheese gets the blues: *Pseudomonas fluorescens* as the causative agent of cheese spoilage. *Journal of Dairy Science*, *94*, 3176–3183. https://doi.org/10.3168/jds.2011-4312

Nogarol, C., Acutis, P. L., Bianchi, D. M., Maurella, C., Peletto, S., Gallina, S., … Decastelli, L. (2013). Molecular characterization of *Pseudomonas fluorescens* isolates involved in the Italian "blue mozzarella" event. *Journal of Food Protection*, *76*(3), 500–504. https://doi.org/10.4315/0362-028X.JFP-12-312

Palleroni, N. J. (2005). Genus I *Pseudomonas* Migula 1894. In D. J. Brenner, N. R. Krieg, & J. T. Staley (Eds.), *Bergey's Manual of Systematic Bacteriology, The Proteobacteria Part B The Gammaproteobacteria. Vol. 2* (2nd ed., pp. 323–379). East Lansing, Michigan:

Springer.

RASFF. (2010). The Rapid Alert System for Food and Feed (RASFF). *Annual Report*. Retrieved from https://ec.europa.eu/food/sites/food/files/safety/docs/rasff_annual_report_2010_en.pdf. Acessed 26 april 2020.

Reichler, S. J., Martin, N. H., Evanowski, R. L., Kovac, J., Wiedmann, M., & Orsi, R. H. (2019). A century of gray: A genomic locus found in 2 distinct *Pseudomonas* spp. is associated with historical and contemporary color defects in dairy products worldwide. *Journal of Dairy Science*, *102*, 5979–6000. https://doi.org/10.3168/jds.2018-16192

Rossi, C., Serio, A., Chaves-López, C., Anniballi, F., Auricchio, B., Goffredo, E., … Paparella, A. (2018). Biofilm formation, pigment production and motility in *Pseudomonas* spp. isolated from the dairy industry. *Food Control*, *86*, 241–248. https://doi.org/10.1016/j.foodcont.2017.11.018

Sechi, P., Vizzani, A., Scuota, S., Zicavo, A., Parmegiani, S., & Cenci Goga, B. (2011). Anomalous blue colouring of mozzarella cheese intentionally contaminated with pigment producing strains of *Pseudomonas fluorescens*. *Italian Journal of Food Safety*, *1*(1), 81–84. https://doi.org/10.4081/ijfs.2011.1.81

Silby, M. W., Winstanley, C., Godfrey, S. A. C., Levy, S. B., & Jackson, R. W. (2011). *Pseudomonas* genomes: diverse and adaptable. *FEMS Microbiology Reviews*, *35*(4), 652–680. https://doi.org/10.1111/j.1574-6976.2011.00269.x

Singh, P. K., Bourque, G., Craig, N. L., Dubnau, J. T., Feschotte, C., Flasch, D. A., … Levin, H. L. (2014). Mobile genetic elements and genome evolution 2014. *Mobile DNA*, *5*, 26. https://doi.org/10.1186/1759-8753-5-26

Stoeckel, M., Lidolt, M., Achberger, V., Glück, C., Krewinkel, M., Stressler, T., … Hinrichs,

J. (2016). Growth of *Pseudomonas weihenstephanensis*, *Pseudomonas proteolytica* and *Pseudomonas* sp. in raw milk: Impact of residual heat-stable enzyme activity on stability of UHT milk during shelf-life. *International Dairy Journal*, *59*, 20–28. https://doi.org/10.1016/j.idairyj.2016.02.045

Woods, R. G., Burger, M., Beven, C.-A., & Beacham, I. R. (2001). The *aprX–lipA* operon of *Pseudomonas fluorescens* B52: a molecular analysis of metalloprotease and lipase production. *Microbiology*, *147*(2), 345–354. https://doi.org/10.1099/00221287-147-2-345

Zhang, C., Bijl, E., & Hettinga, K. (2018). Destabilization of UHT milk by protease AprX from *Pseudomonas fluorescens* and plasmin. *Food Chemistry*, *263*(15), 127–134. https://doi.org/10.1016/j.foodchem.2018.04.128

Zhang, C., Bijl, E., Muis, K. E., & Hettinga, K. (2020). Stability of fat globules in UHT milk during proteolysis by the AprX protease from *Pseudomonas fluorescens* and by plasmin. *Journal of Dairy Science*, *103*, 179–190. https://doi.org/10.3168/jds.2019-17150

CHAPTER I: *Pseudomonas* sp. as the causative agent of anomalous blue discoloration in Brazilian fresh soft cheese (Minas Frescal)

Rafaela da Silva Rodrigues et al.

The content of this part has been published in:

International Dairy Journal
https://doi.org/10.1016/j.idairyj.2021.105020

Title page

Rafaela da Silva Rodrigues[1,2], Solimar Gonçalves Machado[2], Antônio Fernandes de Carvalho[2]*, Luís Augusto Nero[1]*

Pseudomonas sp. as the causative agent of anomalous blue discoloration in Brazilian fresh soft cheese (Minas Frescal)

[1] InsPOA - Laboratório de Inspeção de Produtos de Origem Animal, Departamento de Veterinária, Universidade Federal de Viçosa, Campus Universitário, Centro, Viçosa MG 36570 900, Brazil

[2] INOVALEITE - Laboratório de Pesquisa em Leites e Derivados, Departamento de Tecnologia de Alimentos, Universidade Federal de Viçosa, Campus Universitário, Centro, Viçosa MG 36570 900, Brazil

* Corresponding authors: LA Nero, nero@ufv.br; AF Carvalho, antoniofernandes@ufv.br

Abstract

In this study, we investigated the anomalous blue discoloration in fresh cheeses produced in Brazil (Minas Frescal). Isolates (n = 31) obtained from Minas Frescal cheeses with blue discoloration and identified as *Pseudomonas* spp. through PCR. These isolates were characterised as able to produce a diffusible blue pigment *in vitro* at different temperatures (5, 14 and 25 °C) after 24 h of incubation. Two isolates (A006 and B157) were selected and subjected to whole-genome sequencing; an *in silico* multi-locus sequence typing identified one new allelic profile for B157 and showed that isolates that produce the blue discoloration possibly share a common evolutionary development. Furthermore, algorithms ANIb, GGDC and inferred tree from GBDP distances, allowed the identification of B157 as *Pseudomonas carnis* and A006 as a potential novel species. For the first time, *Pseudomonas* strains that cause blue discoloration in Minas frescal cheese in Brazil were investigated and identified.

Keywords: fresh cheese; blue discoloration; *Pseudomonas*; spoilage

1. Introduction

Bacteria belonging to the genus *Pseudomonas* are common spoilage contaminants associated with dairy, especially in fluid milk and cheese, due to their ability to produce lipolytic and proteolytic enzymes (Arslan, Eyi, & Özdemir, 2011; Capodifoglio et al., 2016; Machado et al., 2015; Meng et al., 2017). Also, *Pseudomonas* strains are usually described as capable of producing pigments, such as pyocyanine, pyoverdin, pyorubin, pyochelin and pyomelanin, that can cause spoilage due to formation of discoloration and/or development of spots in dairy foods, especially cheeses (Andreani et al., 2014; Capodifoglio et al., 2016; del Olmo, Calzada, & Nuñez, 2018; Rossi et al., 2018; Zhang, Bijl, Muis, & Hettinga, 2020). A current concern in the dairy industry is a blue pigment produced by some *Pseudomonas* strains, and described as a blue discoloration (Martin, Murphy, Ralyea, Wiedmann, & Boor, 2011; Nogarol et al., 2013). This specific spoilage usually occurs in fresh cheeses with high moisture content, high pH (5.3–6.0) and produced without starter cultures, and the blue pigment is developed even during storage at low temperatures (Carminati et al., 2019; Carrascosa et al., 2015; Sechi et al., 2011).

Spoilage due to blue discoloration gained attention by the dairy industry after several reports in mozzarella, being named as “blue mozzarella” (Nogarol et al., 2013). As consequence, the Rapid Alert System for Food and Feed (RASFF) described this spoilage and identified *Pseudomonas fluorescens* as its causative agent (RASFF, 2010). Since then, several studies were developed to properly characterise this spoilage in cheeses, targeting the detection and identification of their causative agents. Therefore, blue discoloration was mainly reported in fresh cheeses, as mozzarella (Caputo et al., 2015; Carminati et al., 2019; Sechi et al., 2011), Latin-style fresh Queso Fresco cheese (Martin et al., 2011), and Gran Canaria artisanal fresh cheese (Carrascosa et al., 2015).

The Brazilian dairy industry is increasingly reporting the occurrence of blue discoloration in Minas frescal cheeses, a fresh and soft cheese that contains a high moisture content (>55%) and pH higher than 5.0 (Carvalho, Viotto, & Kuaye, 2007; Magenis et al., 2014). As a fresh cheese, Minas frescal cheese does not undergo maturation, and it is highly susceptible to microbial contamination during production, being highly perishable even during storage at low temperatures (Carvalho et al., 2007; Nascimento, Moreno, & Kuaye, 2008; Silva, Almeida, Alves, & Almeida, 2003). Despite the evidence of routes for microbial contamination and spoilage, blue discoloration in Brazilian cheeses (Fig. 1) is wrongly associated with chemical residues; to the best of our knowledge, there is no scientific report related to blue discoloration spoilage in Minas frescal cheese. Thus, we hypothesised that *Pseudomonas* strains are the causative agents of this specific spoilage described in Minas frescal cheese in Brazil, and here we described the genetic features of some of these strains.

Fig. 1. Blue discoloration in Brazilian fresh soft cheese (Minas Frescal).

2. Material and methods

2.1. Sampling, isolation of *Pseudomonas* and determination of the blue discoloration phenotype

Samples of Minas frescal cheese that presented blue discoloration (n = 2, named A and B) were subjected to microbial analysis to detect the causative agents of this spoilage. Both samples were produced by different Brazilian dairy industries and facilities, without starter cultures and acidified with lactic acid (personal information). Sub-samples of 10 g were diluted with 90 mL sodium citrate (20 g L^{-1}; pH 7.5), and ten-fold diluted with peptone salt solution (0.1% peptone, w/v, 0.85% NaCl, w/v) (ISO, 2010). Selected dilutions were plated in duplicate in plates containing (i) *Pseudomonas* agar base (PAB; Oxoid Ltd., Basingstoke, UK) supplemented with cephaloridine, fucidin and cetrimide (CFC, Oxoid), (ii) PAB-CFC supplemented with bromothymol blue (0.02 g L^{-1}) and (iii) PAB-CFC supplemented with bromothymol blue (0.02 g L^{-1}) and lactose (instead of glycerol, at 10 g L^{-1}). The plates were incubated at 25 ± 1 °C for 48 h, and the plates containing 30 to 300 colonies were selected for further analysis. Only bluish-green colonies were selected from plates containing the culture media supplemented with bromothymol blue (ii) and (iii) (Machado et al., 2015). After incubation, isolated colonies were selected (10% of counts), streaked onto plate count agar (PCA, Oxoid) and incubated at 25 °C for 24 h three consecutive times. After purification, the isolates obtained (n = 98 for sample A; n = 159 for sample B) were subjected to Gram staining and characterised for oxidase and catalase production (Downes & Ito, 2001). Isolates characterised as Gram-negative, oxidase and catalase positive were identified as presumable *Pseudomonas* spp., and the colonies were subjected to longwave ultraviolet light to check their ability in producing fluorescence pigments (Martin et al., 2011). Selected isolates were streaked

onto plates containing PCA and potato dextrose agar (PDA, Oxoid) and incubated at 5, 14, 25 and 32 °C for 168 h, being daily examined to verify the production of pigments, in special the blue discoloration (Andreani et al., 2014; Carminati et al., 2019).

Isolates that presented the blue discoloration were selected (n = 30 for sample A; n = 1 for sample B) and subjected to extraction of genomic DNA as described by Li and Mustapha (2002). DNA concentration was estimated by spectrophotometry (NanoDrop Lite, Thermo Scientific, Waltham, MA, USA) and subjected to a PCR protocol for identification, using primers PA-GS-F (5′-GACGGGTGAGTAATGCCTA-3′) and PA-GS-R (5′-CACTGGTGTTCCTTCCTATA-3′) to detect a genus-specific region of 16S rRNA for *Pseudomonas* spp. (Spilker, Coenye, Vandamme, & LiPuma, 2004).

2.2. Genome sequencing, assembly and gene prediction

One isolate from each cheese sample was subjected to whole-genome sequencing (WGS). The DNA preparation and sequencing were performed by Neoprospecta Microbiome Technologies (Florianópolis, SC, Brazil). The whole-genome shotgun sequencing was performed by Illumina Miseq platform (paired-ends reads; forward of 305 bp and reverse of 205 bp). A preliminary evaluation of the quality of the raw reads was performed in FastQC version 0.11.5 (Andrews, 2010). Then, the reads were trimmed with Trimmomatic v.0.36 to remove the adaptors and to accept 20 as lowest Phred value (Bolger, Lohse, & Usadel, 2014). After trimming, sequence reads were rechecked for quality. *De novo* assembly was performed using five assemblers on the same data, for comparison: Mira Assembler v. 4.9.6, mode "genome, accurate" (Chevreux, Wetter, & Suhai, 1999); SPAdes v. 3.12.0, using the "-careful" parameter (Bankevich et al., 2012); CLC Genomics Workbench v. 20.0.4; SOAPdenovo v. 2.04 with k-mer 127 (Luo et al., 2012); Velvet v. 1.2.10 with k-mer 71 (Zerbino & Birney, 2008).

QUAST v. 5.0.2 was used to check the quality parameters of the assembled genome (Gurevich, Saveliev, Vyahhi, & Tesler, 2013). Contigs shorter than 500 nt were removed from the assembly set, and ContEst16S software was used to check possible contamination (Lee et al., 2017). Finally, the genes of the draft assemblies were predicted using Prokaryotic Genome Annotation System (Prokka) v. 1.14.5 (Seemann, 2014).

2.3. Phylogenetic analysis

To reveal the genetic relationships and to identify a possible connection of the isolates identified here with strains characterised in previous studies as capable of producing blue discoloration, an in silico multi-locus sequence typing (MLST) was performed at pubMLST (https://pubmlst.org/organisms/pseudomonas-fluorescens), and a circular phylogenetic tree-based Neighbor-Joining was generated using iTol v.4.2 (https://itol.embl.de/; Letunic & Bork, 2016). Furthermore, the genomes of the isolates A006 and B157 were related to the genomes of the all-species type strains in the *P. fluorescens* phylogenetic subgroup available in database platforms (Duman et al., 2020; Gomila, Peña, Mulet, Lalucat, & García-Valdés, 2015). Relatedness was estimated, using several approaches: average nucleotide identity (ANI) based on BLASTN algorithm (ANIb); ANI based on the MUMMER ultra-rapid aligning tool (ANIm); tetranucleotide frequency correlation coefficients (TETRA); and, digital DNA:DNA hybridisation (DDH) by the genome-to-genome distance (GGDC) method. ANIb, ANIm and TETRA frequencies were calculated using the JSpeciesWS software tool available on a webpage (http://jspecies.ribohost.com/jspeciesws/; Richter, Rosselló-Móra, Oliver Glöckner, & Peplies, 2016). GGDC was also calculated using a web service (http://ggdc.dsmz.de/; Meier-Kolthoff, Auch, Klenk, & Göker, 2013) and the recommended BLAST method. The GGDC results were based on the recommended formula 2, which is independent of genome length and,

thus, robust when using draft genome sequences. Additionally, a phylogenomic tree was built using the Type Strain Genome Server (TYGS; https://tygs.dsmz.de), for a whole genome-based taxonomic analysis (Meier-Kolthoff & Göker, 2019): the genomes obtained were compared with deposited genomes from the same type strain and a phylogenetic tree was visualised using iTol v.4.2.

3. Results and discussion

PAB-CFC supplemented with bromothymol blue and/or lactose allowed better isolation of colonies with *Pseudomonas* presumptive characteristics when compared with PAB-CFC alone. The production of diffusible blue pigment was already observed during purification procedure of the colonies. All colonies characterised as capable of producing pigment also presented fluorescence under ultraviolet light and PCR products with the expected amplicon size for *Pseudomonas* spp. Isolates were able to produce blue pigment after 24 h of incubation at 5, 14 and 25 °C, and only isolates from sample A were able to produce the blue pigment at 32 °C, after 4 d. Some *Pseudomonas* strains were already described as capable of producing a light blue pigment in PDA at 6, 22 and 31 °C (Andreani et al., 2014; Martin et al., 2011). However, the strains that were attributed to “blue mozzarella” in Europe presented a distinct blue phenotype, producing a diffusible blue pigment at 6 and 22 °C, but not above 30 °C (Andreani et al., 2014; Rossi et al., 2018). Also, we observed that isolates clearly showed visible development of blue discoloration on PDA (Fig. 2).

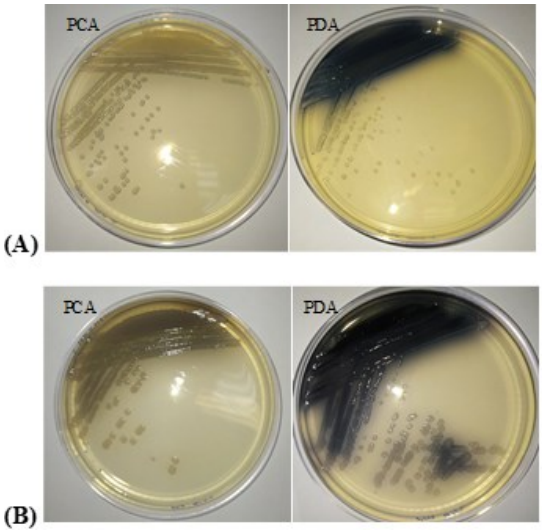


Fig. 2. Blue phenotype and growth of isolates A006 (A) and B157 (B) after incubation at 25 °C for 48 h. PCA, plate count agar; PDA, potato dextrose agar.

This behaviour of blue discoloration at low temperatures has also been observed when the strain was inoculated in mozzarella or only in the brine (Rossi et al., 2018). Chierici et al. (2016) studied the blue discoloration by *Pseudomonas* spp. strains in brine (pH 6.3) incubated at different temperatures, and they observed the presence of the pigment only at 4 °C, after 10 d, and 14 °C, after 3 d; strains did not produce blue discoloration when incubated at 30 °C. Within our results, these data indicate that the temperature affects directly the blue discoloration spoilage, which may occur due to a potential stressing condition induced by low temperatures, as proposed by Caputo et al. (2015).

Based on these results, isolates A006 (from sample A) and B157 (from sample B) were selected as representative from the obtained samples and then subjected to WGS, as described above. The best assembly was obtained with the assembler Mira, based on the observed quality parameters (Table 1). Whole genome assemblies were deposited in GenBank (National Center for Biotechnology Information, Bethesda MD, USA, BioProject PRJNA682387). The prediction with Prokka for A006 allowed the identification of 5951 protein-coding sequences

(CDS), 72 tRNA, 16 rRNA and 1 tmRNA; the same analysis for B157 allowed to predict of 5958 CDS, 66 tRNA, 10 rRNA and 1 tmRNA.

Table 1. Quality metrics of the *Pseudomonas* assembled genomes.

Isolates	Quality parameters	Platform				
		Mira	SPAdes	CLC	SOAPdenovo	Velvet
A006	Contigs	99	144	125	512	590
	Total length (bp)	6,493,137	6,440,882	6,426,093	6,542,594	6,539,935
	Max. contig length (bp)	566,016	445,448	566,695	88,801	142,180
	GC (%)	59.55	59.58	59.58	59.59	59.59
	N50	255,869	128,269	121,359	25,810	45,451
B157	Contigs	68	68	98	303	182
	Total length (bp)	6,567,469	6,525,355	6,509,608	6,528,503	6,506,313
	Max. contig length (bp)	1,217,346	466,006	417,142	141,554	213,190
	GC (%)	59.63	59.65	59.65	59.65	59.65
	N50	385,782	209,876	124,195	43,923	71,152

The sequences obtained were compared with related *P. fluorescens* sequences to search for the corresponding allelic profiles for MLST: various sequences used in this comparison were from strains capable of producing blue discoloration. B157 presented a new allelic profile, and A006 was characterised as belonging to sequence type 96 (ST96). Based on these sequences, all isolates that are capable of producing blue discoloration possibly share a common evolutionary development (Fig. 3), as hypothesised by Andreani et al. (2014).

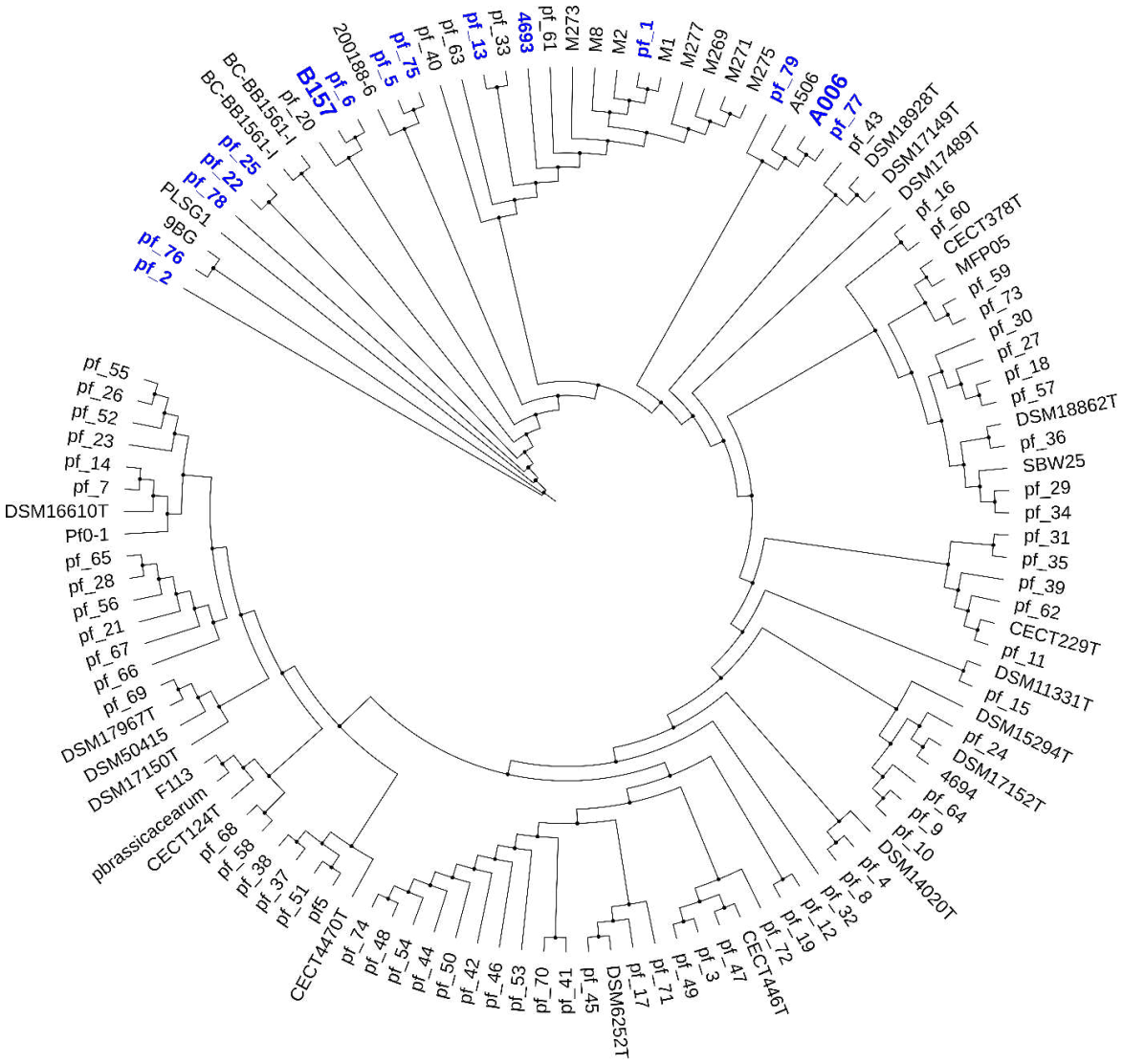


Fig. 3. Neighbour-joining tree obtained from the comparison of sequences by MLST (https://pubmlst.org/organisms/pseudomonas-fluorescens), using iTol (https://itol.embl.de/). Isolated in blue are producers of the blue pigment.

The genome relatedness values of isolate A006 with its closest relatives in the *P. fluorescens* group, based on ANIb, ANIm and GGDC, were usually below the established cutoff for each algorithm (ANIb: 95%, ANIm: 95%, TETRA: 0.999 and GGDC: 70%). The similarity to the closest related strain, *Pseudomonas carnis*, was 94.84% for ANIb, 95.79% for ANIm, 0.99900 for TETRA and 62.90% for GGDC (Supplementary material Table S1). Taxonomic classification requests all values of the algorithms to be above the established cutoff, indicating that this isolate, A006, possibly belongs to a new species. For B157, *P. carnis* was the closest

related species, with similarities of 98.15% by ANIb, 98.53% by ANIm, 0.99965 by TETRA and 86.60% by GGDC. These high similarity values indicate that isolate B157, therefore, belongs to this species, *P. carnis*. Based on the genome comparison in TYGS (Fig. 4), isolate B157 can be included in the group from *P. carnis* and isolate A006 was confirmed as a potential novel species, since it was not included in any group formed by known species.

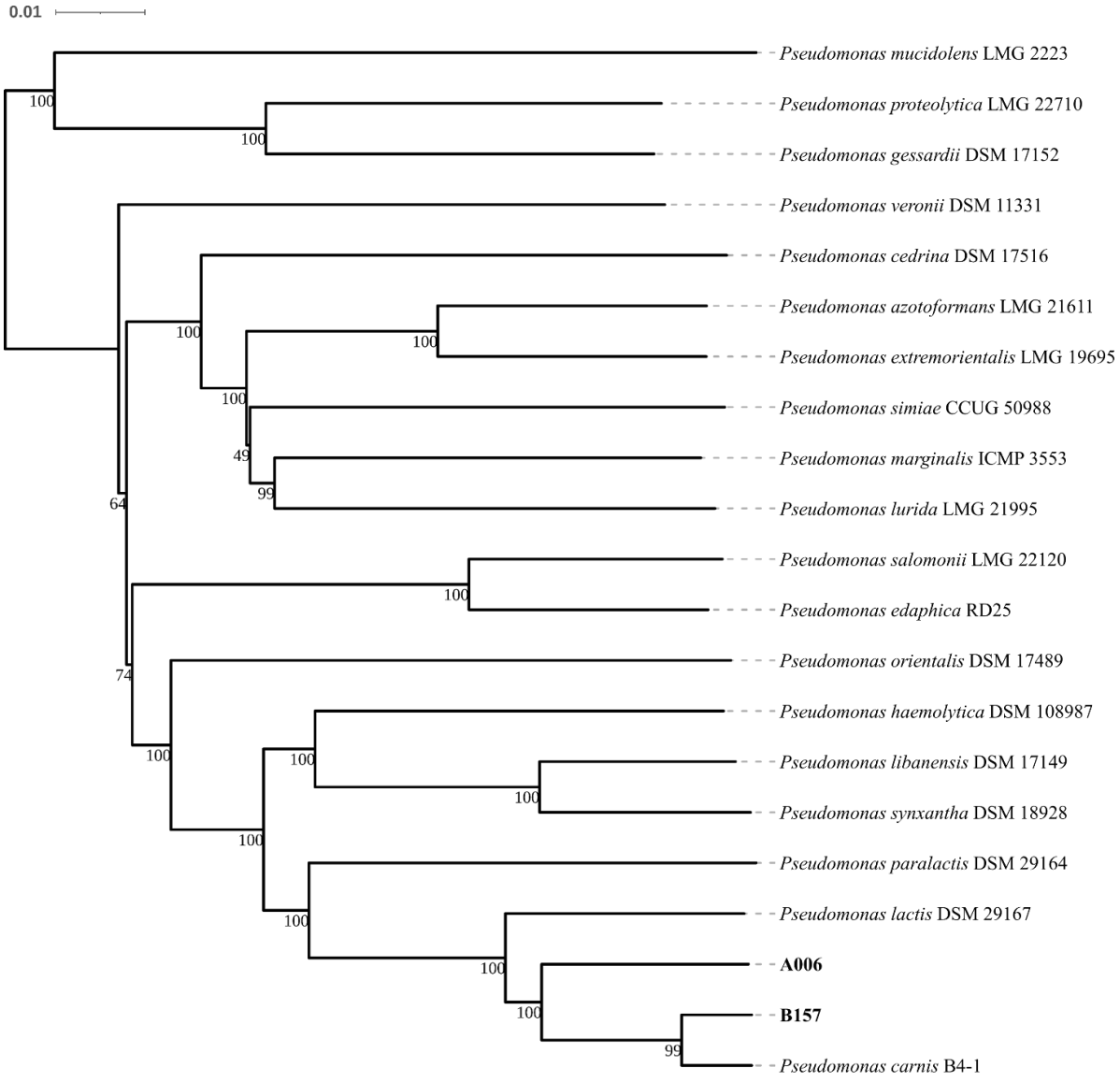


Fig. 4. A phylogenomic tree constructed using the Type Strain Genome Server (TYGS; https://tygs.dsmz.de), for a whole genome-based taxonomic analysis and determination of closely related type strains (Meier-Kolthoff & Göker, 2019). Tree inferred with FastME 2.1.6.1 (Lefort, Desper, & Gascuel, 2015) from Genome Blast Distance

Phylogeny (GBDP) distances calculated from genome sequences. The branch lengths are scaled in terms of GBDP distance formula d_5. The numbers above branches are GBDP pseudo-bootstrap support values > 60% from 100 replications, with an average branch support of 93.6%. The tree was rooted at the midpoint (Farris, 1972).

To verify the similarity of the isolates with *P. carnis*, all genomes of this species available in GenBank were downloaded (n = 5), and compared with the genomes from A006 and B157 based on algorithms ANIb, ANIm, TETRA and GGDC. Isolate B157 presented similar profiles with all GenBank sequences, allowing its identification as *P. carnis*, and isolate A006 presented a same pattern of similarity with other *Pseudomonas* species, supporting the hypothesis of belonging to a novel species (Supplementary material Table S2). Based on this genomic comparison, isolate A006 was grouped in a different cluster from *P. carnis* strains and isolate B157 (both included in a single cluster) (Fig. 5).

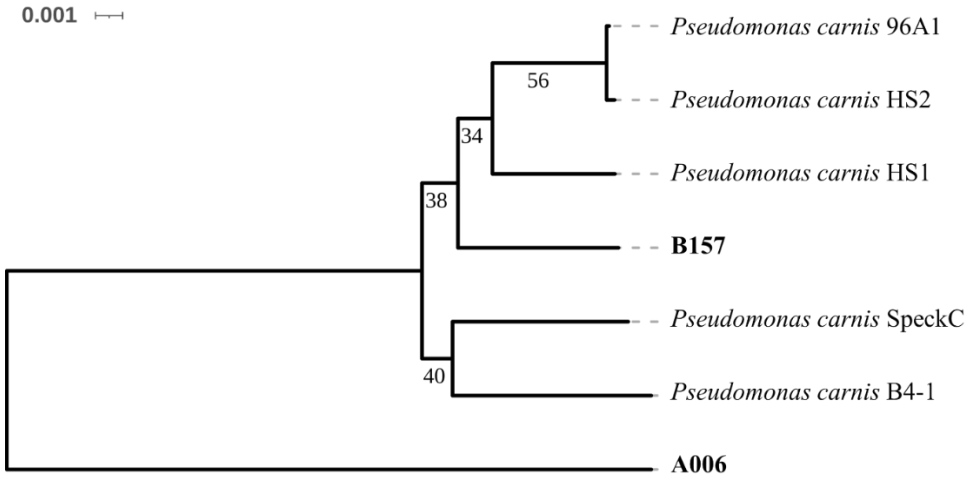


Fig. 5. A phylogenomic tree constructed using the Type Strain Genome Server (TYGS; https://tygs.dsmz.de) for comparison of genome sequences (Meier-Kolthoff et al., 2013). Tree inferred with FastME 2.1.6.1 (Lefort et al., 2015) from Genome Blast Distance Phylogeny (GBDP) distances calculated from genome sequences. The branch lengths are scaled in terms of GBDP distance formula d_5. The numbers above branches are GBDP pseudo-bootstrap support values > 60% from 100 replications, with an average branch support of 42.0%. The tree was rooted at the midpoint (Farris, 1972).

This is the first study that isolates *Pseudomonas* causing anomalous blue discoloration in Brazilian fresh soft cheese and reports similarity of the isolates from cheese in Brazil to those related to the blue mozzarella event in 2010 in Europe. Moreover, ANIb and GGDC and the inferred tree from GBDP distances (by TYGS), indicated that one isolate, B157, belongs to the *P. carnis* species, and the other isolate, A006, possibly belongs to a novel species.

References

Andreani, N. A., Martino, M. E., Fasolato, L., Carraro, L., Montemurro, F., Mioni, R., et al. (2014). Tracking the blue: A MLST approach to characterise the *Pseudomonas fluorescens* group. *Food Microbiology*, *39*, 116–126. https://doi.org/10.1016/j.fm.2014.11.009

Andrews, S. (2010). FastQC - A quality control tool for high throughput sequence data. http://www.bioinformatics.babraham.ac.uk/projects/fastqc/. *Babraham Bioinformatics*.

Arslan, S., Eyi, A., & Özdemir, F. (2011). Spoilage potentials and antimicrobial resistance of *Pseudomonas* spp. isolated from cheeses. *Journal of Dairy Science*, *94*, 5851–5856. https://doi.org/10.3168/jds.2011-4676

Bankevich, A., Nurk, S., Antipov, D., Gurevich, A. A., Dvorkin, M., Kulikov, A. S., et al. (2012). SPAdes: A new genome assembly algorithm and its applications to single-cell sequencing. *Journal of Computational Biology*, *19*(5), 455–477. https://doi.org/10.1089/cmb.2012.0021

Bolger, A. M., Lohse, M., & Usadel, B. (2014). Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics*, *30*(15), 2114–2120. https://doi.org/10.1093/bioinformatics/btu170

Capodifoglio, E., Vidal, A. M. C., Lima, J. A. S., Bortoletto, F., D'Abreu, L. F., Gonçalves, A. C. S., et al. (2016). Lipolytic and proteolytic activity of *Pseudomonas* spp. isolated during milking and storage of refrigerated raw milk. *Journal of Dairy Science*, *99*, 5214–5223. https://doi.org/10.3168/jds.2015-10453

Caputo, L., Quintieri, L., Bianchi, D. M., Decastelli, L., Monaci, L., Visconti, A., & Baruzzi, F. (2015). Pepsin-digested bovine lactoferrin prevents Mozzarella cheese blue discoloration caused by *Pseudomonas fluorescens*. *Food Microbiology*, *46*, 15–24. https://doi.org/10.1016/j.fm.2014.06.021

Carminati, D., Bonvini, B., Rossetti, L., Zago, M., Tidona, F., & Giraffa, G. (2019). Investigation on the presence of blue pigment-producing *Pseudomonas* strains along a production line of fresh mozzarella cheese. *Food Control*, *100*, 321–328. https://doi.org/10.1016/j.foodcont.2019.02.009

Carrascosa, C., Millán, R., Jaber, J. R., Lupiola, P., del Rosario-Quintana, C., Mauricio, C., & Sanjuán, E. (2015). Blue pigment in fresh cheese produced by *Pseudomonas fluorescens*. *Food Control*, *54*, 95–102. https://doi.org/10.1016/j.foodcont.2014.12.039

Carvalho, J. D. G., Viotto, W. H., & Kuaye, A. Y. (2007). The quality of Minas Frescal cheese produced by different technological processes. *Food Control*, *18*, 262–267. https://doi.org/10.1016/j.foodcont.2005.10.005

Chevreux, B., Wetter, T., & Suhai, S. (1999). Genome Sequence Assembly Using Trace Signals and Additional Sequence Information. *Computer Science and Biology: Proceedings of the German Conference on Bioinformatics (GCB) '99, GCB, Hannover, Germany.*, *99*, 45–56.

Chierici, M., Picozzi, C., La Spina, M. G., Orsi, C., Vigentini, I., Zambrini, V., & Foschino, R. (2016). Strain diversity of *Pseudomonas fluorescens* group with potential blue pigment phenotype isolated from dairy products. *Journal of Food Protection*, *79*(8), 1430–1435.

https://doi.org/10.4315/0362-028X.JFP-15-589

del Olmo, A., Calzada, J., & Nuñez, M. (2018). The blue discoloration of fresh cheeses: A worldwide defect associated to specific contamination by *Pseudomonas fluorescens*. *Food Control*, *86*, 359–366. https://doi.org/10.1016/j.foodcont.2017.12.001

Downes, F. P., & Ito, K. (2001). Compendium of Methods for The Microbiological Examination of Foods. In *American Public Health Association*. Washington, DC.

Duman, M., Mulet, M., Saticioglu, I. B., Altun, S., Gomila, M., Lalucat, J., & García-Valdés, E. (2020). *Pseudomonas sivasensis* sp. nov. isolated from farm fisheries in Turkey. *Systematic and Applied Microbiology*, *43*, 126103. https://doi.org/10.1016/j.syapm.2020.126103

Farris, J. S. (1972). Estimating Phylogenetic Trees from Distance Matrices. *The American Naturalist*, *106*(951), 645–668. https://doi.org/10.1086/282802

Gomila, M., Peña, A., Mulet, M., Lalucat, J., & García-Valdés, E. (2015). Phylogenomics and systematics in *Pseudomonas*. *Frontiers in Microbiology*, *6*(Article 214). https://doi.org/10.3389/fmicb.2015.00214

Gurevich, A., Saveliev, V., Vyahhi, N., & Tesler, G. (2013). QUAST: Quality assessment tool for genome assemblies. *Bioinformatics*. https://doi.org/10.1093/bioinformatics/btt086

ISO. (2010). ISO 6887-5. 2010. Microbiology of Food and Animal Feeding Stuffs–Preparation of Test Samples, Initial Suspension and Decimal Dilutions for Microbiological Examination–Part 5: Specific Rules for the Preparation of Milk and Milk Products. *International Standard*.

Lee, I., Chalita, M., Ha, S. M., Na, S. I., Yoon, S. H., & Chun, J. (2017). ContEst16S: An algorithm that identifies contaminated prokaryotic genomes using 16S RNA gene sequences.

International Journal of Systematic and Evolutionary Microbiology, *67*, 2053–2057. https://doi.org/10.1099/ijsem.0.001872

Lefort, V., Desper, R., & Gascuel, O. (2015). FastME 2.0: A comprehensive, accurate, and fast distance-based phylogeny inference program. *Molecular Biology and Evolution*, *32*(10), 2798–2800. https://doi.org/10.1093/molbev/msv150

Letunic, I., & Bork, P. (2016). Interactive tree of life (iTOL) v3: an online tool for the display and annotation of phylogenetic and other trees. *Nucleic Acids Research*, *44*, W242–W245. https://doi.org/10.1093/nar/gkw290

Li, Y., & Mustapha, A. (2002). Evaluation of four template preparation methods for polymerase chain reaction-based detection of *Salmonella* in ground beef and chicken. *Letters in Applied Microbiology*, *35*, 508–512. https://doi.org/10.1046/j.1472-765X.2002.01231.x

Luo, R., Liu, B., Xie, Y., Li, Z., Huang, W., Yuan, J., et al. (2012). SOAPdenovo2: An empirically improved memory-efficient short-read de novo assembler. *GigaScience*, *1*(18). https://doi.org/10.1186/2047-217X-1-18

Machado, S. G., da Silva, F. L., Bazzolli, D. M. S., Heyndrickx, M., Costa, P. M. d. A., & Vanetti, M. C. D. (2015). *Pseudomonas* spp. and *Serratia liquefaciens* as Predominant Spoilers in Cold Raw Milk. *Journal of Food Science*, *80*(8), M1842–M1849. https://doi.org/10.1111/1750-3841.12957

Magenis, R. B., Prudêncio, E. S., Fritzen-Freire, C. B., Stephan, M. P., Silvio do Egito, A. Ô., & Daguer, H. (2014). Rheological, physicochemical and authenticity assessment of Minas Frescal cheese. *Food Control*, *45*, 22–28. https://doi.org/10.1016/j.foodcont.2014.04.012

Martin, N. H., Murphy, S. C., Ralyea, R. D., Wiedmann, M., & Boor, K. J. (2011). When cheese gets the blues: *Pseudomonas fluorescens* as the causative agent of cheese spoilage. *Journal of Dairy Science*, *94*, 3176–3183. https://doi.org/10.3168/jds.2011-4312

Meier-Kolthoff, J. P., Auch, A. F., Klenk, H. P., & Göker, M. (2013). Genome sequence-based species delimitation with confidence intervals and improved distance functions. *BMC Bioinformatics*, *14*(60). https://doi.org/10.1186/1471-2105-14-60

Meier-Kolthoff, J. P., & Göker, M. (2019). TYGS is an automated high-throughput platform for state-of-the-art genome-based taxonomy. *Nature Communications*, *10*. https://doi.org/10.1038/s41467-019-10210-3

Meng, L., Zhang, Y., Liu, H., Zhao, S., Wang, J., & Zheng, N. (2017). Characterization of *Pseudomonas* spp. and associated proteolytic properties in raw milk stored at low temperatures. *Frontiers in Microbiology*, *8*, article 2158. https://doi.org/10.3389/fmicb.2017.02158

Nascimento, M. S., Moreno, I., & Kuaye, A. Y. (2008). Applicability of bacteriocin-producing *Lactobacillus plantarum*, *Enterococcus faecium* and *Lactococcus lactis* ssp. *lactis* as adjunct starter in Minas Frescal cheesemaking. *International Journal of Dairy Technology*, *61*(4), 352–357. https://doi.org/10.1111/j.1471-0307.2008.00426.x

Nogarol, C., Acutis, P. L., Bianchi, D. M., Maurella, C., Peletto, S., Gallina, S., et al. (2013). Molecular characterization of *Pseudomonas fluorescens* isolates involved in the Italian "blue mozzarella" event. *Journal of Food Protection*, *76*(3), 500–504. https://doi.org/10.4315/0362-028X.JFP-12-312

RASFF. (2010). The Rapid Alert System for Food and Feed (RASFF). *Annual Report*. Retrieved from https://ec.europa.eu/food/sites/food/files/safety/docs/rasff_annual_report_2010_en.pdf. Acessed 26 april 2020.

Richter, M., Rosselló-Móra, R., Oliver Glöckner, F., & Peplies, J. (2016). JSpeciesWS: A web server for prokaryotic species circumscription based on pairwise genome comparison.

Bioinformatics, *32*(6), 929–931. https://doi.org/10.1093/bioinformatics/btv681

Rossi, C., Serio, A., Chaves-López, C., Anniballi, F., Auricchio, B., Goffredo, E., et al. (2018). Biofilm formation, pigment production and motility in *Pseudomonas* spp. isolated from the dairy industry. *Food Control*, *86*, 241–248. https://doi.org/10.1016/j.foodcont.2017.11.018

Sechi, P., Vizzani, A., Scuota, S., Zicavo, A., Parmegiani, S., & Cenci Goga, B. (2011). Anomalous blue colouring of mozzarella cheese intentionally contaminated with pigment producing strains of *Pseudomonas fluorescens*. *Italian Journal of Food Safety*, *1*(1), 81–84. https://doi.org/10.4081/ijfs.2011.1.81

Seemann, T. (2014). Prokka: Rapid prokaryotic genome annotation. *Bioinformatics*, *30*(14), 2068–2069. https://doi.org/10.1093/bioinformatics/btu153

Silva, I. M. M., Almeida, R. C. C., Alves, M. A. O., & Almeida, P. F. (2003). Occurrence of Listeria spp. in critical control points and the environment of Minas Frescal cheese processing. *International Journal of Food Microbiology*, (81), 241–248. https://doi.org/10.1016/S0168-1605(02)00223-4

Spilker, T., Coenye, T., Vandamme, P., & LiPuma, J. J. (2004). PCR-Based Assay for Differentiation of *Pseudomonas aeruginosa* from Other *Pseudomonas* Species Recovered from Cystic Fibrosis Patients. *Journal of Clinical Microbiology*, *45*(5), 2074–2079. https://doi.org/10.1128/JCM.42.5.2074-2079.2004

Zerbino, D. R., & Birney, E. (2008). Velvet: Algorithms for de novo short read assembly using de Bruijn graphs. *Genome Research*, *18*, 821–829. https://doi.org/10.1101/gr.074492.107

Zhang, C., Bijl, E., Muis, K. E., & Hettinga, K. (2020). Stability of fat globules in UHT milk during proteolysis by the AprX protease from *Pseudomonas fluorescens* and by plasmin. *Journal of Dairy Science*, *103*, 179–190. https://doi.org/10.3168/jds.2019-17150

Supplementary material

Table S1

Relatedness of the sequenced genomes of the *Pseudomonas* isolates A006 and B157 to whole genome sequences of all *P. fluorescens* subgroup species type strains based on ANIb, ANIm, TETRA and GGDC.

Strains[a]	A006				B157			
	ANIb	ANIm	TETRA	GGDC	ANIb	ANIm	TETRA	GGDC
A006	100.00	100.00	1.00000	100.00	94.76	**95.83**	**0.99922**	63.50
B157	94.91	**95.82**	**0.99922**	63.50	100.00	100.00	1.00000	100.00
P. antarctica DSM 15318^T (UYXQ00000000.1)	85.82	87.94	0.98367	32.40	85.83	87.93	0.98212	32.40
P. azotoformans LMG 21611^T (LT629702.1)	86.23	88.33	0.99185	32.90	86.15	88.28	0.99212	32.80
P. brenneri DSM 15294^T (VFIL00000000.1)	83.84	87.06	0.98293	29.20	83.68	87.04	0.98493	29.30
P. carnis B4-1^T (CABIVL000000000.1)	94.84	**95.79**	**0.99900**	62.90	**98.15**	**98.53**	**0.99965**	**86.60**
P. cedrina DSM 17516^T (UYXV00000000.1)	85.41	88.03	0.99336	32.10	85.60	88.04	0.99383	32.20
P. congelans DSM 14939^T (FNJH00000000.1)	76.18	84.51	0.91097	22.50	76.19	84.51	0.91401	22.50
P. costantinii LMG 22119^T (MDDR00000000.1)	86.09	87.89	0.98577	32.80	86.01	87.89	0.98471	32.70
P. edaphica RD25^T (VBVZ00000000.1)	86.13	88.27	0.98884	33.10	86.14	88.25	0.98792	33.00
P. extremaustralis DSM 17835^T (LT629689.1)	85.44	88.13	0.98987	32.70	85.50	88.21	0.99011	32.60
P. extremorientalis LMG 19695^T (MDGK00000000.1)	86.10	88.27	0.99207	32.90	85.95	88.27	0.99224	32.80
P. fildesensis KG01^T (LFMW00000000.1)	86.15	88.34	0.99101	33.10	86.04	88.32	0.99053	33.00
P. fluorescens ATCC 13525^T (LT907842.1)	85.76	87.95	0.98717	32.20	85.81	87.97	0.98663	32.10
P. gessardii DSM 17152^T (VFEW00000000.1)	83.54	87.09	0.98101	29.30	83.80	87.14	0.98203	29.50
P. grimontii DSM 17515T^T (VFES00000000.1)	86.15	88.42	0.99129	33.40	86.17	88.42	0.99054	33.20
P. haemolytica DSM 108987^T (VOIW00000000.1)	88.95	90.30	0.99647	39.00	88.76	90.28	0.99663	38.90
P. kairouanesis KC12^T (QUZU00000000.1)	85.37	87.98	0.99242	32.00	85.29	88.02	0.99190	32.10
P. lactis DSM 29167^T (JYLO00000000.1)	94.46	**95.27**	0.99896	60.00	93.95	94.97	**0.99915**	58.00
P. libanensis DSM 17149^T (JYLH00000000.1)	89.19	90.42	0.99547	39.70	88.92	90.41	0.99623	39.70
P. lurida LMG 21995^T (PDJB01000001.1)	85.81	88.03	0.99151	32.40	85.97	88.04	0.99163	32.50
P. marginalis ICPM 3553^T (LKEG00000000.1)	86.28	88.24	0.98993	32.90	86.24	88.26	0.98926	32.80
P. mediterranea DSM 16733^T (LT629790.1)	79.43	85.57	0.96957	24.70	79.35	85.58	0.97186	24.90
P. mucidolens LGM 2223^T (LT629802.1)	82.51	86.30	0.97448	27.80	82.40	86.28	0.97578	27.90
P. nabeulensis E10B^T (QUZT00000000.1)	85.35	88.03	0.98916	31.80	85.38	88.03	0.98801	32.00
P. orientalis DSM 17489^T (JYLM00000000.1)	87.08	88.92	0.99107	34.50	86.94	88.89	0.99198	34.20
P. palleroniana LMG 23076^T (PYWX00000000.1)	85.65	87.83	0.99468	31.80	85.66	87.82	0.99549	31.80
P. panacis DSM 18529^T (VFER00000000.1)	86.45	88.50	0.98814	33.50	86.41	88.51	0.98793	33.40
P. paralactis DSM 29164^T (JYLN00000000.1)	89.95	91.12	0.99576	41.60	89.77	91.19	0.99662	41.60
P. poae LMG 21465^T (MDFK00000000.1)	85.22	87.82	0.98900	32.00	85.24	87.82	0.98849	32.10
P. proteolytica CCUG 51515^T (VXJX00000000.1)	83.49	87.12	0.98597	29.20	83.61	87.15	0.98712	29.30
P. rhodesiae DSM 14020^T (VFEU00000000.1)	84.83	87.59	0.98342	31.00	84.96	87.64	0.98224	31.20
P. salomonii LMG 22120^T (MDFI00000000.1)	86.34	88.23	0.98779	33.00	86.20	88.17	0.98686	32.70
P. simiae CCUG 50988^T (MDFH00000000.1)	86.02	88.00	0.99219	32.30	85.97	87.98	0.99242	32.30
P. sivasensis P7^T (JAAOWU000000000.1)	85.62	87.95	0.98618	32.20	85.46	87.92	0.98498	32.00
P. synxantha DSM 18928^T (JYLJ00000000.1)	88.78	90.15	0.99471	38.50	88.43	90.12	0.99520	37.90
P. tolaasii NCPPB 2192^T (PHHD00000000.1)	85.52	87.93	0.98007	32.30	85.57	87.91	0.98006	32.00
P. trivialis DSM 14937^T (JYLK00000000.1)	85.35	87.93	0.98508	32.30	85.32	87.88	0.98478	32.10
P. veronii DSM 11331^T (JYLL00000000.1)	86.33	88.34	0.98630	33.30	86.22	88.38	0.98584	33.40

[a] The accession numbers (GenBank) of the whole genome sequences used in the study are indicated in brackets.

Table S2

Relatedness of the sequenced genomes of the *Pseudomonas* isolates A006 and B157 to whole genome sequences of all *P. carnis* strains based on ANIb, ANIm, TETRA and GGDC.

Strains[a]	A006				B157			
	ANIb	ANIm	TETRA	GGDC	ANIb	ANIm	TETRA	GGDC
A006	100.00	100.00	1.00000	100.00	94.76	**95.83**	**0.99922**	63.50
B157	94.91	**95.82**	**0.99922**	63.50	100.00	100.00	1.00000	100.00
P. carnis 96A1 (JAAMRA010000001.1)	94.70	**95.76**	**0.99904**	62.60	**98.52**	**98.98**	**0.99974**	**86.80**
P. carnis B4-1^T (CABIVL010000001.1)	94.84	**95.79**	**0.99900**	65.20	**98.15**	**98.53**	**0.99965**	**87.30**
P. carnis HS1 (JAAMRH010000001.1)	94.88	**95.81**	**0.99919**	60.30	**98.62**	**98.93**	**0.99974**	**86.60**
P. carnis HS2 (JAAMRI010000001.1)	94.59	**95.75**	**0.99905**	64.30	**98.47**	**98.98**	**0.99972**	**88.50**
P. carnis SpeckC (UYJC01000001.1)	94.95	**95.77**	0.99888	68.40	**98.23**	**98.59**	**0.99963**	**86.00**

[a] The accession numbers (GenBank) of the whole genome sequences used in the study are indicated in brackets.

CHAPTER II: Comparative genomic and functional annotation of *Pseudomonas* spp. genomes responsible for blue discoloration of Brazilian fresh soft cheese

Rafaela da Silva Rodrigues et al.

The content of this part has been published in:

International Dairy Journal
https://doi.org/10.1016/j.idairyj.2023.105605

Title page

Rafaela da Silva Rodrigues[a], Solimar Gonçalves Machado[a], Antônio Fernandes de Carvalho[a], Luís Augusto Nero[b]*

Comparative genomic and functional annotation of *Pseudomonas* spp. genomes responsible for blue discoloration of Brazilian fresh soft cheese

[a] Department of Food Technology, Federal University of Viçosa, Av. P. H. Rolfs s/n, 36570900 Viçosa, MG, Brazil

[b] Department of Veterinary Medicine, Federal University of Viçosa, Av. P. H. Rolfs s/n, 36570900 Viçosa, MG, Brazil

* Corresponding author: LA Nero, nero@ufv.br, tel. +55 31 3612 5615.

Abstract

In this study, we investigated by genomic analysis the strains related to blue discoloration in Brazilian fresh soft cheese (Minas Frescal) and compared their genomes with those of strains from other countries associated with spoilage. *Pseudomonas carnis* B157 and *Pseudomonas* sp. A006 were studied; A006 was identified as *Pseudomonas paracarnis*. Orthogroup analysis was performed between predicted proteins from the genomes of 11 pigment-producing and 4 non-producing strains. We found 29 unique genes contained in a region of 74 genes, including tryptophan biosynthesis (*trpABCDF*) and mobile genetic element genes, such as those involved in horizontal gene transfer (HGT; e.g., *pilM*, *pilV*, *pilX*, *tadA*, *traD*). Thus, we show a large genomic region of discoloration-causing strains containing several HGT genes. Comprehension at the genomic level will help advance the detection of these spoilages, such as in the design of markers that will allow their tracking and control in the dairy industry.

Keywords: Minas Frescal, food spoilers, blue pigment, genome annotation

1. Introduction

Cheese spoilage caused by *Pseudomonas* spp. has increased in recent years, mainly due to the growing number of reports of a blue discoloration on the surface of fresh cheeses (Andreani et al., 2014; Carminati et al., 2019; Carrascosa et al., 2015, 2021; Chierici et al., 2016; Martin et al., 2011; Sechi et al., 2011). This discoloration has been associated with strains belonging to the phylogenetic group of *Pseudomonas fluorescens* and it is mainly characterized by a pigment that is usually produced at low temperatures (Andreani et al., 2014; Carrascosa et al., 2015; Rossi et al., 2018).

Previous studies on the blue discoloration of cheese have been focused on identifying metabolic pathways leading to the formation of blue pigment. The use of whole-genome sequencing (WGS) and comparative genomics has allowed for more robust investigations into this problem. Andreani et al. (2015) performed a pangenome analysis and demonstrated that blue-pigmenting strains had copies of genes involved in the tryptophan biosynthesis pathway (*trpABCDF*). Reichler et al. (2019), using comparative genomics, also noted the presence of an accessory locus encoding tryptophan biosynthesis genes, present in all isolates that produced grey or blue pigment.

However, we have observed that the production of this pigment and subsequent genomic identification of the accessory locus *trpABCDF* is not species-dependent, as it has already been detected in isolates identified as *Pseudomonas lactis*, *Pseudomonas putida* (Reichler et al., 2019), and *P. fluorescens* (Andreani et al., 2014, 2019). Also, the emergence of new pigment-producing species, such as *Pseudomonas carnis* (Lick et al., 2020), which have already been reported to be responsible for blue discoloration in cheese (Rodrigues et al., 2021), shows that there are genes conserved among different species of *Pseudomonas* that are responsible for blue

pigment synthesis. We therefore hypothesize that genes conserved between different species may come from mobile genetic elements. A characterization of genomes could provide us with further insights.

To further identify the genes related to this blue discoloration, we conducted a genomic characterization of the draft genomes of *Pseudomonas* sp. A006 and *P. carnis* B157, and genomic comparison with 11 pigment-producing strains against 4 non-producing strains.

2. Material and methods

2.1. Retrieval of genome data and bacterial strains

Pseudomonas sp. A006 and *P. carnis* B157 were isolated from Minas Frescal cheese, a fresh soft cheese produced in Brazil, and the whole genome was sequenced, trimmed, and assembled by Rodrigues et al. (2021). We retrieved the draft genomes of A006 and B157 strains from the GenBank database under accession numbers JAEFBF000000000 and JAEFBE000000000, respectively. The draft genomes were annotated using Prokka v.1.14.5 software (Seemann, 2014) with the default settings. The Prokka pipeline includes the prediction of coding sequences using Prodigal v.2.6.3 (Hyatt et al., 2010), identification of transfer RNA genes and transfer-messenger RNA genes with Aragorn v.1.2.38 (Laslett & Canback, 2004) and of non-coding RNA with Infernal v.1.1.2 (Kolbe & Eddy, 2011), and prediction of ribosomal RNA with Barrnap v.0.9 (https://github.com/tseemann/barrnap).

The relatedness of strain A006 was determined with the current species identified for *Pseudomonas*, and the following approaches were used: average nucleotide identity (ANI), based on the BLASTn algorithm (ANIb) and the MUMMER ultra-rapid aligning tool (ANIm); tetranucleotide frequency correlation coefficients (TETRA); digital DNA–DNA hybridization

(dDDH). ANIb, ANIm, and TETRA frequencies were calculated using JSpeciesWS (Richter, Rosselló-Móra, Oliver Glöckner, & Peplies, 2016). The dDDH values were calculated using GGDC 2.1 (Genome-to-Genome Distance Calculator; Meier-Kolthoff et al., 2013), with the recommended BLAST method and recommended formula 2. Additionally, a phylogenomic tree was built using the Type Strain Genome Server (TYGS; Meier-Kolthoff & Göker, 2019), and the phylogenetic tree was visualized using iTol v.4.2 (Letunic & Bork, 2016).

2.2. Functional annotation

Gene Ontology (GO) terms were assigned using Blast2GO Basic v.5.2.5 software (Conesa et al., 2005) based on BLASTp searches against the NR database with an e-value cutoff of 1e-10, and using InterProScan against the InterPro collection of databases. For graphical presentation and comparison of the functional annotations of the two strains, the GO terms were functionally classified using the online tool WEGO 2.0 (Ye et al., 2006, 2018). Annotation with Clusters of Orthologous Groups of proteins (COG) was performed using the Batch CD-search tool in NCBI (Marchler-Bauer & Bryant, 2004), with default parameters (expected value threshold of 0.01, composition-corrected scoring turned on, and maximum number of hits of 500). Then, the results were assigned to COG categories, based on the files 'cog-20.def.tab' (file with COG descriptions) and 'fun-20.tab' (file with descriptions of COG functional categories; https://ftp.ncbi.nih.gov/pub/COG/COG2020/data/).

Pathway analysis was performed by KEGG Orthology (KO) annotation using the KEGG Automatic Annotation Server (Kanehisa & Goto, 2000; Moriya, Itoh, Okuda, Yoshizawa, & Kanehisa, 2007), with the parameters: BLAST ('search program'); all the available *Pseudomonas* marked in 'genes data set'. The KO files were used to retrieve the pathways, KO description, and EC numbers associated with the genes after extraction of information from

KEGG databases via the KEGG REST API (https://www.kegg.jp/kegg/rest/keggapi.html), using ‘link’ and ‘list’. KO numbers were also mapped to pathways using KEGG Mapper (Kanehisa & Sato, 2020). For graphical presentation and comparison of the KEGG pathways, COG classification, and GO terms, graphics were made using R software (Version 4.0.2), with the ‘ggplot2’ (Wickham, 2016) and ‘ggstance’ (Henry, Wickham, & Chang, 2020) packages. antiSMASH (ANTIbiotics & Secondary Metabolite Analysis Shell; Blin et al., 2019) was used to predict the biosynthetic gene clusters (BGCs) for secondary metabolites, with the detection strictness parameter in default mode and all extra features on.

2.3. Identification of mobile genetic elements

PlasmidSPAdes from SPAdes v.3.12.0 (Antipov et al., 2016; Bankevich et al., 2012) was used to obtain possible plasmid DNA sequences from trimmed, paired-end reads (parameter ‘--careful’ and the others in default). The obtained scaffolds were analysed on the PlasmidFinder 2.0 server (Carattoli et al., 2014) and using Mash dist v.2.1.1 (Ondov et al., 2016) against the PLSDB plasmid database v.2020_11_1 (Galata, Fehlmann, Backes, & Keller, 2019). The putative plasmid sequences were compared with the closest identified plasmid (downloaded from GenBank) using BLASTn and progressiveMauve (Darling et al., 2010). This allowed for the identification of scaffolds that were not assigned to the identified plasmids. Integrative and conjugative elements (ICEs) and integrative and mobilizable elements (IMEs) were predicted using the ICEfinder web-based tool with the default parameters, and predicted sequences were searched via BLASTn against the ICEberg database v.2.0 (Liu et al., 2019). Then, the region containing the putative ICE/IME was aligned with the closest identified sequences by ICEberg, using the BLASTn algorithm (https://blast.ncbi.nlm.nih.gov/Blast.cgi). In addition, the oriTfinder tool (Li et al., 2018) was employed to identify the origin of transfer

site (*oriT*) on sequences. Figures comparing the closest identified plasmids with putative plasmids or ICE/IME were generated using BLAST Ring Image Generator (BRIG), with an e-value of 1e-10 (Alikhan, Petty, Ben Zakour, & Beatson, 2011).

Phage-related sequences were detected in each draft genome using VirSorter v.1.0.6 (Roux, Enault, Hurwitz, & Sullivan, 2015) and the output was filtered using the PHAge Search Tool (PHASTER; Arndt et al., 2016). VirSorter was run using the reference database of viral genomes of archaeal and bacterial viruses from RefSeq (RefSeqABVir), and only categories 1 and 2 were considered.

2.4. Identification of pigment biosynthesis

To identify potential pigments, six biosynthetic pathways were investigated: pyoverdine, pyochelin, phenazines inducing pyocyanin, violacein, indigoidine, and tryptophan. Three approaches were utilized. First, protein sequences of genes related to the biosynthetic pathways of interest were downloaded from NCBI and a local protein database was created using the 'makeblastdb' tool, available at NCBI-BLAST v.2.9.0+ (Camacho et al., 2009). Predicted proteins from each strain were used as the BLASTp query. The 'awk' command was used to filter the output file with the following cutoff values: query coverage ≥ 70%, e-value ≤ 1e-10, and identity ≥ 60%. The second approach was to verify the predicted proteins identified as possible genes of the biosynthetic pathways studied with UniProt-TrEMBL. A database was built using all downloaded UniProt-TrEMBL sequences (ftp://ftp.uniprot.org/pub/databases/uniprot/current_release/knowledgebase/complete/uniprot_trembl.fasta.gz). A local BLASTp was performed using the same cutoff parameters described above. Finally, the HMMER package v.3.2.1 (Eddy, 1998) was used to search Pfam profiles (downloaded from the Pfam database;

ftp://ftp.ebi.ac.uk/pub/databases/Pfam/releases/Pfam33.1/Pfam-A.hmm.gz), using 'hmmpress' and 'hmmscan' tools (Eddy, 2011; El-Gebali et al., 2019), using the maximum sensitivity mode and an e-value of < 1e-5 to allow evaluation of the protein or domain families.

2.5. Comparative genomic analysis

To allow for the identification of possible genomic regions contained only in blue or black pigment producers, OrthoFinder v.2.3.3 with default parameters (Emms & Kelly, 2015) was used to help determine groups of orthologous gene pairs (orthogroups). Predicted proteomes from the genomes studied here and from other previously identified strains were used (nine pigment-producing and four non-pigment-producing strains; Table 1). All genomes were annotated using Prokka v.1.14.5 (Seemann, 2014) for standardization. Orthogroups were visualized as UpSet plots using R software (Version 4.0.2), with the 'UpSetR' package (Conway, Lex, & Gehlenborg, 2017). Besides that, dDDH was calculated using GGDC (Meier-Kolthoff et al., 2013) to analyse genomic similarity and determine relatedness between strains. The 'ggplot2' package (Wickham, 2016) was used to create the heatmap for the results.

Table 1. Summary table of genomes used for comparative genomic analyses.

Strains	GenBank - ID	Source	Reference
Blue/black pigment producers			
Pseudomonas fluorescens Ps_22	LCYA00000000	Blue mozzarella cheese	Andreani et al., 2015
Pseudomonas fluorescens Ps_77	LCYB00000000	Blue beef	
Pseudomonas carnis B4-1^T	CABIVL000000000	Refrigerated raw poultry	Lick et al., 2020
Pseudomonas paracarnis UBT403	CAJFCM000000000	Minced beef	Lick et al., 2021
Pseudomonas fluorescens ITEM 17298	NPKB00000000	Mozzarella cheese	Fanelli et al., 2017
Pseudomonas lactis FSL E2-0548	RXHZ00000000	Pasteurized milk	Reichler et al., 2019
Pseudomonas lactis FSL E2-8864	RWKC00000000	Pasteurized milk	
Pseudomonas lactis FSL R10-2514	RWKD00000000	Pasteurized milk	
Pseudomonas lactis FSL W5-0203	RXHY00000000	Queso fresco cheese curds	
Non-blue/black pigment producers			
Pseudomonas fluorescens Ps_20	LCYD00000000	UHT-milk	Andreani et al., 2015
Pseudomonas fluorescens Ps_40	LCYC00000000	Ricotta cheese	
Pseudomonas fluorescens A506	CP003041	Pear tree leaf	Loper et al., 2012
Pseudomonas lactis SS101	CM001513	Wheat roots	

The protein-coding genes *trpCBAFD1* (central genome) and *trpCBAFD2* (accessory) were used for phylogenetic analysis. The nucleotide sequences were separately aligned using the WebPrank web server (Löytynoja & Goldman, 2010) and trimmed using the Gblocks server (Castresana, 2002). The genes were concatenated (according to identification: central genome or accessory). Phylogenetic analysis of the concatenated sequences was performed in MEGA v. 11 (Tamura, Stecher, & Kumar, 2021), with bootstrap analyses for 1,000 replicates (Felsenstein, 1985). This software was used to identify the most appropriate substitution model for the maximum likelihood (ML) method (Felsenstein, 1981). The phylogenetic tree was visualized using iTol v.4.2 (Letunic & Bork, 2016).

A BLASTn comparison using Easyfig v.2.2.6 software (Sullivan, Petty, & Beatson, 2011) was performed between unique genomic regions identified in discoloration-causing strains.

3. Results

3.1. Identification

The strain A006 was identified here as belonging to the species *Pseudomonas paracarnis*, a new species recently described (Lick et al., 2021). The genome relatedness values of strain A006 with the only two genomes of *P. paracarnis* available at NCBI, based on ANIb, ANIm, TETRA, and GGDC, were above the established cutoff for each algorithm (95%, 95%, 0.999, and 70%, respectively). The relatedness values with *P. paracarnis* UBT403 (GenBank: CAJFCM00000000) were 99.62%, 99.88%, 0.99962, and 98.90%, respectively, and with *P. paracarnis* V5/DAB/2/5 (GenBank: CAJFCN00000000) were 98.63%, 99.13%, 0.99949, and 92.50%, respectively. Based on the TYGS genome comparison, using the genomes of the type strains of the closest related species (except *P. paracarnis* UBT403, which is not a type strain), strain A006 was included in the cluster of *P. paracarnis* (Fig. S1, Supplementary Material).

3.2. Genomic features and functional annotation

Annotation using Prokka software revealed the genic content of sequences. The draft genomes annotated had the following content: 5,951 protein-coding sequences, 72 tRNAs, 16 rRNAs, and 1 transfer-messenger RNA (tmRNA); and 5,958 protein-coding sequences, 66 tRNAs, 10 rRNAs, and 1 tmRNA, for strains A006 and B157, respectively.

A total of 4,673 (78.5%) predicted genes of *P. paracarnis* A006 were annotated by GO terms, COG categories, and/or KEGG pathways (Table 2), with 2,260 (38.0%) being annotated in all three databases. Of these, we obtained the annotation of 4,093 (68.8%) for GO terms, which were categorized into 45 functional groups (Fig. S2, Supplementary Material). The most abundant were ‘Metabolic process’ and ‘Cellular process’ within ‘Biological Process’ and

'Catalytic activity' within 'Molecular Function'. A total of 3,386 genes had a COG classification (56.9%). Among the 24 categories of COG obtained (Fig. S3, Supplementary Material), the category [E] for 'Amino acid transport and metabolism' represented the largest group followed by [K] for 'Transcription'. Categories [A] ('RNA processing and modification') and [Z] ('Cytoskeleton') represented the smallest groups. Of the 3,135 genes annotated with KO (52.7%), we assigned 2,436 of them to 255 KEGG pathways, which were divided into 6 functional categories and 44 subcategories (Fig. S4, Supplementary Material).

Table 2. Gene functional annotation statistics of *Pseudomonas paracarnis* A006 and *Pseudomonas carnis* B157 for Gene Ontology (GO), Clusters of Orthologous Genes (COG), and KEGG Orthology (KO).

Database (DB)	A006		B157	
	Annotated genes	Percent	Annotated genes	Percent
GO	4,093	68.8%	3,975	66.7%
COG	3,386	56.9%	3,440	57.7%
KO	3,135	52.7%	3,166	53.1%
At least DB	4,673	78.5%	4,708	79.0%
In all DB	2,260	38.0%	2,206	37.0%
By none	1,278	21.5%	1,250	21.0%
Total genes	5,951	100.0%	5,958	100.0%

For strain *P. paracarnis* B157, a total of 4,708 (79.0%) predicted genes were annotated by GO terms, COG categories, and/or KEGG pathways (Table 2), with 2,206 (37.0%) being annotated in all three databases. We obtained the annotation of 3,975 (66.7%) and 3,440 (57.7%) genes for GO terms and COG categories, respectively. The most abundant GO terms and COG categories were the same as for strain A006 (Fig. S2 and Fig. S3, Supplementary Material). Of the 3,166 genes annotated with KO (53.1%), we assigned 2,457 of them to 253 KEGG pathways, which were also divided into 6 functional categories and 44 subcategories (Fig. S4, Supplementary Material).

AntiSMASH predicted 12 putative secondary metabolite clusters for *P. paracarnis* A006 and 15 for *P. carnis* B157 (Table S1, Supplementary Material). These included gene clusters for siderophore synthesis, non-ribosomal peptide synthase (NRPS)-related functions, aryl polyene, homoserine lactone (Hserlactone), beta-lactone-containing protease inhibitor (betalactone), and *N*-acetylglutaminylglutamine amide (NAGGN).

3.3. Identification of mobile genetic elements

PlasmidSPAdes assembled several scaffolds of possible plasmid sequences for both strains, but no hits were found using PlasmidFinder. However, we employed other tools and databases because PlasmidFinder is a database with plasmids from Gram-positive bacteria and *Enterobacteriaceae* (https://cge.cbs.dtu.dk/services/PlasmidFinder/). Using the plasmid database PLSDB, it was possible to identify similarities between the sequences obtained from strain A006 and the plasmid pD2RT from *Pseudomonas migulae* (NCBI Reference Sequence: NC_021250.1; Fig. 1A). For strain B157, two hits were verified: plasmid pCFR-1cb6 from *Citrobacter freundii* (NZ_CP026240.1) and plasmid pCAV1043-58 from *Enterobacter asburiae* (NZ_CP011588.1). An alignment using BRIG software with pCFR-1cb6 as a reference can be seen in Fig. 1B.

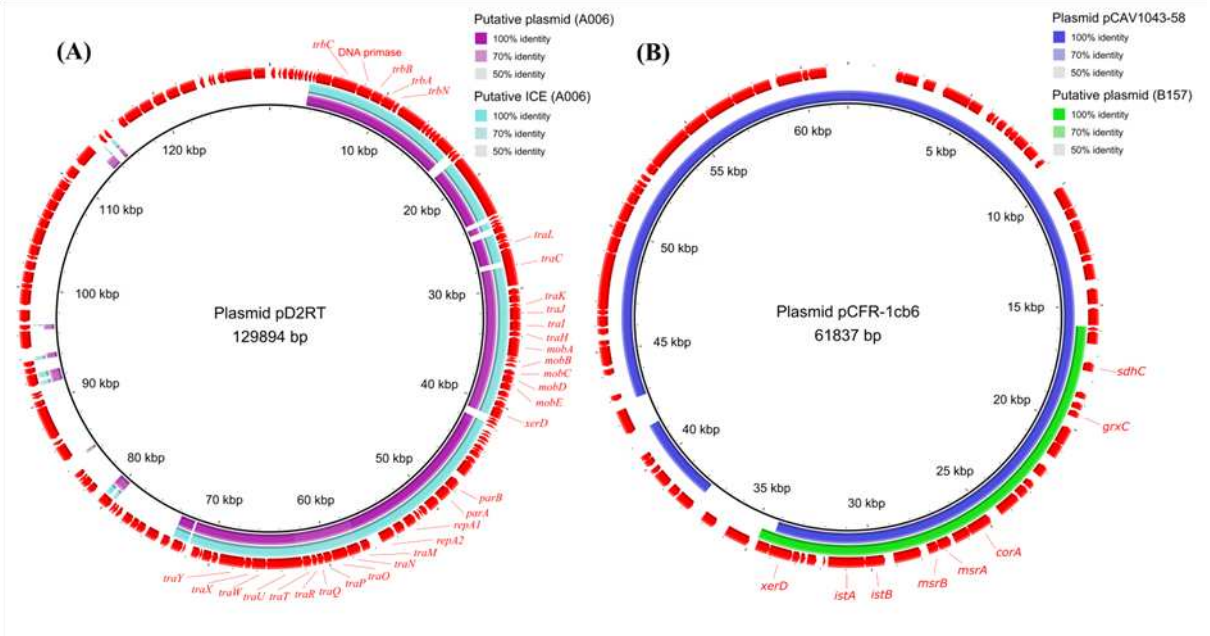


Fig. 1. Plasmid comparison using BLAST Ring Image Generator (BRIG). **(A)** The putative plasmid and putative integrative conjugative element (ICE), represented in purple and light blue, respectively, found for *Pseudomonas paracarnis* A006, were compared with the plasmid pD2RT, used as a reference. **(B)** The putative plasmid found for *Pseudomonas carnis* B157, represented in green, were compared with the plasmid pCAV1043-58 (dark blue) and the plasmid pCFR-1cb6, used as a reference. For both, the outer red ring refers to the annotation of reference, with some indicated genes.

One putative ICE of 139,269 bp was identified for A006 by ICEfinder and oriTfinder, containing genes related to type IV secretion system (T4SS), relaxase, integrase, and type IV coupling protein (T4CP), but not containing origin site of DNA transfer (*oriT*). oriTfinder, which is also able to identify predicted determinants of acquired antibiotic resistance, and predicted virulence factors, does not predict any of these in the ICE. As for strain B157, one IME of 8,623 bp was identified, containing only genes related to integrase (product: phage integrase) and relaxase, because those without T4SS but with integrase and relaxase are tagged as putative IMEs (Liu et al., 2019). Furthermore, it was only possible to identify sequences close to that from the ICEberg database with the ICE sequence of strain A006. The closest sequence to the identified ICE was that of *Pseudomonas aeruginosa* UCBPP-PA14 (GenBank CP000438.1). In order to identify closer sequences, a BLASTn search was performed against

the non-redundant nucleotide database of NCBI. The alignment showing the highest query coverage was that with *P. migulae* strain D2RT plasmid pD2RT, with an e-value of 0.0 and percentage identity of 97.79%. It is the same plasmid sequence previously identified with the PLSDB plasmid database as close to the predicted plasmid sequence (Fig. 1A).

Prophages and phage sequences were identified by VirSorter and validated by PHASTER. We identified one complete phage contig and two predicted prophages for strain A006 and two prophages for strain B157. We removed those that were predicted as 'not so sure' (category 3) by VirSorter. The sequences considered to be 'somewhat sure' (category 2) by VirSorter and then indicated as 'questionable' or 'incomplete' by PHASTER were also removed from the results. The only exception was for prophages identified by VirSorter but which were divided into two phages by PHASTER, one 'intact' and the other 'questionable'. However, the final nucleotide sequences were those obtained by VirSorter. A summary of the genetic features of predicted intact prophages and phages per genome are shown in Table S2 (Supplementary Material).

3.4. Identification of pigment biosynthesis

No genes associated with violacein and indigoidine biosynthesis were found for strains A006 and B157. For phenazines/pyocyanin, only one gene related to biosynthesis was identified (*phzF*), and the same was observed for pyochelin (*pchB*). We identified 20 genes and operons for the biosynthesis of pyoverdine (Fig. 2) and 14 genes for the biosynthesis of tryptophan (Fig. 3) as well as duplication of the *trpA*, *trpB*, *trpC*, *trpD*, and *trpF* genes.

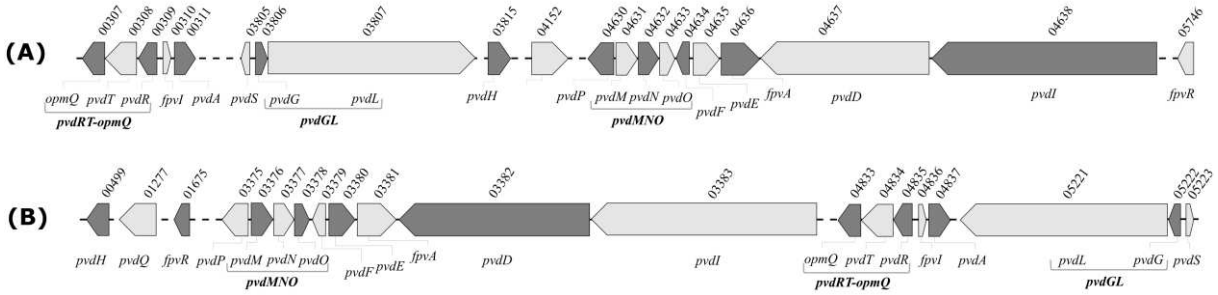


Fig. 2. Schematic representation of gene clusters and genes involved in biosynthesis of pyoverdine, found in the draft genome of **(A)** *Pseudomonas paracarnis* A006 and **(B)** *Pseudomonas carnis* B157. Annotated gene number is reported for each strain.

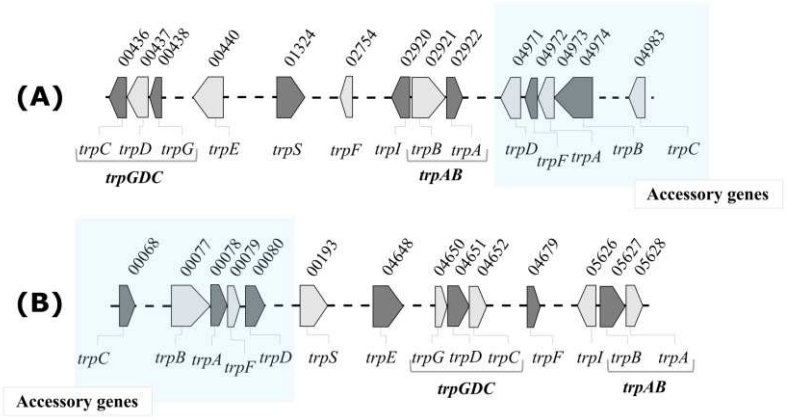


Fig. 3. Schematic representation of gene clusters and genes involved in biosynthesis of tryptophan, found in the draft genome of **(A)** *Pseudomonas paracarnis* A006 and **(B)** *Pseudomonas carnis* B157. Annotated gene number is reported for each strain.

3.5. Comparative genomic analysis

We performed an orthogroup analysis to identify genomic regions contained only in the genomes of blue or black pigment-producing strains. Using OrthoFinder, we identified a total of 29 unique orthogroups for the 11 pigment-producing strains (Fig. 4). However, we did not observe a phylogenomic relationship between them, based on the calculation of *in silico* dDDH with the GGDC web tool for genome-to-genome comparison. The established same-species delineation thresholds of 70% (Meier-Kolthoff et al., 2013) were used for identification of genomic similarity. The dendrogram generated with dDDH values in Fig. 5 displays that in the

three large clades created there are strains that do not cause blue discoloration along with those that do.

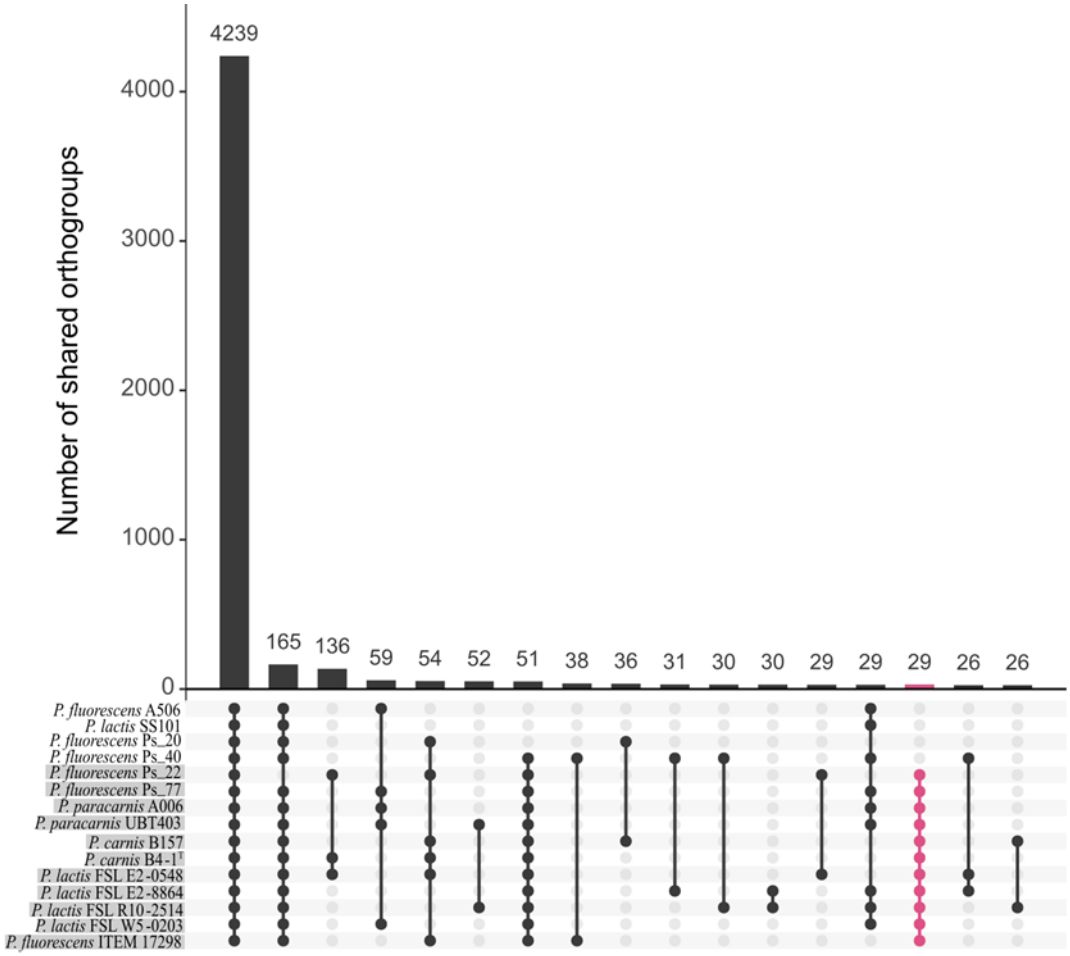


Fig. 4. Plot visualizing shared orthologous protein groups. The 29 shared orthogroups across all blue or black pigment producers are highlighted in pink. Only 17 sets of orthogroups are shown, the rest are omitted.

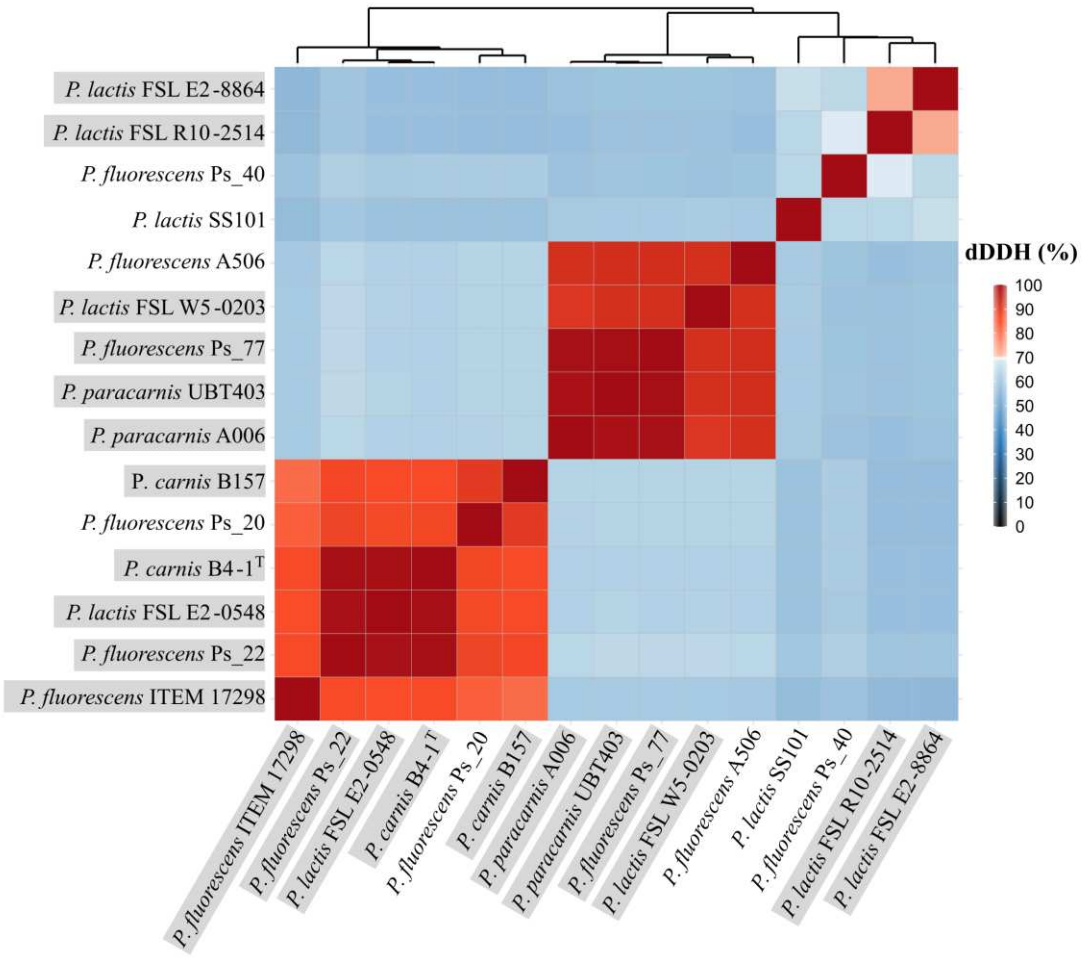


Fig. 5. Heatmap of *in silico* DNA-DNA hybridization (dDDH). The similarity values are presented in a color gradient heatmap (red gradients > 70; white = 70; blue gradients < 70). Values of >70% group together genomes from strains of the same species. Drawn dendrograms can be used to visualize the difference between strains. Discoloration-causing strains are highlighted in gray.

Of the 29 unique orthogroups identified for the 11 strains (Fig. 4), 4 orthogroups contained the proteins of previously identified as accessory genes of tryptophan biosynthesis (*trpACDF*). Thus, we selected those draft genomes that had these 29 unique genes in the same contig and without paralogs to check all the surrounding genes. The strains *P. carnis* B4-1, *P. fluorescens* ITEM 17298, *P. fluorescens* Ps_77, *P. lactis* E2-0548, and *P. lactis* E2-8864 were selected. We verified that there is a genomic region with 74 homologous genes among the strains, which includes the genes of the 29 unique identified orthogroups, and an orthogroup containing TrpB protein.

In order to verify if duplicate *trp* genes could have arisen from some duplication event, we studied the phylogenetic relationship of *trpCBAFD* genes from the central genome (*trp1*) and the accessory genome (*trp2*), including *trp* genes from strains that do not have their duplication. Genes from strain Ps_22 were not included because this draft genome lacks the *trpB* gene in the core genome. The phylogenetic tree constructed with a concatenated alignment was divided into two major clades: *trp1* and *trp2* (Fig. 6). It displayed that there is an evolutionary relationship between *trp2* genes and that they possibly evolved from a common ancestor. Furthermore, we observed that the *trp* genes of the central genome have the same length (*trpA*: 810 bp; *trpB*: 1,227 bp; *trpC*: 837 bp; *trpD*: 1,050 bp; *trpF*: 633 bp) for all strains studied here but differ from the accessory genes present in the strains causing discoloration (*trpA*: 801 bp; *trpB*: 1,887 bp; *trpC*: 789 bp; *trpD*: 966 bp; *trpF*: 603 bp).

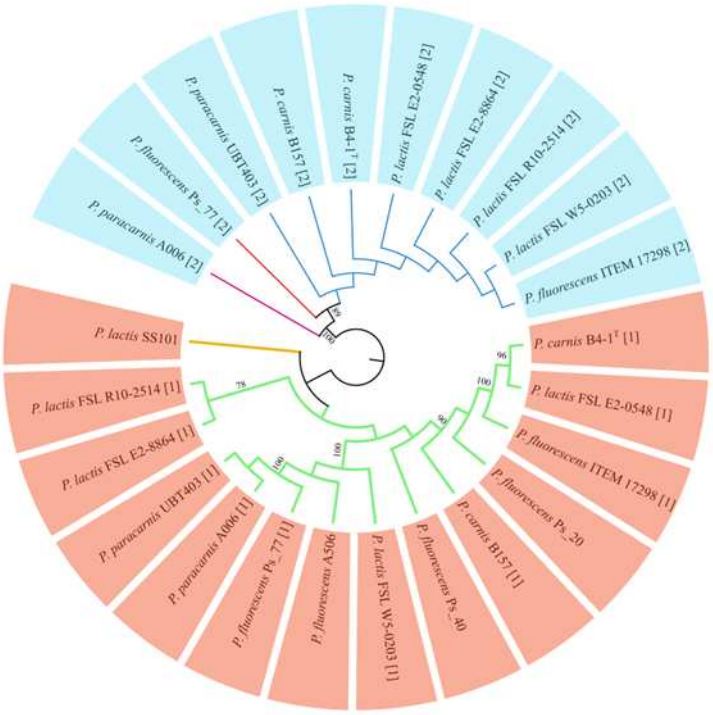


Fig. 6. Phylogenetic tree constructed with a concatenated alignment of genes sequences from *trpCBAFD* genes from the central genome (*trp1* [1]) and the accessory genome (*trp2* [2]). The evolutionary history was inferred by using the Maximum Likelihood method and Tamura 3-parameter model (K. Tamura, 1992). A discrete Gamma distribution was used to model evolutionary rate differences among sites (5 categories (+G, parameter = 0.4957)). The percentage of trees in which the associated taxa clustered together (≥70%) is shown below the branches and

considering this value of minimum for the formation of clades, internal colors highlight grouped clades. The two large clades colored in red and blue highlight the *trp1* and *trp2* genes, respectively.

The unique genomic region identified that contains 74 genes has a length (bp) ranging from 66,080 to 66,099 (Table S3, Supplementary Material). Interestingly, it has a lower GC content (%) than the draft genomes, ranging from 55.67 to 55.68, with that of the studied genomes ranging from 59.66 to 59.78. A BLASTn comparison performed using Easyfig software showed a high degree of similarity at nucleotide level between the five identical unique genomic regions studied, of the five selected strains (Fig. S5, Supplementary Material).

We annotated the 74 proteins from coding sequences in this region using BLASTp, an NCBI BLAST tool (https://blast.ncbi.nlm.nih.gov/Blast.cgi), against the reference protein database (cutoff values: query coverage $\geq 70\%$, e-value $\leq$ 1e-10, and identity $\geq 60\%$). A schematic representation of protein-coding gene organization in that genomic region is shown in Fig. 7. We found several genes encoding proteins involved with mobile elements and horizontal gene transfer. Altogether, we identified seven genes that encode proteins belonging to families of ICEs (TIGR03752, TIGR03749, TIGR03746, TIGR03745, TIGR03758, TIGR03747, and TIGR03759) and two genes that encode proteins belonging to families of conjugal transfer (TIGR03751 and TIGR03750). We also observed the presence of protein-coding genes for type IV pilus biogenesis (PilM), one involved in the conjugative transfer of pilus (PilV), pilus assembly (TadA and PilX), type II secretion system F family, and coupling of the type IV conjugative transfer system (TraD). Another notable result is an annotated protein containing an autoinducer binding domain, which could be related to bacterial communication by quorum sensing. However, 23 proteins from the unique region were annotated with hypothetical proteins.

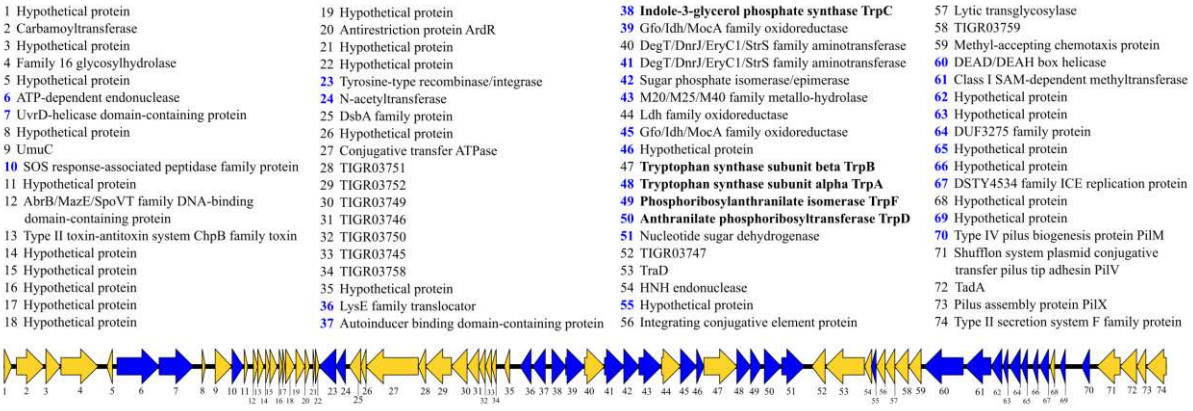


Fig. 7. Schematic representation of protein-coding genes organization in unique genomic region identified in drafts genomes of discoloration-causing strains. The direction of the arrows represents the coding strand of the coding sequence. Number below the arrows correspond to annotation, shown at the top of the figure. The arrows and text colored in blue represent 29 unique genes identified by orthogroup analysis.

4. Discussion

In recent years, species belonging to the genus *Pseudomonas*, such as *P. fluorescens* (Andreani et al., 2015) and *P. lactis* (Reichler et al., 2019), have been responsible for blue discoloration in different types of cheese. However, we found in a recent study (Rodrigues et al., 2021) that an isolate responsible for cheese discoloration belonged to *P. carnis*, which is associated with meat spoilage (Lick et al., 2020). Moreover, a new species of *Pseudomonas* with a high degree of genomic similarity to *P. carnis*, *P. lactis*, and *Pseudomonas paralactis* has been described: *P. paracarnis* (Lick et al., 2021). The strain A006 studied here showed a high degree of similarity with these species (Rodrigues et al., 2021). We investigated and confirmed that strain A006 belongs to this new species. The high degree of heterogeneity of the genus *Pseudomonas*, especially the phylogenetic group of *P. fluorescens*, has made it difficult to correctly assign strains to species (Mulet, Lalucat, & García-Valdés, 2010). Conducting studies using whole genome sequences has allowed for correct taxonomies (Garrido-Sanz et al., 2016; Gomila et al., 2015).

Understanding genetic differences between strains allows the identification of possible relationships with specific phenotypic traits. Some studies have linked the appearance of the blue pigment with genes in the tryptophan biosynthesis pathway involved in indole production (*trpABCDF* genes; Andreani et al., 2015, 2019; Reichler et al., 2019). Andreani et al. (2015, 2019) claim that the chromogenic reaction of pigment producers with Kovac's reagent indicates that an indole derivative is a precursor of the blue pigment. Here, we observed that strains A006 and B157 were predicted to have genes pertaining to tryptophan metabolism and tryptophan biosynthesis in amino acid metabolism. When we investigated genes that had previously been linked to pigment production, we found that the draft genomes had duplicate genes related to tryptophan biosynthesis (Fig. 3). Andreani et al. (2019) used transposon mutagenesis to identify *trp* genes as being involved in the production of blue pigment by *P. fluorescens* Ps_77. However, in a previous study employing genomic analysis, Andreani et al. (2015) showed the absence of the tryptophanase gene (*tnaA*) that encodes the enzyme responsible for converting tryptophan to indole (observed in indole-producing bacteria). We also have not identified the presence of this gene.

There are other hypotheses that relate the production of these pigments to additional metabolic pathways. Andreani et al. (2015) reported that in strains that produce blue pigment, the genes involved in iron uptake and in the biosynthesis of the siderophore pyoverdine are more expressed. Lippolis et al. (2021) studied the proteome of *P. fluorescens* pf59, isolated from mozzarella cheese, in response to different environmental conditions and found a relationship between the appearance of blue colour and the overexpression of proteins involved in redox balance and osmotic-shock defence response. Quintieri et al. (2020) also studied the proteome of a blue pigment-producing strain that was isolated from mozzarella cheese, *P. fluorescens* ITEM 17298 (included here in this work). They identified proteins associated with pyomelanin synthesis: proteins involved in the conversion of *L*-tyrosine to homogentisate, a

precursor of pyomelanin in the extracellular environment. Andreani et al. (2014) showed, using multilocus sequence typing (MLST) analysis, that strains characterized as pyomelanin producers are phylogenetically distant from those that produce blue discoloration of cheese. However, we found that the strain *P. fluorescens* ITEM 17298 is grouped into the so-called 'blue branch' by *in silico* MLST (data not shown), and that this strain also shares a region with strains that have previously been identified as producing an indigo derivative, which would be responsible for the blue discoloration.

We also identified genes encoding proteins involved with mobile elements and horizontal gene transfer (Fig. 7). The presence of mobile genetic elements has already been reported in another study. Chierici (2016) performed comparative genomic analysis and a total of 24 genes were recognized as specific for the 5 pigment-producing strains. Among these genes, the researcher evidenced the presence of genes from bacteriophage structures. However, we identified here that a larger region, with 74 genes, is conserved in discoloration-causing strains, and it presents protein-coding sequences related to DNA uptake from an extracellular source (Piepenbrink, 2019), such as the genes coding for PilM, PilV, and PilX which are essential for the formation of functional type IV pili (Asikyan, Kus, & Burrows, 2008). Furthermore, the unique genomic region identified in discoloration-causing strains shows a GC content below the average GC content of the genomes (Table S3, Supplementary Material). Mobile genetic elements usually have a lower GC content than their bacterial hosts, and this is indicative of the presence of recently acquired mobile elements (Millan et al., 2015; Nishida, 2012). Probably, this region has not come from an ancestor, since it is not widespread in all members of the species of *P. lactis* and *P. paracarnis* (Lick et al., 2021; Von Neubeck et al., 2017). Therefore, one hypothesis is that there was a horizontal acquisition of the region and that a phylogenetic proximity between the strains already described allowed this to happen.

The putative plasmid of strain A006 showed similarity to pD2RT, a toluene-degrading plasmid (Jutkina et al., 2013). Therefore, the main constituents of this plasmid are genes responsible for the degradation of toluene (*xyl*), which were not found in the putative plasmid and putative ICE identified for the A006 strain (Fig. 1A). Similarity was only verified for a region that contained genes encoding transposition and integration elements, and putative plasmid backbone functions, such as the *repA* gene, which was identified by Jutkina et al. (2013) as responsible for plasmid replication. The putative plasmid for strain B157 showed similarity with a small region of the plasmids pCFR-1cb6 and pCAV1043-58. Fang et al. (2018) showed that plasmid pCAV1043-58 has an *msrAB* region which has been characterized as an accessory module harbouring resistance genes. The *msrAB* region is involved in resistance to oxidative damage, because *msrAB* genes encode the enzyme peptide methionine sulphoxide reductase which catalyses the reduction of oxidized methionine residues on proteins (Skaar et al., 2002; Stohl, Criss, & Seifert, 2005). In Fig. 1B it can be seen that the similarity between the putative plasmids of B157 and the plasmids pCFR-1cb6 and pCAV1043-58 was exactly in this region.

In addition, to further characterize the strains, and in order to help predict possible secondary metabolites, we investigated their protein-coding regions using antiSMASH. We have identified a cluster for aryl polyenes, a class of yellow pigments related to antioxidative carotenoids (Schöner et al., 2016), and NAGGN is related to osmotic stress (Sagot et al., 2010). We also found a predicted pyoverdine biosynthetic gene cluster. Pyoverdine is a yellow-green fluorescent siderophore (Meyer & Hornsperger, 1978). By investigating the predicted proteins of draft genomes, we have identified 20 genes related to biosynthesis, secretion, and uptake of pyoverdine, classified into: genes for biosynthesis (*pvdLG*, *pvdP*, *pvdE*, *pvdQ*, *pvdF*, *pvdA*, *pvdH*, *pvdD*, *pvdI*), regulation (*fpvI*, *fpvR*, and *fvdS* – sigma factor), modification (*pvdMNO*),

transport systems to secretion (*pvdRT-opmQ*), and ferripyoverdine receptor (*fpvA*) (Lamont & Martin, 2003; Ringel & Brüser, 2018).

5. Conclusions

In conclusion, the investigations described here present a firm foundation for determining the unique genomic region in blue discoloration-causing strains which contains conserved coding sequences for proteins related to horizontal gene transfer and tryptophan biosynthesis. To improve our understanding of blue discoloration-causing strains, it is necessary to study and explore genomes, and this can also help to design more effective strategies for the control and detection of spoilers in dairy products, such as advances in their detection by designing specific primers. Further studies should focus on sequencing and assembling the complete genome of a discoloration-causing strain, as it would allow for more genomic clarification, as well as help several researchers working with omics approaches.

Research data

The dataset and relevant documents pertaining to analyses are publicly available, via the Open Science Framework (Foster & Deardorff, 2017), and can be found at https://osf.io/trehz/?view_only=7c56e214da354c52bdb991e608bd2c0d.

References

Alikhan, N. F., Petty, N. K., Ben Zakour, N. L., & Beatson, S. A. (2011). BLAST Ring Image Generator (BRIG): Simple prokaryote genome comparisons. *BMC Genomics*, *12*(402). https://doi.org/10.1186/1471-2164-12-402

Andreani, N. A., Carraro, L., Martino, M. E., Fondi, M., Fasolato, L., Miotto, G., … Cardazzo, B. (2015). A genomic and transcriptomic approach to investigate the blue pigment phenotype in *Pseudomonas fluorescens*. *International Journal of Food Microbiology*, *213*, 88–98. https://doi.org/10.1016/j.ijfoodmicro.2015.05.024

Andreani, N. A., Carraro, L., Zhang, L., Vos, M., & Cardazzo, B. (2019). Transposon mutagenesis in *Pseudomonas fluorescens* reveals genes involved in blue pigment production and antioxidant protection. *Food Microbiology*, *82*, 497–503. https://doi.org/10.1016/j.fm.2019.03.028

Andreani, N. A., Martino, M. E., Fasolato, L., Carraro, L., Montemurro, F., Mioni, R., … Cardazzo, B. (2014). Tracking the blue: A MLST approach to characterise the *Pseudomonas fluorescens* group. *Food Microbiology*, *39*, 116–126. https://doi.org/10.1016/j.fm.2014.11.009

Antipov, D., Hartwick, N., Shen, M., Raiko, M., Lapidus, A., & Pevzner, P. A. (2016). PlasmidSPAdes: Assembling plasmids from whole genome sequencing data. *Bioinformatics*, *32*(22), 3380–3387. https://doi.org/10.1093/bioinformatics/btw493

Arndt, D., Grant, J. R., Marcu, A., Sajed, T., Pon, A., Liang, Y., & Wishart, D. S. (2016). PHASTER: a better, faster version of the PHAST phage search tool. *Nucleic Acids Research*, *44*, W16–W21. https://doi.org/10.1093/nar/gkw387

Asikyan, M. L., Kus, J. V, & Burrows, L. L. (2008). Novel proteins that modulate type IV pilus

retraction dynamics in *Pseudomonas aeruginosa*. *Journal of Bacteriology*, *190*(21), 7022–7034. https://doi.org/10.1128/JB.00938-08

Bankevich, A., Nurk, S., Antipov, D., Gurevich, A. A., Dvorkin, M., Kulikov, A. S., … Pevzner, P. A. (2012). SPAdes: A new genome assembly algorithm and its applications to single-cell sequencing. *Journal of Computational Biology*, *19*(5), 455–477. https://doi.org/10.1089/cmb.2012.0021

Blin, K., Shaw, S., Steinke, K., Villebro, R., Ziemert, N., Lee, S. Y., … Weber, T. (2019). AntiSMASH 5.0: Updates to the secondary metabolite genome mining pipeline. *Nucleic Acids Research*, *47*, W81–W87. https://doi.org/10.1093/nar/gkz310

Camacho, C., Coulouris, G., Avagyan, V., Ma, N., Papadopoulos, J., Bealer, K., & Madden, T. L. (2009). BLAST+: Architecture and applications. *BMC Bioinformatics*, *10*(421). https://doi.org/10.1186/1471-2105-10-421

Carattoli, A., Zankari, E., Garciá-Fernández, A., Larsen, M. V., Lund, O., Villa, L., … Hasman, H. (2014). *In Silico* detection and typing of plasmids using plasmidfinder and plasmid multilocus sequence typing. *Antimicrobial Agents and Chemotherapy*, *58*(7), 3895–3903. https://doi.org/10.1128/AAC.02412-14

Carminati, D., Bonvini, B., Rossetti, L., Zago, M., Tidona, F., & Giraffa, G. (2019). Investigation on the presence of blue pigment-producing *Pseudomonas* strains along a production line of fresh mozzarella cheese. *Food Control*, *100*, 321–328. https://doi.org/10.1016/j.foodcont.2019.02.009

Carrascosa, C., Martínez, R., Sanjuán, E., Millán, R., del Rosario-Quintana, C., Acosta, F., … Jaber, J. R. (2021). Identification of the *Pseudomonas fluorescens* group as being responsible for blue pigment on fresh cheese. *Journal of Dairy Science*, *104*(6). https://doi.org/10.3168/jds.2020-19517

Carrascosa, C., Millán, R., Jaber, J. R., Lupiola, P., del Rosario-Quintana, C., Mauricio, C., & Sanjuán, E. (2015). Blue pigment in fresh cheese produced by *Pseudomonas fluorescens*. *Food Control*, *54*, 95–102. https://doi.org/10.1016/j.foodcont.2014.12.039

Castresana J. (2002). Gblocks, v. 0.91b. Retrieved January 6, 2022, from http://molevol.cmima.csic.es/castresana/Gblocks_server.html

Chierici, M. (2016). *Investigation on the blue phenotype in* Pseudomonas *species involved in blue discoloration defect of fresh cheese*. Università degli studi di Milano.

Chierici, M., Picozzi, C., La Spina, M. G., Orsi, C., Vigentini, I., Zambrini, V., & Foschino, R. (2016). Strain diversity of *Pseudomonas fluorescens* group with potential blue pigment phenotype isolated from dairy products. *Journal of Food Protection*, *79*(8), 1430–1435. https://doi.org/10.4315/0362-028X.JFP-15-589

Conesa, A., Götz, S., García-Gómez, J. M., Terol, J., Talón, M., & Robles, M. (2005). Blast2GO: A universal tool for annotation, visualization and analysis in functional genomics research. *Bioinformatics*, *21*(18), 3674–3676. https://doi.org/10.1093/bioinformatics/bti610

Conway, J. R., Lex, A., & Gehlenborg, N. (2017). UpSetR: An R package for the visualization of intersecting sets and their properties. *Bioinformatics*, *33*(18), 2938–2940. https://doi.org/10.1093/bioinformatics/btx364

Darling, A. E., Mau, B., & Perna, N. T. (2010). Progressivemauve: Multiple genome alignment with gene gain, loss and rearrangement. *PLoS ONE*, *5*(6), e11147. https://doi.org/10.1371/journal.pone.0011147

Eddy, S. R. (1998). Profile hidden Markov models. *BIOINFORMATICS REVIEW*, *14*(9), 755–763.

Eddy, S. R. (2011). Accelerated profile HMM searches. *PLoS Computational Biology*, *7*(10), e1002195. https://doi.org/10.1371/journal.pcbi.1002195

El-Gebali, S., Mistry, J., Bateman, A., Eddy, S. R., Luciani, A., Potter, S. C., … Finn, R. D. (2019). The Pfam protein families database in 2019. *Nucleic Acids Research*, *47*, D427–D432. https://doi.org/10.1093/nar/gky995

Emms, D. M., & Kelly, S. (2015). OrthoFinder: solving fundamental biases in whole genome comparisons dramatically improves orthogroup inference accuracy. *Genome Biology*, *16*(157). https://doi.org/10.1186/s13059-015-0721-2

Fang, H., Feng, J., Xu, Y., Zhang, Y., Zhan, Z., Yin, Z., … Zhou, D. (2018). Sequencing of pT5282-CTXM, p13190-KPC and p30860-NR, and comparative genomics analysis of IncX8 plasmids. *International Journal of Antimicrobial Agents*, *52*, 210–217. https://doi.org/10.1016/j.ijantimicag.2018.04.012

Felsenstein, J. (1981). Evolutionary trees from DNA sequences: A maximum likelihood approach. *Journal of Molecular Evolution*, *17*, 368–376. https://doi.org/https://doi.org/10.1007/BF01734359

Felsenstein, J. (1985). Confidence limits on phylogenies: an approach using the Bootstrap. *Evolution*, *39*(4), 783–791. https://doi.org/10.1111/j.1558-5646.1985.tb00420.x

Foster, E. D., & Deardorff, A. (2017). Open Science Framework (OSF). *Journal of the Medical Library Association*, *105*(2), 203–206. https://doi.org/https://dx.doi.org/10.5195%2Fjmla.2017.88

Galata, V., Fehlmann, T., Backes, C., & Keller, A. (2019). PLSDB: A resource of complete bacterial plasmids. *Nucleic Acids Research*, *47*, D195–D202. https://doi.org/10.1093/nar/gky1050

Garrido-Sanz, D., Meier-Kolthoff, J. P., Göker, M., Martín, M., Rivilla, R., & Redondo-Nieto, M. (2016). Genomic and genetic diversity within the *Pseudomonas fluorescens* complex. *PLoS ONE*, *11*(2). https://doi.org/10.1371/journal.pone.0150183

Gomila, M., Peña, A., Mulet, M., Lalucat, J., & García-Valdés, E. (2015). Phylogenomics and systematics in *Pseudomonas*. *Frontiers in Microbiology*, *6*(Article 214). https://doi.org/10.3389/fmicb.2015.00214

Henry, L., Wickham, H., & Chang, W. (2020). *ggstance: Horizontal "ggplot2" Components*. Retrieved from https://cran.r-project.org/package=ggstance

Hyatt, D., Chen, G.-L., LoCascio, P. F., Land, M. L., Larimer, F. W., & Hauser, L. J. (2010). Prodigal: prokaryotic gene recognition and translation initiation site identification. *BMC Bioinformatics*, *11*(1), 119. https://doi.org/10.1186/1471-2105-11-119

Jutkina, J., Hansen, L. H., Li, L., Heinaru, E., Vedler, E., Jõesaar, M., & Heinaru, A. (2013). Complete nucleotide sequence of the self-transmissible TOL plasmid pD2RT provides new insight into arrangement of toluene catabolic plasmids. *Plasmid*, *70*, 393–405. https://doi.org/10.1016/j.plasmid.2013.09.003

Kanehisa, M., & Goto, S. (2000). KEGG: Kyoto Encyclopedia of Genes and Genomes. *Nucleic Acids Research*, *28*(1), 27–30. https://doi.org/10.1093/nar/28.1.27

Kanehisa, M., & Sato, Y. (2020). KEGG Mapper for inferring cellular functions from protein sequences. *Protein Science*, *29*, 28–35. https://doi.org/10.1002/pro.3711

Kolbe, D. L., & Eddy, S. R. (2011). Fast filtering for RNA homology search. *Bioinformatics*, *27*(22), 3102–3109. https://doi.org/10.1093/bioinformatics/btr545

Lamont, I. L., & Martin, L. W. (2003). Identification and characterization of novel pyoverdine synthesis genes in *Pseudomonas aeruginosa*. *Microbiology*, (149), 833–842.

https://doi.org/10.1099/mic.0.26085-0

Laslett, D., & Canback, B. (2004). ARAGORN, a program to detect tRNA genes and tmRNA genes in nucleotide sequences. *Nucleic Acids Research*, *32*(1), 11–16. https://doi.org/10.1093/nar/gkh152

Letunic, I., & Bork, P. (2016). Interactive tree of life (iTOL) v3: an online tool for the display and annotation of phylogenetic and other trees. *Nucleic Acids Research*, *44*, W242–W245. https://doi.org/10.1093/nar/gkw290

Li, X., Xie, Y., Liu, M., Tai, C., Sun, J., Deng, Z., & Ou, H. Y. (2018). OriTfinder: A web-based tool for the identification of origin of transfers in DNA sequences of bacterial mobile genetic elements. *Nucleic Acids Research*, *46*, W229–W234. https://doi.org/10.1093/nar/gky352

Lick, S., Kröckel, L., Wibberg, D., Winkler, A., Blom, J., Bantleon, A., … Kalinowski, J. (2020). *Pseudomonas carnis* sp. nov., isolated from meat. *International Journal of Systematic and Evolutionary Microbiology*. https://doi.org/10.1099/ijsem.0.003928

Lick, S., Wibberg, D., Winkler, A., Blom, J., Grimmler, C., Goesmann, A., … Kröckel, L. (2021). *Pseudomonas paracarnis* sp. nov., isolated from refrigerated beef. *International Journal of Systematic and Evolutionary Microbiology*, *71*, 004652. https://doi.org/10.1099/ijsem.0.004652

Lippolis, R., Rossi, C., De Angelis, M., Minervini, F., Paparella, A., & Chaves-López, C. (2021). Adaptive remodelling of blue pigmenting *Pseudomonas fluorescens* pf59 proteome in response to different environmental conditions. *Food Control*, *127*, 108105. https://doi.org/10.1016/j.foodcont.2021.108105

Liu, M., Li, X., Xie, Y., Bi, D., Sun, J., Li, J., … Ou, H. Y. (2019). ICEberg 2.0: An updated database of bacterial integrative and conjugative elements. *Nucleic Acids Research*, *47*,

D660–D665. https://doi.org/10.1093/nar/gky1123

Löytynoja, A., & Goldman, N. (2010). WebPRANK: A phylogeny-aware multiple sequence aligner with interactive alignment browser. *BMC Bioinformatics*, *11*, 579. https://doi.org/10.1186/1471-2105-11-579

Marchler-Bauer, A., & Bryant, S. H. (2004). CD-Search: Protein domain annotations on the fly. *Nucleic Acids Research*, *32*, W327–W331. https://doi.org/10.1093/nar/gkh454

Martin, N. H., Murphy, S. C., Ralyea, R. D., Wiedmann, M., & Boor, K. J. (2011). When cheese gets the blues: *Pseudomonas fluorescens* as the causative agent of cheese spoilage. *Journal of Dairy Science*, *94*, 3176–3183. https://doi.org/10.3168/jds.2011-4312

Meier-Kolthoff, J. P., Auch, A. F., Klenk, H. P., & Göker, M. (2013). Genome sequence-based species delimitation with confidence intervals and improved distance functions. *BMC Bioinformatics*, *14*(60). https://doi.org/10.1186/1471-2105-14-60

Meier-Kolthoff, J. P., & Göker, M. (2019). TYGS is an automated high-throughput platform for state-of-the-art genome-based taxonomy. *Nature Communications*, *10*, 2182. https://doi.org/10.1038/s41467-019-10210-3

Meyer, J. M., & Hornsperger, J. M. (1978). Role of pyoverdine(pf), the iron-binding fluorescent pigment of *Pseudomonas fluorescens*, in iron transport. *Journal of General Microbiology*, *107*, 329–331. https://doi.org/10.1099/00221287-107-2-329

Millan, A. S., Toll-Riera, M., Qi, Q., & MacLean, R. C. (2015). Interactions between horizontally acquired genes create a fitness cost in *Pseudomonas aeruginosa*. *Nature Communications*, *6*, 4–11. https://doi.org/10.1038/ncomms7845

Moriya, Y., Itoh, M., Okuda, S., Yoshizawa, A. C., & Kanehisa, M. (2007). KAAS: An automatic genome annotation and pathway reconstruction server. *Nucleic Acids Research*,

35, W182–W185. https://doi.org/10.1093/nar/gkm321

Mulet, M., Lalucat, J., & García-Valdés, E. (2010). DNA sequence-based analysis of the *Pseudomonas* species. *Environmental Microbiology*, *12*(6), 1513–1530. https://doi.org/10.1111/j.1462-2920.2010.02181.x

Nishida, H. (2012). Comparative Analyses of Base Compositions, DNA Sizes, and Dinucleotide Frequency Profiles in Archaeal and Bacterial Chromosomes and Plasmids. *International Journal of Evolutionary Biology*, *2012*, 1–5. https://doi.org/10.1155/2012/342482

Ondov, B. D., Treangen, T. J., Melsted, P., Mallonee, A. B., Bergman, N. H., Koren, S., & Phillippy, A. M. (2016). Mash: Fast genome and metagenome distance estimation using MinHash. *Genome Biology*, *17*(132). https://doi.org/10.1186/s13059-016-0997-x

Piepenbrink, K. H. (2019). *DNA uptake by Type IV filaments*. *6*(February), Article 1. https://doi.org/10.3389/fmolb.2019.00001

Quintieri, L., Fanelli, F., Zühlke, D., Caputo, L., Logrieco, A. F., Albrecht, D., & Riedel, K. (2020). Biofilm and pathogenesis-related proteins in the foodborne *P. fluorescens* ITEM 17298 with distinctive phenotypes during cold storage. *Frontiers in Microbiology*, *11*, Article 991. https://doi.org/10.3389/fmicb.2020.00991

Reichler, S. J., Martin, N. H., Evanowski, R. L., Kovac, J., Wiedmann, M., & Orsi, R. H. (2019). A century of gray: A genomic locus found in 2 distinct *Pseudomonas* spp. is associated with historical and contemporary color defects in dairy products worldwide. *Journal of Dairy Science*, *102*, 5979–6000. https://doi.org/10.3168/jds.2018-16192

Richter, M., Rosselló-Móra, R., Oliver Glöckner, F., & Peplies, J. (2016). JSpeciesWS: A web server for prokaryotic species circumscription based on pairwise genome comparison. *Bioinformatics*, *32*(6), 929–931. https://doi.org/10.1093/bioinformatics/btv681

Ringel, M. T., & Brüser, T. (2018). The biosynthesis of pyoverdines. *Microbial Cell*, *5*(10), 424–437. https://doi.org/10.15698/mic2018.10.649

Rodrigues, R. da S., Machado, S. G., de Carvalho, A. F., & Nero, L. A. (2021). *Pseudomonas* sp. as the causative agent of anomalous blue discoloration in Brazilian fresh soft cheese (Minas Frescal). *International Dairy Journal*, *117*, 105020. https://doi.org/10.1016/j.idairyj.2021.105020

Rossi, C., Serio, A., Chaves-López, C., Anniballi, F., Auricchio, B., Goffredo, E., … Paparella, A. (2018). Biofilm formation, pigment production and motility in *Pseudomonas* spp. isolated from the dairy industry. *Food Control*, *86*, 241–248. https://doi.org/10.1016/j.foodcont.2017.11.018

Roux, S., Enault, F., Hurwitz, B. L., & Sullivan, M. B. (2015). VirSorter: Mining viral signal from microbial genomic data. *PeerJ*, *3*, e985. https://doi.org/10.7717/peerj.985

Sagot, B., Gaysinski, M., Mehiri, M., Guigonis, J. M., Le Rudulier, D., & Alloing, G. (2010). Osmotically induced synthesis of the dipeptide N-acetylglutaminylglutamine amide is mediated by a new pathway conserved among bacteria. *Proceedings of the National Academy of Sciences of the United States of America*, *107*(28), 12652–12657. https://doi.org/10.1073/pnas.1003063107

Schöner, T. A., Gassel, S., Osawa, A., Tobias, N. J., Okuno, Y., Sakakibara, Y., … Bode, H. B. (2016). Aryl polyenes, a highly abundant class of bacterial natural products, are functionally related to Antioxidative carotenoids. *ChemBioChem*, *17*, 247–253. https://doi.org/10.1002/cbic.201500474

Sechi, P., Vizzani, A., Scuota, S., Zicavo, A., Parmegiani, S., & Cenci Goga, B. (2011). Anomalous blue colouring of mozzarella cheese intentionally contaminated with pigment producing strains of *Pseudomonas fluorescens*. *Italian Journal of Food Safety*, *1*(1), 81–

84. https://doi.org/10.4081/ijfs.2011.1.81

Seemann, T. (2014). Prokka: Rapid prokaryotic genome annotation. *Bioinformatics*, *30*(14), 2068–2069. https://doi.org/10.1093/bioinformatics/btu153

Skaar, E. P., Tobiason, D. M., Quick, J., Juddt, R. C., Weissbach, H., Etienne, F., … Seifert, H. S. (2002). The outer membrane localization of the *Neisseria gonorrhoeae* MsrA/B is involved in survival against reactive oxygen species. *Proceedings of the National Academy of Sciences of the United States of America*, *99*(15), 10108–10113. https://doi.org/10.1073/pnas.152334799

Stohl, E. A., Criss, A. K., & Seifert, H. S. (2005). The transcriptome response of *Neisseria gonorrhoeae* to hydrogen peroxide reveals genes with previously uncharacterized roles in oxidative damage protection. *Molecular Microbiology*, *58*(2), 520–532. https://doi.org/10.1111/j.1365-2958.2005.04839.x

Sullivan, M. J., Petty, N. K., & Beatson, S. A. (2011). Easyfig: A genome comparison visualizer. *Bioinformatics*, *27*(7), 1009–1010. https://doi.org/10.1093/bioinformatics/btr039

Tamura, K., Stecher, G., & Kumar, S. (2021). MEGA11: Molecular Evolutionary Genetics Analysis Version 11. *Molecular Biology and Evolution*, *38*(7), 3022–3027. https://doi.org/10.1093/molbev/msab120

Von Neubeck, M., Huptas, C., Glück, C., Krewinkel, M., Stoeckel, M., Stressler, T., … Wenning, M. (2017). *Pseudomonas lactis* sp. nov. and *Pseudomonas paralactis* sp. nov., isolated from bovine raw milk. *International Journal of Systematic and Evolutionary Microbiology*. https://doi.org/10.1099/ijsem.0.001836

Wickham, H. (2016). *ggplot2: Elegant Graphics for Data Analysis*. Retrieved from https://ggplot2.tidyverse.org

Ye, J., Fang, L., Zheng, H., Zhang, Y., Chen, J., Zhang, Z., … Wang, J. (2006). WEGO: A web tool for plotting GO annotations. *Nucleic Acids Research*, *34*, W293–W297. https://doi.org/10.1093/nar/gkl031

Ye, J., Zhang, Y., Cui, H., Liu, J., Wu, Y., Cheng, Y., … Shi, C. (2018). WEGO 2.0: A web tool for analyzing and plotting GO annotations, 2018 update. *Nucleic Acids Research*, *46*, W71–W75. https://doi.org/10.1093/nar/gky400

Supplementary Material

Supplementary Table S1. Summary of secondary metabolites of *Pseudomonas paracarnis* A006 and *Pseudomonas carnis* B157 predicted by ANTIbiotics & Secondary Metabolite Analysis Shell (antiSMASH).

Strain	Type	Size (nt)	Most similar known cluster (similarity %)
A006	RiPP-like	10,879	–
	Arylpolyene	43,576	APE Vf biosynthetic gene cluster (45%)
	Hserlactone	20,579	–
	RiPP-like	10,846	–
	Siderophore	11,926	–
	NRPS	52,897	Pyoverdin biosynthetic gene cluster (11%)
	NRPS	43,712	Pseudomonine biosynthetic gene cluster (100%)
	NRPS	64,525	Pyoverdin biosynthetic gene cluster (10%)
	Betalactone	23,173	Fengycin biosynthetic gene cluster (13%)
	NAGGN	14,842	–
	Redox-cofator	19,821	Lankacidin C biosynthetic gene cluster (13%)
	NRPS-like	17,595	Fragin biosynthetic gene cluster (25%)
B157	Hserlactone	20,576	–
	NAGGN	14,864	–
	NRPS-like, betalactone	33,997	Pyoverdin biosynthetic gene cluster (2%)
	Siderophore	11,926	–
	Arylpolyene	43,576	APE Vf biosynthetic gene cluster (45%)
	RiPP-like	10,879	–
	NRPS	65,893	Pyoverdin biosynthetic gene cluster (11%)
	Betalactone	23,175	Fengycin biosynthetic gene cluster (13%)
	NRPS	64,236	Viscosin biosynthetic gene cluster (43%)
	RiPP-like	10,846	–
	Redox-cofator	22,148	Lankacidin C biosynthetic gene cluster (13%)
	NRPS	46,270	Anikasin biosynthetic gene cluster (55%)
	Terpene	22,226	–
	NRPS	33,392	Pyoverdin biosynthetic gene cluster (3%)
	NRPS-like	19,206	Fragin biosynthetic gene cluster (25%)

Supplementary Table S2. Phage and prophages regions detected in draft genome of *Pseudomonas paracarnis* A006 and *Pseudomonas carnis* B157 using Virsorter v.1.0.6 and only categories 1 and 2 were considered. The output was filtered using the PHAge Search Tool (PHASTER).

	Virsorter						PHASTER								
Genome	**Prediction**	**Length (Kb)**	**Category**	**Fragment**	**N genes**	**N phage hallmark genes**	**Completeness[a]**	**Score**	**Specific Keyword[b]**	**N ORF**	**Phage Hit Proteins**	**Attachment Sited[c]**	**GC %**	**Length (Kb)**	**Most Common Phage**
A006	Complete phage contig	89.639	2	All contig	98	4	intact	150	transposase, tail, head, capsid, terminase, lysis, integrase	57	41	no	57.11%	41.9	PHAGE_Entero_HK022_NC_002166(8)
A006	Prophage	15.169	2	[gene 85 - gene 106] [115588 - 130757]	22	5	intact	130	tail, plate	23	19	no	59.47%	14.8	PHAGE_Vibrio_VP882_NC_009016(14)
A006	Prophage	117.643	2	[gene 89 - gene 270] [83183 - 200826]	182	8	two intact regions	150	tail, terminase, plate, capsid, head, portal, lysin, integrase	78	53	yes	57.03%	47.4	PHAGE_Iodoba_phiPLPE_NC_011142(14)
								150	virion, tail, head, capsid, portal, integrase	66	44	yes	58.22%	50	PHAGE_Entero_HK022_NC_002166(11)
B157	Prophage	62.564	2	[gene 326 - gene 411] [350283-412847]	86	8	intact	140	tail, plate	21	18	no	59.28%	14.2	PHAGE_Vibrio_VP882_NC_009016(14)
							questionable	90	virion, capsid, portal, terminase, lysis	42	28	no	58.32%	26.4	PHAGE_Pseudo_F10_NC_007805(6)
B157	Prophage	83.804	2	[gene 285 - gene 385] [284577-368381]	101	4	intact	130	capsid, lysin, lysis, terminase, head, virion, tail, integrase	48	36	yes	58.75%	39.8	PHAGE_Pseudo_AF_NC_019923(14)

[a] Intact (score > 90); Questionable (score 70-90); Incomplete (score < 70)

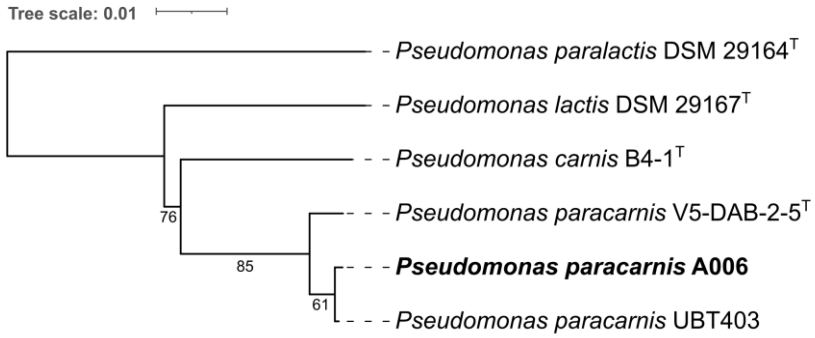
[b] The specific phage-related keyword(s) found in protein name(s) in the region

ORF: Open Reading Frame

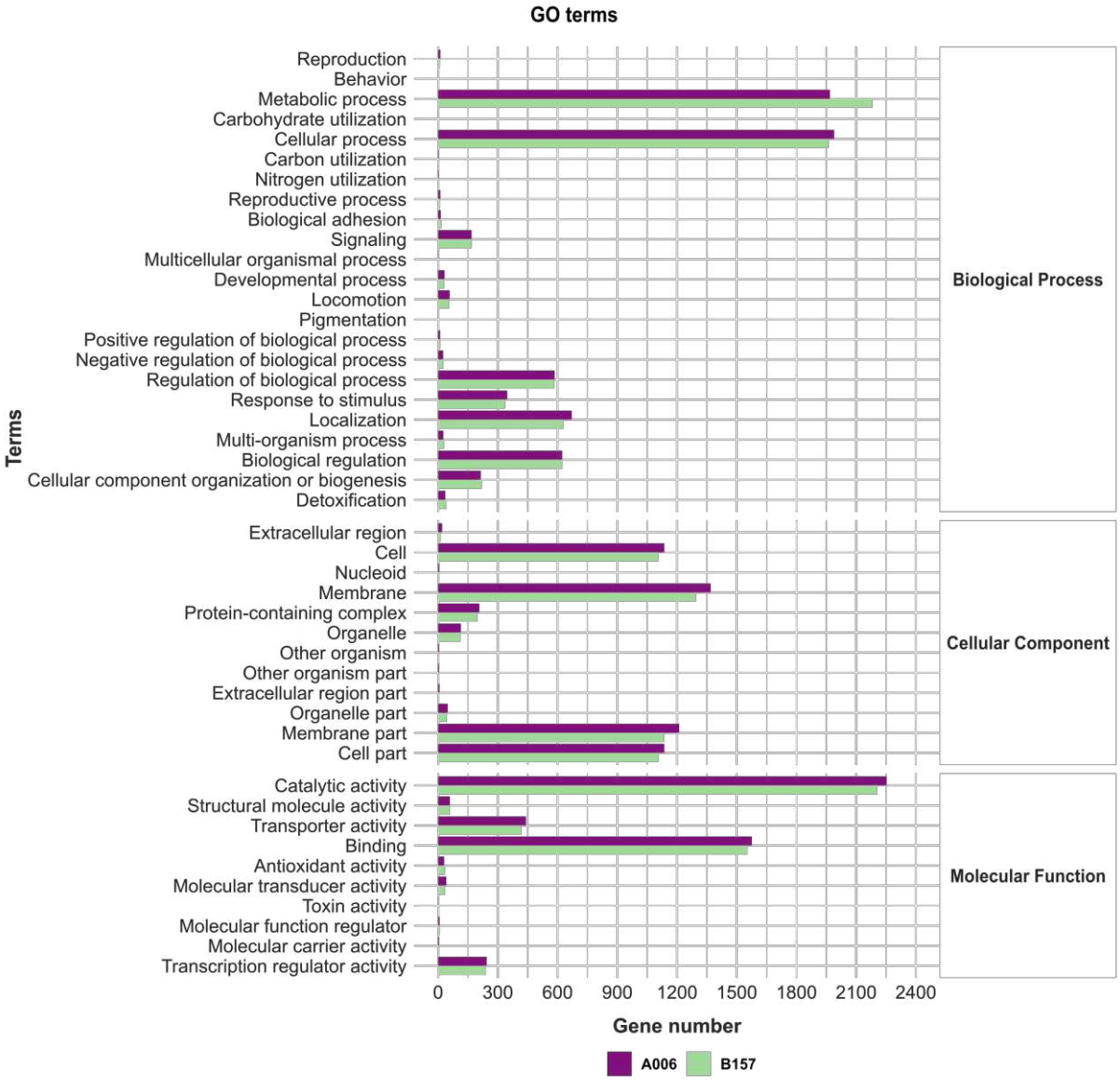
[c] The putative phage attachment site

Supplementary Table S3. Position in draft genome, length (bp), and GC content (%) of unique genomic regions identified in blue discoloration-causing strains.

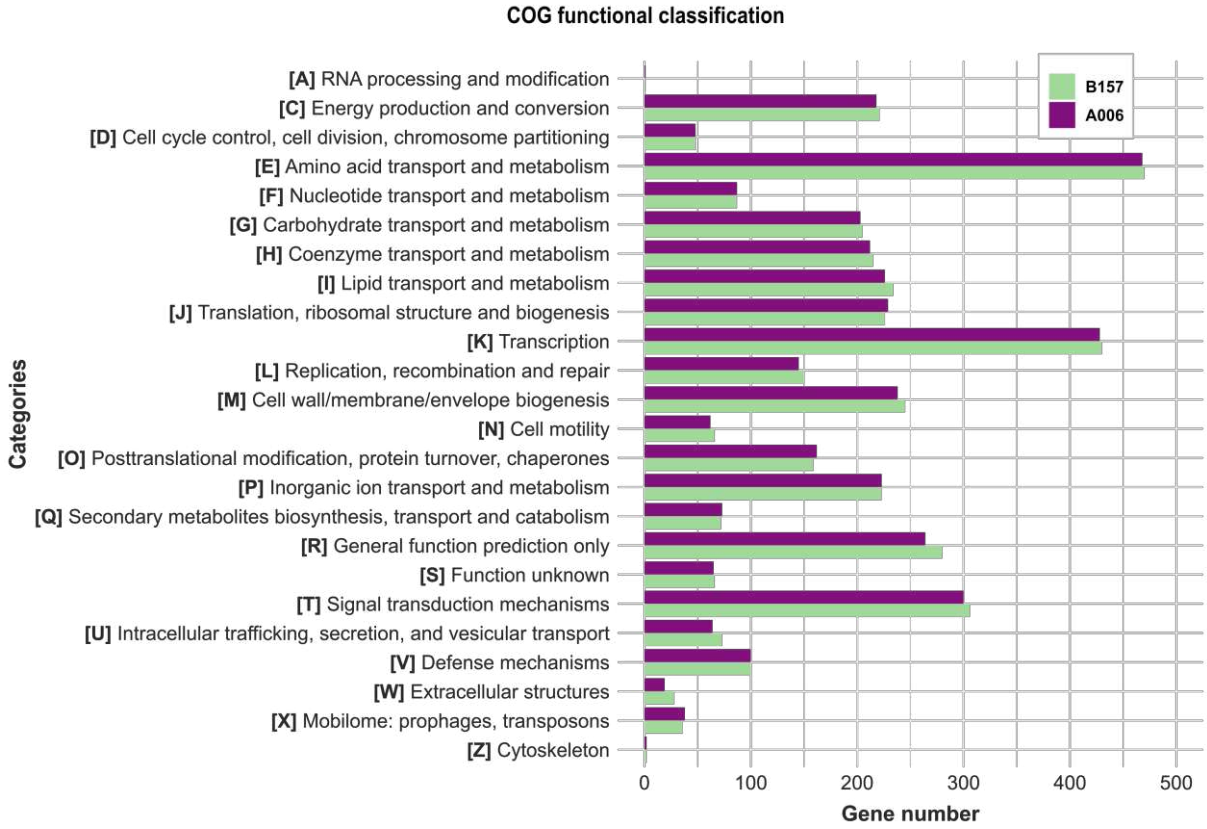
Strains	Exclusive region position in genome (Contig: start - end)	Length (bp)	GC content (%)	
			Exclusive region	Genome
B4-1	CABIVL010000025.1: 1277 - 67360	66,083	55.68	59.78
ITEM 17298	NPKB01000028.1: 14062 - 80159	66,097	55.67	59.75
Ps_77	gi_984325036_gb_LCYB01000004.1_1907245: 628383 - 694463	66,080	55.68	59.66
E2-0548	RXHZ01000014.1: 27432 - 93514	66,082	55.68	59.66
E2-8864	RWKC01000002.1: 225259 - 291358	66,099	55.67	59.75



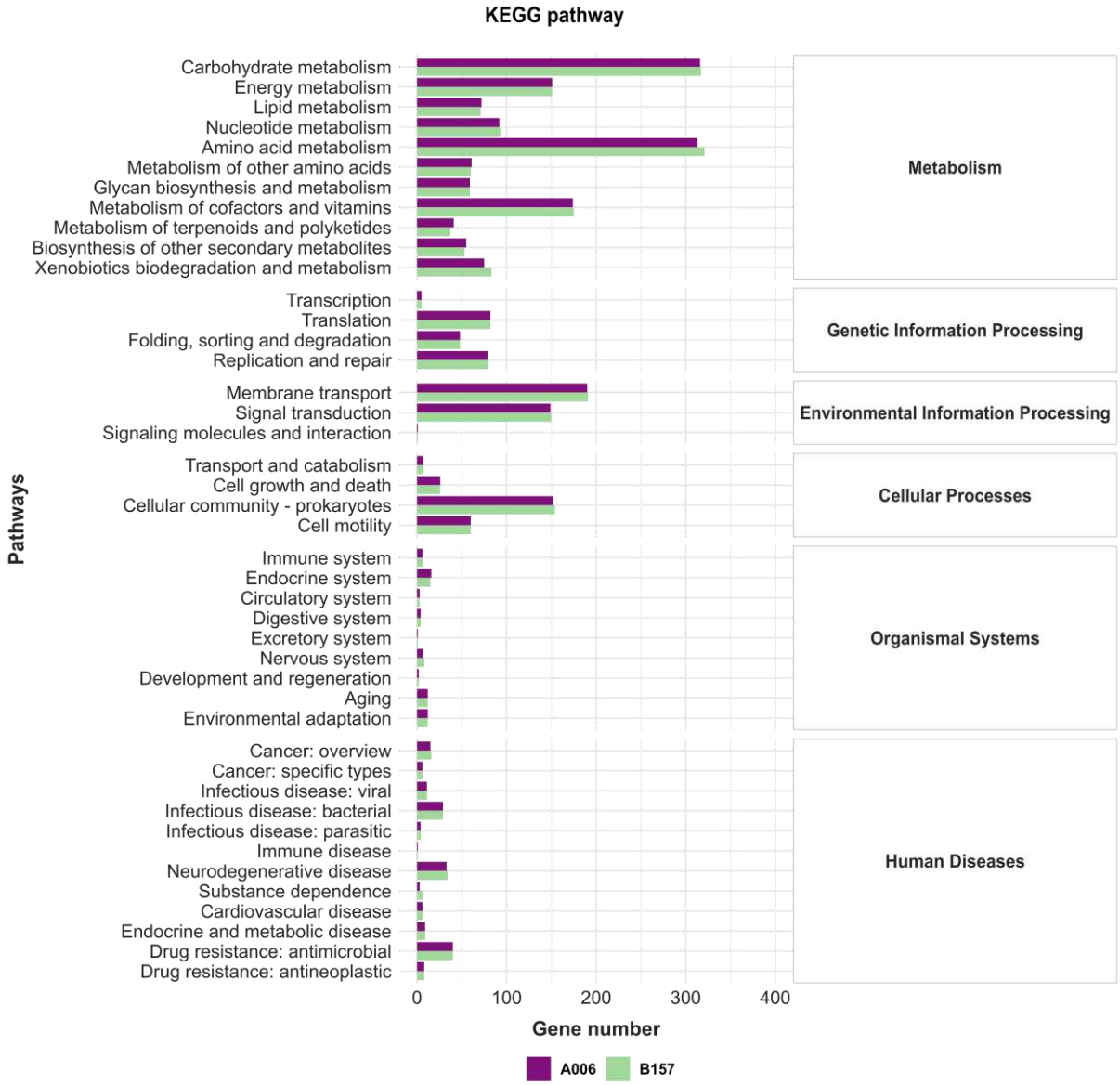
Supplementary Fig. S1. A phylogenomic tree constructed using the Type Strain Genome Server (TYGS; Meier-Kolthoff and Göker 2019). Tree inferred with FastME 2.1.6.1 (Lefort, Desper, & Gascuel, 2015) from Genome Blast Distance Phylogeny (GBDP) calculated from genome sequences. The branch lengths are scaled in terms of GBDP distance formula d_5. The numbers above branches are GBDP pseudo-bootstrap support values >60% from 100 replications, with an average branch support of 74.0 %. The tree was rooted at the midpoint (Farris, 1972).



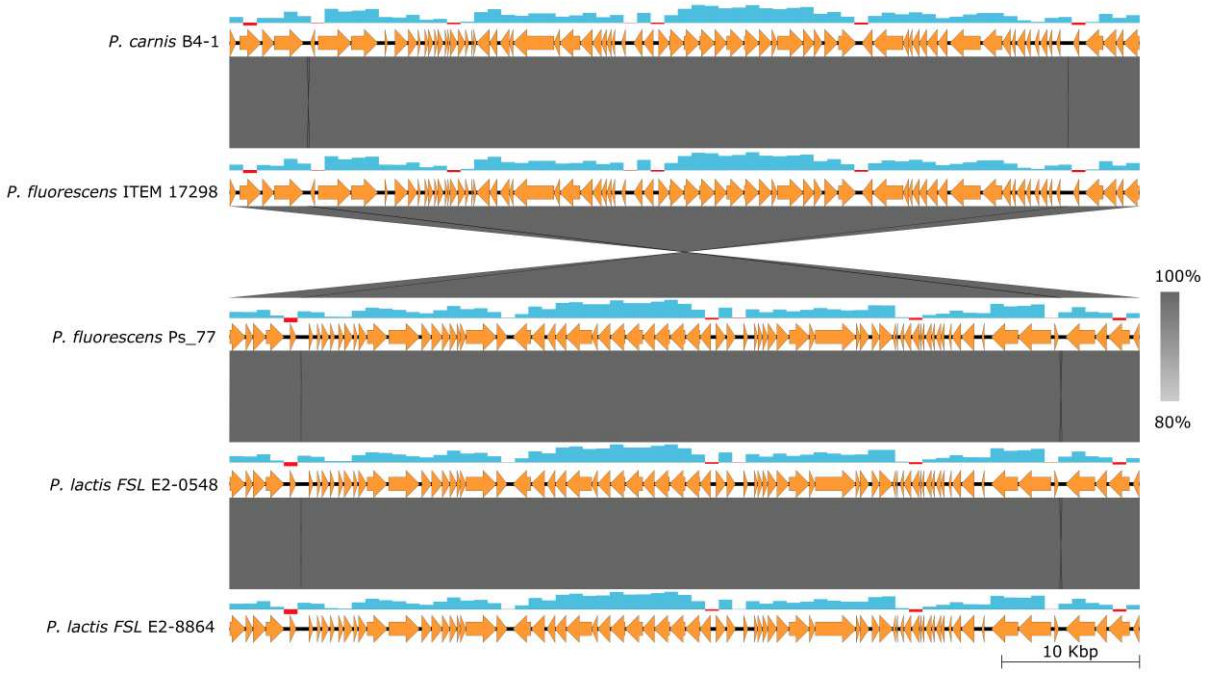
Supplementary Fig. S2. Coding-sequences analyses of draft genome of *Pseudomonas paracarnis* A006 and *Pseudomonas carnis* B157, through the Gene Ontology (GO) functional with classification in three main GO categories (biological process, cellular component, and molecular functional) and 45 subcategories. The right y-axis represents the categories, the left represents the subcategories, and the x-axis represents the number of genes.



Supplementary Fig. S3. Coding-sequences analyses of draft genome of *Pseudomonas paracarnis* A006 and *Pseudomonas carnis* B157, in Clusters of Orthologous Genes (COG), summarized in 24 categories. The y-axis represents the categories and the x-axis represents the number of genes in each category.



Supplementary Fig. S4. Coding-sequences analyses of draft genome of *Pseudomonas paracarnis* A006 and *Pseudomonas carnis* B157, through the Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway with classification in six categories (metabolism, genetic information processing, environmental information processing, cellular process, organismal systems, and human diseases) and 43 subcategories. The right y-axis represents the categories, the left represents the subcategories, and the x-axis represents the number of genes.



Supplementary Fig. S5. Schematic representation of alignment between unique genomic regions identified in draft genomes from blue discoloration-causing strains. The direction of the arrows represents the coding strand of the coding sequence. A BLASTn comparison was performed using Easyfig v.2.2.6 software, and the levels of similarity are shown in the gray gradient scale. The histogram graph above the arrows represents positive and negative values compared to the GC content of the total region, represented in blue and red, respectively.

Farris, J. S. (1972). Estimating Phylogenetic Trees from Distance Matrices. *The American Naturalist*, *106*(951), 645–668. https://doi.org/10.1086/282802

Lefort, V., Desper, R., & Gascuel, O. (2015). FastME 2.0: A comprehensive, accurate, and fast distance-based phylogeny inference program. *Molecular Biology and Evolution*, *32*(10), 2798–2800. https://doi.org/10.1093/molbev/msv150

Meier-Kolthoff, J. P., & Göker, M. (2019). TYGS is an automated high-throughput platform for state-of-the-art genome-based taxonomy. *Nature Communications*, *10*, 2182. https://doi.org/10.1038/s41467-019-10210-3

CHAPTER III: Genome-based analysis of *Pseudomonas paracarnis* RQ057, a strain responsible for blue discoloration spoilage in processed cheese

Rafaela da Silva Rodrigues et al.

Manuscript prepared for submission to *International Journal of Dairy Technology*

Title page

Rafaela da Silva Rodrigues[1], Solimar Gonçalves Machado[1], Antônio Fernandes de Carvalho[1], Luís Augusto Nero[2]*

Genome-based analysis of *Pseudomonas paracarnis* RQ057, a strain responsible for blue discoloration spoilage in processed cheese

[1] Department of Food Technology, Federal University of Viçosa, Av. P. H. Rolfs s/n, 36570900 Viçosa, MG, Brazil

[2] Department of Veterinary Medicine, Federal University of Viçosa, Av. P. H. Rolfs s/n, 36570900 Viçosa, MG, Brazil

* Corresponding author: LA Nero, nero@ufv.br, tel. +55 31 3612 5615.

Abstract

Some *Pseudomonas* species are responsible for causing diverse types of spoilage in dairy products, such as the blue discoloration. Here, we isolated and identified a strain of *Pseudomonas* from processed cheese "requeijão em barra" as the cause of blue discoloration and studied its genome. By whole-genome sequencing, we obtained a draft genome assembled in 52 contigs accounting for 6,185,699 bp, with 59.7% G+C content and harbors 5,566 predicted protein-coding genes. The strain was identified as *Pseudomonas paracarnis* (named as RQ057) using the TYGS web server, the Genome-to-Genome Distance Calculator to estimate the *in-silico* DNA-DNA hybridization, and the average nucleotide identity values based on BLAST, ANIb. Through genomic analysis, two pyoverdine biosynthetic clusters, integrative and conjugative elements (ICE), prophage regions, and genomic islands (GIs) were predicted. Furthermore, we observed that copies of genes related to the tryptophan biosynthetic pathway (*trp*CBAFD), which are unique to blue pigments, are contained in the region identified as ICE and which was also identified in GIs. This genomic region also contains genes for: transposase, type IV transport system (T4SS), and pili construction. Thereby, we can assume that this region contains potential conjugative-transposon-like mobile genetic elements and it has plasmid-like conjugative properties. Overall, the sequencing and characterization this genome to provide a promising resource to conduct molecular investigations of this spoilage pigmenting *Pseudomonas*, especially to select genetic markers to identify and to track potentially spoilage strains.

Keywords: Genomic analysis, *Pseudomonas*, blue discoloration, processed cheese

1. Introduction

Pseudomonas is a bacterial genus belonging to the class *Gammaproteobacteria* of phylum *Proteobacteria*, nonsporulating aerobic cells with rod-shaped morphology, and catalase positive (Palleroni, 2005). It is the genus of Gram-negative bacteria with the highest number of valid species, which currently contains more than 309 names listed by the List of Prokaryotic names with Standing in Nomenclature, LPSN (https://lpsn.dsmz.de/genus/pseudomonas). Many of its species have received special attention in the food industry for being psychotropic and forming strong biofilms, which ensures growth and persistence in the food processing environment (Rossi et al., 2018). Moreover, the presence of this bacterium in the industry causes numerous losses, because many of them are capable of producing proteolytic enzymes that destabilize milk products during storage (Glück et al., 2016; Machado et al., 2017) and pigment molecules (Carminati et al., 2019).

Greater concern about the pigments produced by *Pseudomonas* has been observed in the dairy industry. Cheeses with high moisture content, even if kept at cold storage, are still prone to the growth of some members of this genus and these food products have been most related to the production of pigments (Carrascosa et al., 2021, 2015; Cenci-Goga et al., 2014). In recent years, *Pseudomonas* species has been linked to the cheese spoilage by causing blue discoloration, with particular relevance to *P. lactis*, *P. fluorescens*, *P. carnis* and *P. paracarnis* (Andreani et al., 2014; Reichler et al., 2019; Rodrigues et al., 2021). However, this spoilage is not species-dependent, as only some strains are blue-pigment producers and, thereby, the use of genome-based analysis can help to understand what these strains have so distinct.

It is well known that the whole-genome sequencing and the exploration of *Pseudomonas* genomes allow the identification of key attributes related to: understanding of spoilage-

associated behaviors and biofilm formation (Quintieri et al., 2020), search for gene clusters related to proteolytic activity (Maier et al., 2020), identification of mobile genetic elements (Loper et al., 2012), and identification and phylogenomic distribution (Garrido-Sanz et al., 2016), like several other possibilities. Besides, the genomic-based analyses of blue pigment *Pseudomonas* have shown that the genome of these strains contains some duplicated tryptophan biosynthesis pathway genes (Reichler et al., 2019). Thus, we aimed to isolate and identify *Pseudomonas* from processed cheese ("requeijão e barra") with blue discoloration and to use whole-genome sequencing to characterize this pigment producer.

2. Material and methods

2.1 Isolation history and bacterial strain

In 2020 September, a dairy industry located in the Zona da Mata region, Minas Gerais state, Brazil, reported a blue discoloration spoilage in "requeijão em barra", a processed cheese widely produced and consumed in Brazil. By the time of the report, three samples that presented blue discoloration were obtained, ten-fold diluted in peptone NaCl solution (0.1% peptone, w/v; 0.85% NaCl, w/v) on and plated on *Pseudomonas* agar base (PAB; Oxoid Ltd., Basingstoke, UK) supplemented with cephaloridine, fucidin and cetrimide (CFC, Oxoid), and bromothymol blue (0.02 g L^{-1}), and then incubated at 25±1 °C for 48h. Isolates were tested for catalase, oxidase, and Gram staining. The presumptive *Pseudomonas* isolates were stored at -80 °C in Brain Heart Infusion (BHI – Kasvi, São José dos Pinhais, PR, Brazil) with 20% v/v glycerol in the bacteria culture collection at InovaLeite (Laboratory of Milk and Dairy Products, Federal University of Viçosa, Viçosa, Minas Gerais, Brazil).

2..1.1 Pigment producing

The presumptive *Pseudomonas* isolates were tested to pigment producing in *Pseudomonas* minimal medium (PMM), consisting of 35 mM $K_2HPO_4 \cdot 3H_2O$, 22 mM KH_2PO_4, 8 mM $(NH_4)_2SO_4$, 1,2 mM $MgSO_4$ (Kirner et al., 1996), and 30 mM glucose as a carbon source (Chierici et al., 2016). The prepared PMM was filtered through 0.22 μm pore size filter (Kasvi, São José dos Pinhais, PR, Brazil). The cultures, after incubation in BHI broth at 25 °C for 24 h, were harvested by centrifugation (4,500 RPM for 15 min at 4 °C) and washed twice with saline solution 0.85 % (w/v). Then, they were inoculated into PMM at final concentration of approximately 10^5 CFU/mL in 24-well cell culture plates and incubated at 10, 25 and 32 °C for 10 days to evaluate the production of pigment.

2.1.2 Molecular identification

The presumptive *Pseudomonas* isolates were cultured in BHI broth at 25 °C *overnight*, and the obtained cultures were subjected to DNA extraction using the Wizard® Genomic DNA purification kit (Promega Corp., Madison, WI, USA). DNA was analyzed by electrophoresis on a 0.8% (w/v) agarose gel. Also, DNA concentration and quality were measured by the NanoDrop™ Lite Spectrophotometer (Thermo Scientific, Madison WI, EUA).

Pseudomonas spp. were initially identified using primers PA-GS-F (5'-GACGGGTGAGTAATGCCTA-3') and PA-GS-R (5'-CACTGGTGTTCCTTCCTATA-3') by PCR amplification of the genus-specific region of 16S rRNA (Spilker et al., 2004). PCR reactions consisted of Go Taq Master Mix 2x (Promega), 10 pMol of each pair of primers, 100 ng of DNA, and ultra-pure PCR water (Promega). The PCR amplification was performed with a thermocycler (MaxyGene® Gradient Thermal Cycler; Axygen Scientific, USA) using the following conditions: initial denaturation step at 95 °C for 2 min; 25 cycles at 94 °C for 20 s,

54 °C for 20s, 72 °C for 40s; and with a final extension at 72 °C for 1 min. The PCR products were electrophoresed on agarose gels (1,5 % w/v) in 0.5 × tris/borate/EDTA buffer (TBE) at constant voltage (75 V) for 1 h, and stained with Diamond™ Nucleic Acid Dye (Promega). The gels were visualized under UV light using 100 pb DNA ladder (Invitrogen^{TM}) as a molecular size marker. Amplicons of ~618 bp were considered indicative of *Pseudomonas* spp.

2.1.3 Rep-PCR fingerprinting

To analyze the genetic variability between *Pseudomonas* isolates, repetitive extragenic palindromic PCR (Rep-PCR) assay was employed using a single primer $(GTG)_5$ (5′-GTGGTGGTGGTGGTG-3′), according (Versalovic et al., 1994) with modification. The amplification reaction mixture contained 1X *Taq* reaction Buffer (New England BioLabs), 2 U of *Taq* DNA polymerase (New England BioLabs), 500 µM deoxynucleoside triphosphates (Promega), 50 pmol of primer $(GTG)_5$, 4% of DMSO (v/v), and 100 ng of DNA. The rep-PCR conditions were: initial denaturation step at 94 °C for 5 min; 30 cycles at 94 °C for 60 s, 38 °C for 75s, and 68 °C for 8 min; and with a final extension at 68 °C for 16 min. PCR products were electrophoresed on agarose gels (1.2 % w/v) in 0.5 × TBE at constant voltage (60 V) for 150 min, and stained with Diamond™ Nucleic Acid Dye (Promega). The gels were photographed under UV light using 1 Kb DNA ladder (Invitrogen^{TM}) as a molecular size marker. Fingerprints were analysed using BioNumerics version 6.6 (Applied Maths, Kortrijk, Belgium). Similarity analysis was performed using the Dice coefficient, and a dendrogram was constructed using the Unweighted Pair Group Method with Arithmetic Mean (UPGMA) at 5% optimization and 1.5% tolerance. A cut-off value of 90% similarity was applied to compare the DNA fingerprint patterns.

2.2 Whole-genome sequencing and genome assembly

Based on rep-PCR profiles and similarities, one isolate (RQ057) producing blue pigment was selected for genomic analysis. The DNA extraction, construction of paired-end libraries, and genome sequencing were carried out by Neoprospecta Microbiome Technologies (Florianopolis, SC, Brazil). Illumina Miseq platform was used for whole-genome shotgun sequencing and generated 4,741,240 paired-end reads with a mean length of 305 bp forward and 205 bp reverse. The reads were check using FastQC version 0.11.9 (Andrews, 2010) and trimming using Trimmomatic v.0.39 (Bolger et al., 2014) to accept 20 as lowest Phred value and to remove adaptors. We tested four assemblers for *de novo* genome assembly with the following parameters: Mira Assembler v.4.9.6, mode 'genome, accurate' (Chevreux et al., 1999); SPAdes v.3.15.2, in careful mode (Bankevich et al., 2012); SOAPdenovo v2.04 with k-mer value of 127 (Luo et al., 2012); and, Velvet v.1.2.10 with k-mer value of 71 (Zerbino and Birney, 2008). Genome assembly quality metrics were generated using QUAST v.5.0.2 (Gurevich et al., 2013) for each assembled draft genome. Contigs with a length smaller than 500 bp were excluded. Furthermore, potential contamination of the draft genome sequence was checked by the Contamination Estimator by 16S (ContEst16S) algorithm (Lee et al., 2017).

2.3 Gene prediction and functional annotation

The draft genome was annotated using Prokaryotic Genome Annotation System (Prokka) v.1.14.5 (Seemann, 2014) software with default settings. The output of Prokka was used to further analyses. Encoded proteins were annotated with Clusters of Orthologous Genes (COG) performed in online server Batch CD-search tool (https://www.ncbi.nlm.nih.gov/Structure/bwrpsb/bwrpsb.cgi) (Marchler-Bauer and Bryant, 2004), with default parameters. The files 'cog-20.def.tab' and 'fun-20.tab' from

https://ftp.ncbi.nih.gov/pub/COG/COG2020/data/ were used to assign the results to COG category description. The gene ontology (GO) annotation was predicted using Blast2GO Basic v.5.2.5 (Conesa et al., 2005) based on BLASTp searches against the nonredundant protein database with an e-value cutoff of 1e-10, and using InterProScan. The online tool WEGO 2.0 were used for functionally classified of the terms GO (Ye et al., 2006, 2018). KEGG orthology annotation were performed using the KEGG Automation Annotation Server (KAAS), https://www.genome.jp/kegg/kaas/by the BBH method (Kanehisa and Goto, 2000; Moriya et al., 2007), and KEGG pathways was assign to the results using information extracted from KEGG databases via the KEGG REST API (https://www.kegg.jp/kegg/rest/keggapi.html). To identify putative biosynthetic gene clusters for secondary metabolites, we used antiSMASH (ANTIbiotics & Secondary Metabolite Analysis Shell) v.6.0.1 (Blin et al., 2019), with the detection strictness parameter in default mode and all extra features on. Sequences that were predicted to belong to the pyoverdine biosynthetic cluster were aligned with the pyoverdin sequences from *Pseudomonas fluorescens* A506 (CP003041.1), using BLASTp (e-value ≤ 1e-10; identity ≥ 70%; query coverage ≥ 80%).

2.4 Phylogenomic analysis

The draft genome sequence was uploaded to the Type Genome Server (TYGS) for a whole genome-based taxonomic analysis (Meier-Kolthoff and Göker, 2019). Then, the best 14 matching type strains were used to calculate *in silico* DNA-DNA hybridization (dDDH) with Genome-to-Genome Distance Calculator v.3.0 (GGDC) web tool, using BLAST+ and the recommended formula 2 (Meier-Kolthoff et al., 2013), and average nucleotide identity (ANI) based in BLAST (ANIb) using JSpeciesWS online service (Richter et al., 2016).

2.5 Mobile genetic elements search

2.5.1 Prediction of integrative and conjugative elements

The ICEFinder web-based tool was used for identification of integrative and conjugative element (ICE) and/or integrative and mobilizable elements (IMEs) (Liu et al., 2019), with upload the nucleotide sequence file in the Fasta format and using the default parameters. The sequence identified was extracted from genome Fasta file using BEDTools v.2.26.0 (Quinlan and Hall, 2010) and predicted proteins were extracted from protein Fasta file using SeqKit v.0.14.0 (Shen et al., 2016). The proteins were annotated by BLASTp against NCBI non-redundant protein database (e-value ≤ 1e-10; identity ≥ 60%; query coverage ≥ 70%). For classification, we also annotated the proteins using family and domain databases: Pfam (Mistry et al., 2021), with Hmmer and e-value ≤ 1e-5 (http://pfam.xfam.org/); InterPro using InterProScan (http://www.ebi.ac.uk/interpro/); and UniProt (Bateman et al., 2021), when necessary, using BLAST and default parameters (https://www.uniprot.org/blast/).

2.5.2 Prediction of prophage regions

The draft genome was also screened for prophage sequences using PHASTER (PHAge Search Tool Enhanced Release) web server (Arndt et al., 2016). The candidate sequences for intact prophage were annotated with multiPhATE v.2.0 command-line program (Ecale Zhou et al., 2019), using Phanotate (Mcnair et al., 2019) to predict open reading frames (ORFs), and tRNAscan for tRNA gene. In this program, PhAnToMe (Phage Annotation Tools and Methods) and pVOGs (Prokaryotic Virus Orthologous Groups) databases were used for automatic functional annotation (BLASTp, e-value ≤ 1e-10, and other default parameters). The classification of phage proteins was performed by VIRFAM web server performing analyses of the head-neck-tail module genes and recombinase genes server (Lopes et al., 2010; Lopes et

al., 2014). Additionally, BLASTp against the NCBI (National Center for Biotechnology Information) non-redundant protein database was used to assign the functions of proteins (e-value ≤ 1e-10; identity ≥ 60%; query coverage ≥ 70%). Lastly, the predicted prophages were assigned taxonomy to level of family using VipTree web server v.1.9 (Nishimura et al., 2017), computed by tBLASTx, and also assessed with VIRFAM.

2.5.3 Prediction of genomic islands

The draft genome was subjected to IslandViewer4 web server (Bertelli et al., 2017) for the prediction of genomic islands (GIs) using *P. fluorescens* A506 as reference genome.

2.5.4 Identification of plasmid sequences

We searched for plasmid in the trimmed and paired-end reads using PlasmidSPAdes pipeline in SPAdes v.3.15.2 (Antipov et al., 2016; Bankevich et al., 2012) with careful mode and the other parameters in default. The draft genome was also screened by PlasmidFinder (Carattoli et al., 2014) on September 8, 2021.

2.6 Data processing and visualization

A circular representation to the *P. paracarnis* RQ057 draft genome was constructed using the Circos software v.0.69 (Krzywinski et al., 2009). For that, GC content and GC skew were calculated using a python script GCcalc.py (https://github.com/WenchaoLin/GCcalc) with a 4,000 bp window size and a step size of 2,000 bp. The phylogenomic tree obtained from the whole genome-based taxonomic analysis in TYGS was visualized using the online tool iTOL v.4.2 (Letunic and Bork, 2016). The R software v.4.0.2 (R Core Team, 2020) and RStudio

v.1.3.959 (RStudio Team, 2020) were used to data organization of COG, KEGG, and GO annotation using 'dplyr' v.1.0.7 (Wickham et al., 2021) and 'plyr' (Wickham, 2011) packages. Other R packages were used for graphical presentation: 'ggplot2' (Wickham, 2016) and 'ggstance' (Henry et al., 2020) to KEGG pathways and GO terms; 'VennDiagram' (Chen, 2018) to generate Venn diagram of annotated coding sequences, after construction by web-based program Venny (https://bioinfogp.cnb.csic.es/tools/venny/) (Oliveros, 2007); and, the 'ggplot2' package were also used to create the heatmaps for the GGDC and ANIb results and to generate the schematic representation of coding sequences in association with 'gggenes' (Wilkins, 2020). Furthermore, a BLASTn comparison was performed using Easyfig v.2.2.6 software (Sullivan et al., 2011) between tryptophan biosynthesis gene clusters.

2.7 Data availability

The raw data of whole-genome sequencing of *P. paracarnis* RQ057 is available in Sequence Read Archive (SRA) database of NCBI with accession number SRR16642493. The draft genome sequence has been deposited at DDBJ/ENA/GenBank under the accession JAJGWQ000000000. The version described in this paper is version JAJGWQ010000000.

3. Results

3.1 Isolation history and bacterial strain

High concentration of *Pseudomonas* spp. in samples of processed cheese was detected (7.69 ± 0.54 log CFU/g). Thus, we obtained 65 presumptive *Pseudomonas* isolates from these samples, showing blue discoloration spoilage. Altogether, 34 isolates (52%) produced blue

pigment in *Pseudomonas* minimal medium (PMM), at 10 °C and 25 °C after 4 and 2 days, respectively, but no blue pigment was observed with incubation at 32 °C for 10 days.

The 65 isolates were identified as belonging to the genus *Pseudomonas* by PCR amplification of the genus-specific region of 16S rRNA, and the rep-PCR fingerprint analysis resulted in 5 different clusters with similarity indices of at least 90% and 6 solitary fingerprints (Supplementary Fig. S1). All blue pigmenting isolates were clustered together, showing an identical pattern with 100% similarity.

3.2 Genomic features

One blue pigmenting isolate was selected for whole-genome sequencing (WGS). The best final assembly obtained from the WGS of *P. paracarnis* RQ057 was using the MIRA assembler (Table 1). The draft genome was assembled in 52 contigs accounting for 6,185,699 bp, an estimated G+C content of 59.7%, and the results of ContEst16S algorithm indicated that it did not have contamination of another prokaryotic genome. The prediction using the Prokka software revealed that the draft genome contains 5,566 protein-coding sequences, 70 tRNAs, 16 rRNAs, and 1 transfer-messenger RNA (tmRNA). Predicted proteins of *P. paracarnis* RQ057 were further annotated using the COG, GO, and KEGG orthology (KO) databases (Table 2), which we annotated 60.0%, 71.7%, and 55.6% of predicted proteins, respectively. Furthermore, 2,246 (40.4%) were annotated by all databases (Supplementary Fig. S2). The assignment of genes into COGs functional categories is illustrated using Circos in Fig. 1. Among the categories with more genes are: amino acid transport and metabolism with 465 genes; transcription with 421; signal transduction mechanisms with 290; general function prediction only with 265; cell wall/membrane/envelope biogenesis with 238; translation,

ribosomal structure and biogenesis with 229; and, lipid transport and metabolism with 226 (Supplementary Fig. S3).

Table 1. Quality metrics of the draft assembled genome of *Pseudomonas paracarnis* RQ057 by different assemblers.

Quality parameters	Assemblers			
	Mira	SPAdes	SOAPdenovo	Velvet
Contigs	52	58	144	76
Total length (bp)	6,185,699	6,145,775	6,134,296	6,133,540
Largest contig (bp)	1,090,179	1,006,335	241,428	1,005,776
GC (%)	59.7	59.73	59.74	59.73
N50	476,617	242,829	93,319	185,998
L50	5	6	22	7
N count*	29	0	0	355

* Number of uncalled bases

Table 2. Summary of the functional annotation of predicted protein-coding genes from the *Pseudomonas paracarnis* RQ057 draft genome.

Database (DB)	**Coding sequences**	**Percent**
GO	3,993	71.7%
KO	3,092	55.6%
COG	3,340	60.0%
At least DB	4,523	81.3%
In all DB	2,246	40.4%
By none	1,043	18.7%
Total	5,566	100.0%

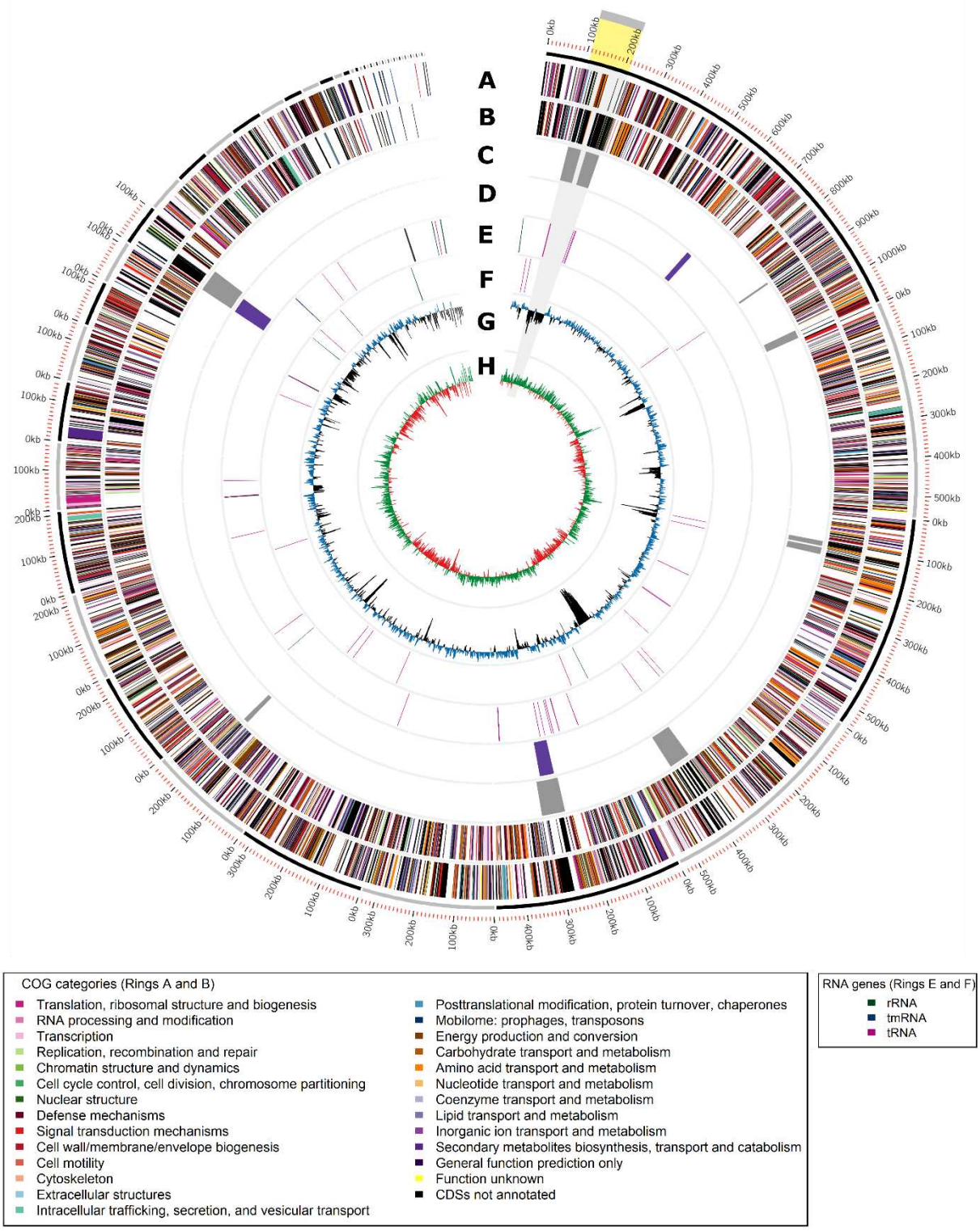


Fig. 1. Circular representation of the *Pseudomonas paracarnis* RQ057 draft genome. All the 52 contigs are shown, and they are ordered by size. Contigs smaller than 100kb had no expressed size. Each coding-sequence (CDS) is color-coded by its COG category. **A** – CDS on the forward strand. **B** – CDS on the reverse strand. **C** – Genomic islands predicted by the IslandViewer, using IslandPath-DIMOB. **D** – Intact prophages predicted by PHASTER webserver. **E** – RNA on the forward strand. **F** – RNA on the reverse strand. **G** – displays GC content, where blue lines indicate regions above or equal to 59.7% (average GC-content of the genome) and black lines indicate regions

below to average GC-content; **H** – displays GC skew where green lines indicate a skew greater than or equal to zero and red lines indicate a skew less zero. The highlight in gray (with a yellow edge) shows two genomic island that surround the exclusive region of blue discoloration strains.

Through annotation using GO database, coding sequences were assigned to GO terms for three categories: biological process, cellular component, and molecular function (Fig. 2, A). The dominant terms in biological process category are cellular process (1,964), metabolic process, (1,949), localization (664), and biological regulation (624). Within cell component, most of the coding sequences were assigned to the terms like membrane (1,305), membrane part (1,131), cell (1,126), and cell part (1,126). Among the molecular function, the most highly represented GO terms were catalytic activity (2,192), binding (1,547), transporter activity (436), and transcription regulator activity (250). Regarding coding-sequences annotation by KEGG database, proteins were assigned to six categories in KEGG (Fig. 2, B), subdivided into 43 subcategories, and the majority of them belonged to the following classifications: 317 protein sequences in carbohydrate metabolism, 313 in amino acid metabolism, 191 in membrane transport, 174 in metabolism of cofactors and vitamins, 152 in energy metabolism, 152 in cellular community - prokaryotes, and 149 in signal transduction. Overall, the genes assigned to metabolism accounted for the highest proportion.

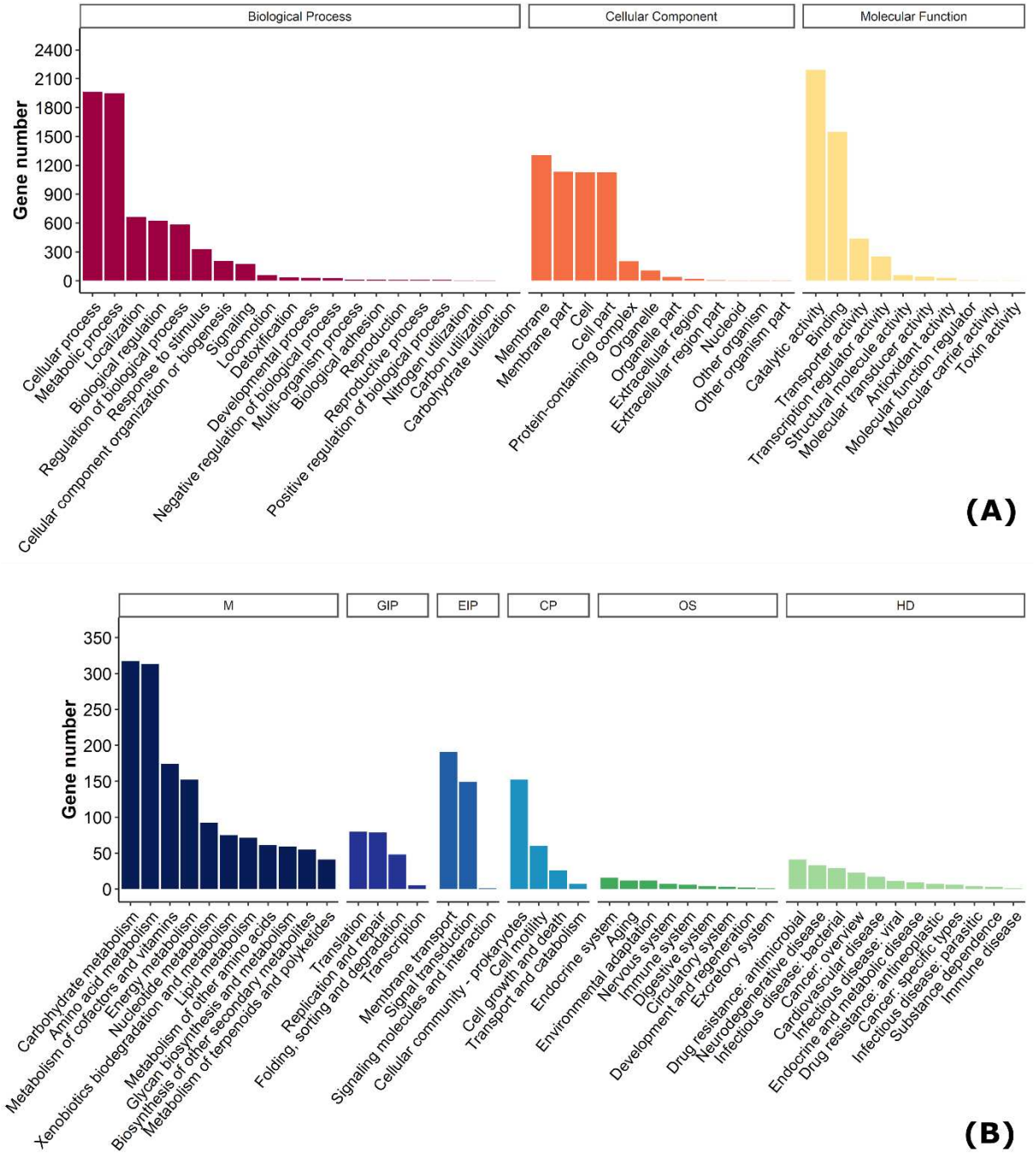


Fig. 2. Coding-sequences analyses of *Pseudomonas paracarnis* RQ057 draft genome. The top x-axis represents the categories, the bottom represents the subcategories, and the y-axis represents the number of genes. **(A)** Gene Ontology (GO) functional classification in three main GO categories (biological process, cellular component, and molecular functional) and 42 subcategories. **(B)** Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway classification in five categories (M: metabolism; GIP: genetic information processing; EIP: environmental information processing; CP: cellular process; OS: organismal systems; HD: human diseases) and 43 subcategories.

Antibiotic and secondary metabolite cluster analysis using AntiSMASH indicated that the draft genome contains 12 putative secondary metabolite clusters (Table 3). Two clusters, cluster 4 and cluster 8, show 10% and 9% similarity with the most similar known pyoverdine biosynthetic gene cluster, respectively. A comparation of the predicted proteins from these clusters with the pyoverdin sequences from *P. fluorescens* A506 showed that four genes involved in the biosynthesis of this siderophore are contained in cluster 4: *pvdS*, *pvdG*, *pvdL* and pvdH (Fig. 3, A). In cluster 8, we identified nine genes: *pvdI*, *pvdD*, *fpvA*, *pvdE*, *pvdF*, *pvdO*, *pvdN*, *pvdM* and *pvdP* (Fig. 3, B).

Table 3. Summary of the secondary metabolite prediction from the *Pseudomonas paracarnis* RQ057 draft genome using antiSMASH server.

Clusters	**Type**	**Size (nt)**	**Most similar known cluster (Similarity %)**
1	RiPP-like	10,879	ND
2	Arylpolyene	43,576	APE Vf biosynthetic gene cluster (45%)
3	Hserlactone	20,579	ND
4	NRPS	52,392	Pyoverdin biosynthetic gene cluster (10%)
5	NAGGN	14,842	ND
6	Redox-cofator	22,148	Lankacidin C biosynthetic gene cluster (13%)
7	Siderophore	11,926	ND
8	NRPS	57,263	Pyoverdin biosynthetic gene cluster (9%)
9	Betalactone	23,173	Fengycin biosynthetic gene cluster (13%)
10	RiPP-like	10,846	ND
11	NRPS	47,513	Pseudomonine biosynthetic gene cluster (100%)
12	NRPS-like	18,308	Fragin biosynthetic gene cluster (25%)

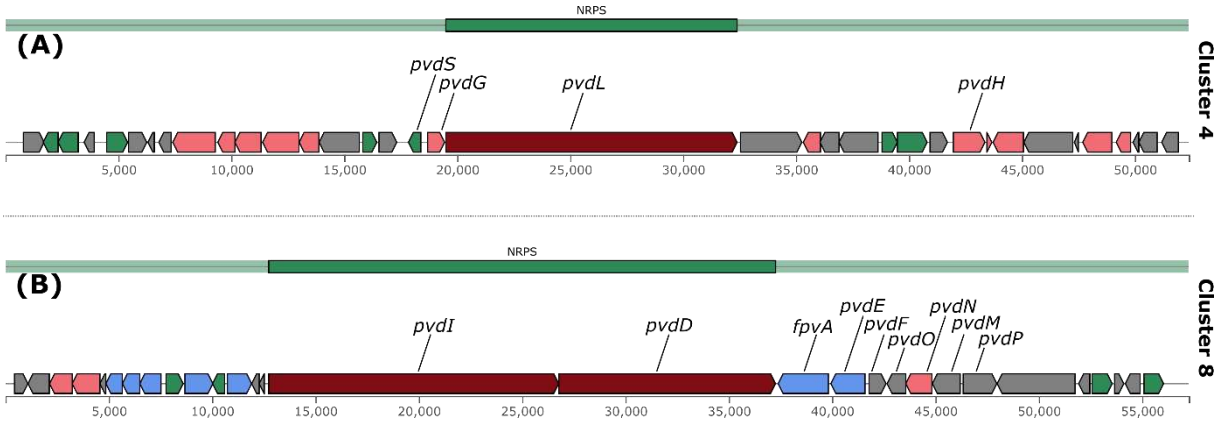


Fig. 3. The two putative pyoverdine biosynthetic gene cluster in *P. paracarnis* RQ057 draft genome, identified using AntiSMASH web server. The names of genes are shown according to the annotation using BLASTp against *P. fluorescens* A506 proteins.

3.3 Phylogenomic analysis

The draft genome sequence data was uploaded to the Type (Strain) Genome Server (TYGS), a free bioinformatics platform for a whole genome-based taxonomic analysis (Meier-Kolthoff and Göker, 2019). This analysis was able to identify the strain under study as belonging to the *Pseudomonas paracarnis* species (Fig. 4). It was also possible to obtain the same result calculating the *in silico* dDDH with GGDC web tool (Fig. 5, A), in which the dDDH values was greater than 70% (92.60%), and ANI based in BLAST (ANIb) (Fig. 5, B), in which it was greater than 95% (98.81%), with *P. paracarnis* V5/DAB/2/5^{T}.

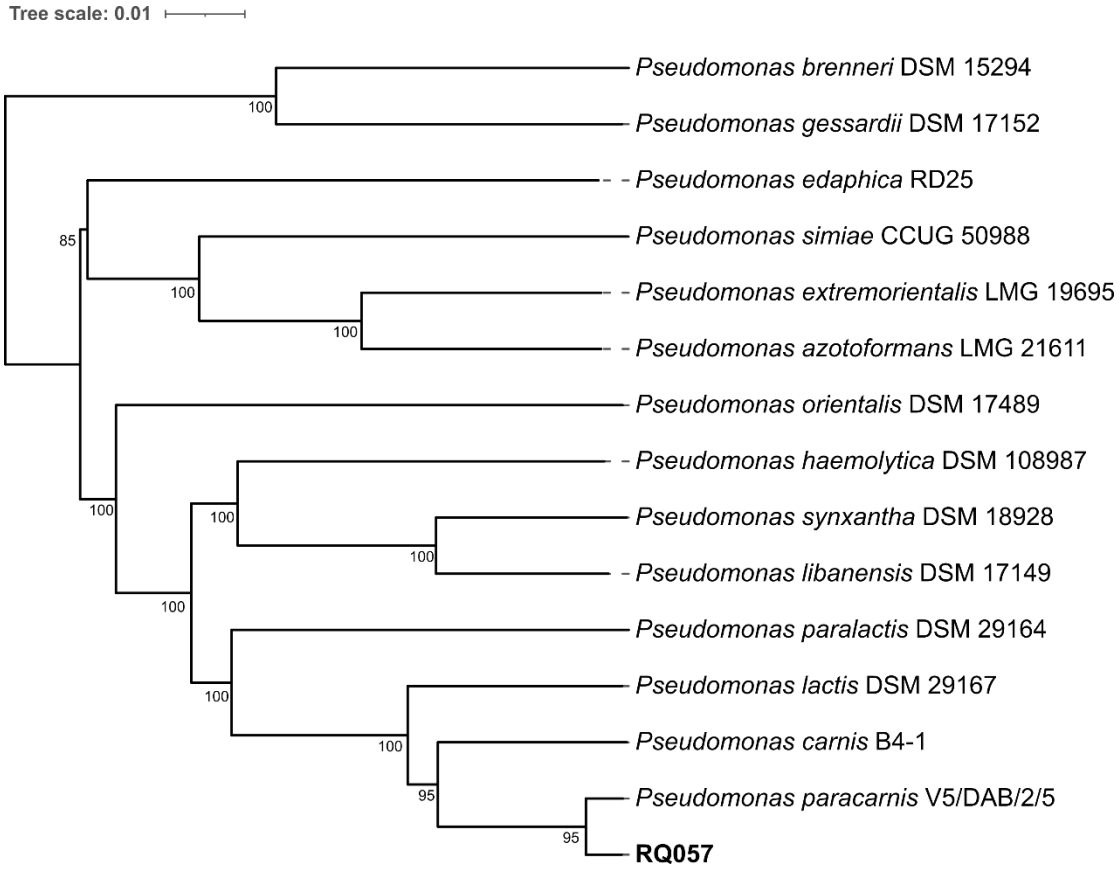


Fig. 4. A phylogenomic tree constructed using the Type Strain Genome Server (TYGS; https://tygs.dsmz.de), for a whole genome-based taxonomic analysis and determination of closely related type strains (Meier-Kolthoff and Göker 2019). Tree inferred with FastME 2.1.6.1 (Lefort et al. 2015) from GBDP distances calculated from genome sequences. The branch lengths are scaled in terms of GBDP distance formula d5. The numbers on branches are GBDP pseudo-bootstrap support values >60% from 100 replications, with an average branch support of 97.9%. The tree was rooted at the midpoint (Farris 1972).

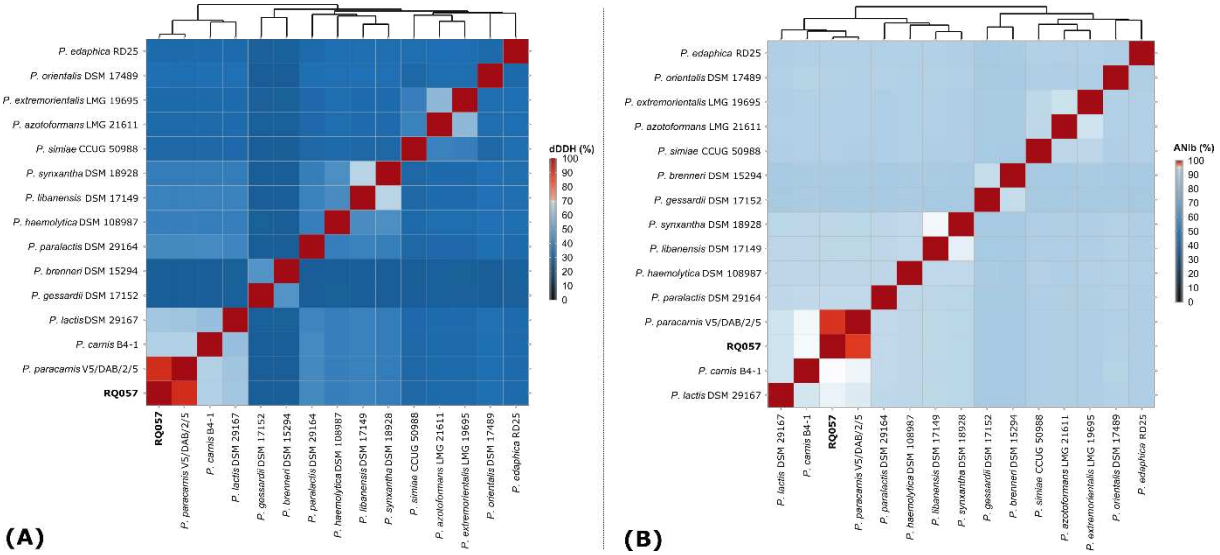


Fig. 5. Heatmap of pairwise comparisons between 15 *Pseudomonas* genomes sequences. **(A)** Heatmap of in silico DNA-DNA hybridization (dDDH). **(B)** Heatmap of average nucleotide identity based on BLASTn algorithm (ANIb). The similarity values are presented in a color gradient heatmap (dDDH: red gradients > 70; white = 70; blue gradients < 70) (ANIb: red gradients > 95; white = 95; blue gradients < 95). The dDDH values of >70% and the ANIb values of > 95% group together genomes from strains of the same species. Drawn dendrograms can be used to visualize the difference between strains.

3.4 Mobile genetic elements search

We identified five prophage regions using PHASTER, two incomplete and three intact, however only intact regions were used for others analyses (Supplementary Table S1). Two regions, termed regions 2 and 3, were the ones with the largest score (150) and the most specific phage-related keywords. Furthermore, they were categorized by VirFam and assigned to Canonical Types according to the attribution described by Lopes et al., (2014). Region 2 was assigned in family *Siphoviridae* of Type 1 (Cluster 3); and Region 3 was assigned in family *Myoviridae* of Type 1 (Cluster 7); both of the order *Caudovirales* (Supplementary Fig. S4 and S5). The Region 1 was also recognized as a *Myoviridae* but could not be further assigned to any of the 4 canonical Types. The VipTree also identified the 2 and 3 regions in the families *Siphoviridae* and *Myoviridae*, respectively, both showed *Gammaproteobacteria* as host group

(Supplementary Fig. S6 and S7). Genomic maps of the intact-prophage regions can be seen in Fig. 6.

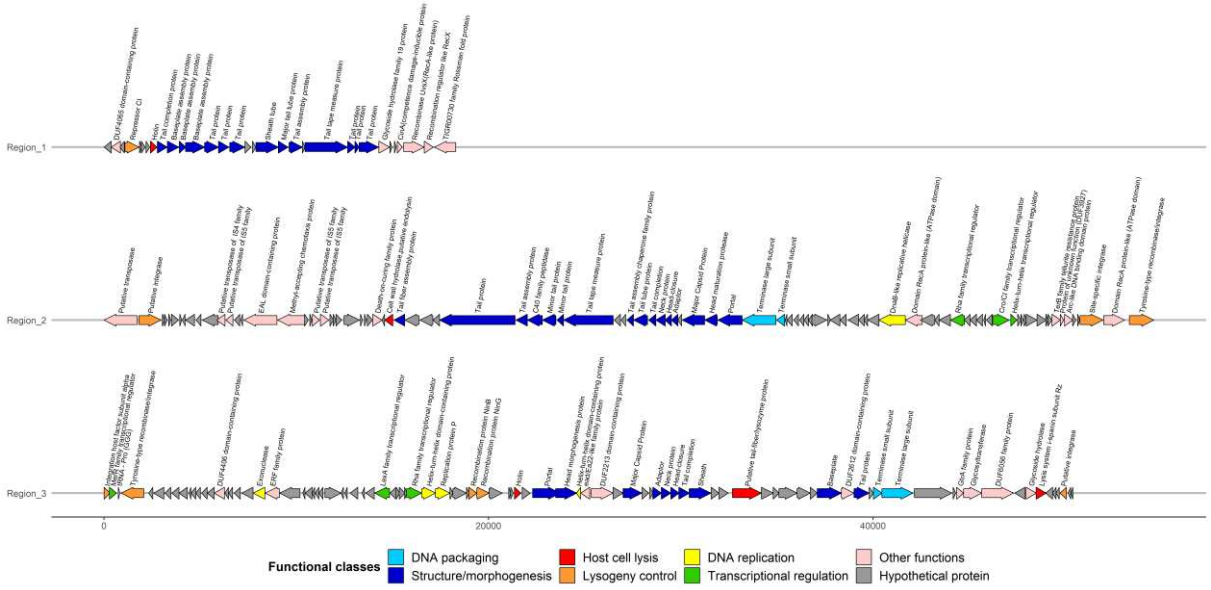


Fig. 6. Schematic representation of protein-coding genes organization in three intact prophages identified in the *P. paracarnis* QR057 draft genome. The direction of the arrows represents the coding strand of the coding sequence. Name above the arrows correspond to annotation, and colored arrows represent putative functional classification for genes.

No putative plasmids were found using PlasmidSPAdes and PlasmidFinder. However, we identified a putative region for integrative conjugative elements (ICE), although the ICEfinder indicated it as a putative ICE with no directed repeats (DR). This region has 52,884 bp, with 52 coding sequences. This putative ICE has genes for proteins identified as type IV transport system – T4SS – (TraD, TraI, TraG, TraU) and also has genes for transposase. Interestingly, *trp* genes (*trp*CBAFD) previously identified as belonging to an exclusive region of blue pigments (Andreani et al., 2019; Reichler et al., 2019) are included therein. The organization of coding sequences of the predicted ICE in *P. paracarnis* RQ057 can be seen in Fig. 7.

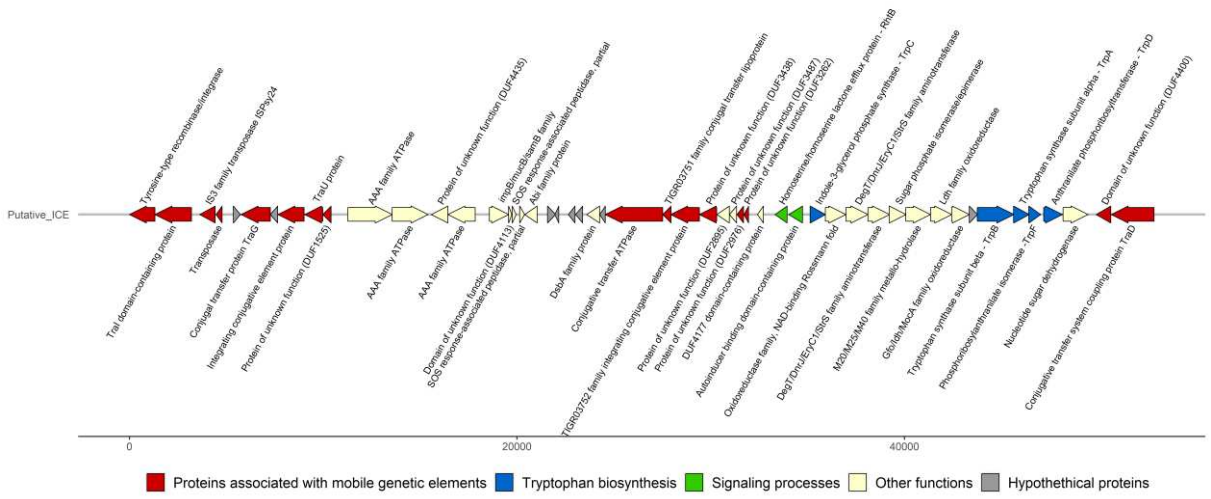


Fig. 7. Schematic representation of protein-coding genes organization in the sequence containing integrative conjugative elements (ICE), identified by ICEfinder, in the *P. paracarnis* QR057 draft genome. The direction of the arrows represents the coding strand of the coding sequence. Name next to the arrows correspond to annotation, and colored arrows represent putative functional classification for genes.

IslandViewer predicted genomic islands (GIs) using IslandPath-DIMOB and SIGI-HMM prediction methods (Supplementary Fig. S8). The IslandPath-DIMOB method, that have better accuracy (Langille et al., 2010), was able to predict 10 entire GIs within contigs. Two particular GIs (named GI_3 and GI_4) had genes identified by ICEfinder, however three *trp* genes (*trp*BAF) were in the interval between those two GIs. Furthermore, in the GI_3, we identified several genes involved with pili construction proteins (as PilM, PilV, PilX, PilP, PilO2, PilN). The genomic maps of the GI_3, the interval, and GI_4 can be seen in Supplementary Fig. S9 and also in Fig. 1.

The *P. paracarnis* strain RQ057 genome has three main clusters which contains genes related with tryptophan biosynthesis: (1) *trpE*, *trpG*, *trpD* and *trpC*; (2) *trpA*, *trpB*, and *trpI*; and, (3) *trpC*, *trpB*, *trpA*, *trpF* and *trpD*. The first cluster is also contained in *P. fluorescens* A506 genome and showed high similarity value (>80%) with the same region from *P. paracarnis* RQ057 (Fig. 8, A). The second cluster also is contained in strain A506, and showed similarity value greater than 92% (Fig. 8, B). Nonetheless, the third region is not contained in

the *P. fluorescens* A506 but it is within the regions identified as ICE and GIs, as explained previously. Furthermore, when comparing this third cluster with the other two clusters within the *P. paracarnis* genome, only small sequences showed similarity values close to 65% (Fig. 8, C and D).

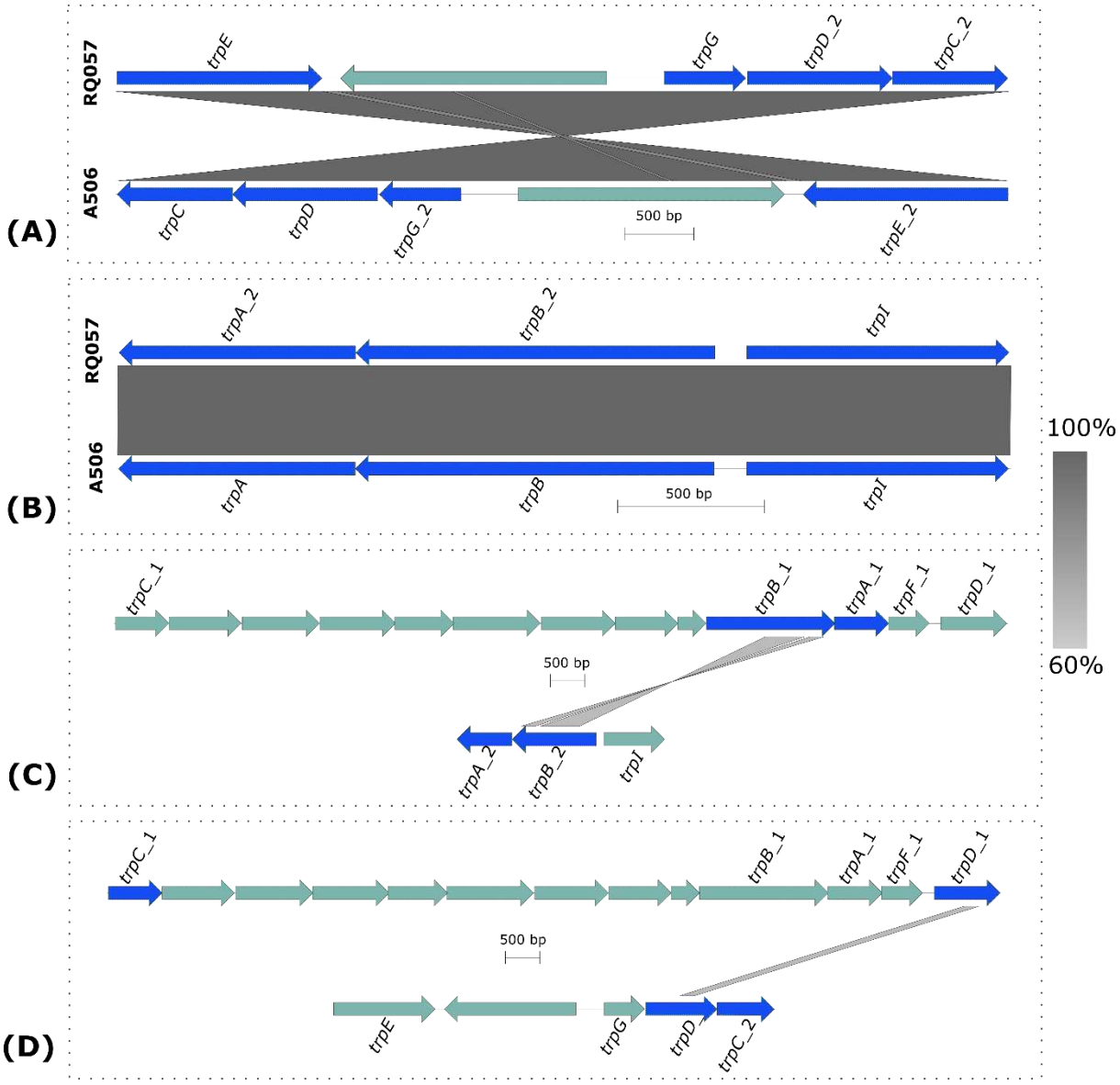


Fig. 8. Schematic representation of gene clusters related to tryptophan biosynthesis in genomes of the *P. paracarnis* RQ057 and *P. fluorescens* A506 (CP003041). The direction of the arrows represents the coding strand of the coding sequence. The dark blue arrows on each show the genes of interest in each alignment. A BLASTn comparison was performed using Easyfig v.2.2.6 software, and the levels of similarity are shown in the gray gradient scale. **(A)** and **(B)** show the alignment of the gene clusters of *P. paracarnis* RQ057 with the same cluster from *P. fluorescens* A506. **(C)** and **(D)** show the alignment of gene cluster contained within exclusive region and the other cluster region of genes related to tryptophan biosynthesis of *P. paracarnis* RQ057.

4. Discussion

Pseudomonas is a genus commonly found in milk and dairy products and its presence is related to the spoilage of these foods (Machado et al., 2015; Scatamburlo et al., 2015). Its ability to produce different pigments has caused important economic losses for cheese manufacturers, mainly due to the production of a blue pigment (Carminati et al., 2019; Martin et al., 2011). The samples analyzed here showed blue discoloration, and the mean count of *Pseudomonas* spp. was approximately 7.7 log CFU/g. High cell counts of *Pseudomonas* are associated with the appearance of blue pigment. Rossi et al. (2018) studying the blue pigment, *in situ*, in artificially contaminated mozzarella cheese samples, observed that the blue color appeared when the samples reached values between 6.5 and 7.5 Log CFU/g storage at 10 °C. In addition, the ability to express this blue phenotype at optimal growth temperature and at low temperatures has also been reported in other studies (Andreani et al., 2014; Carminati et al., 2019). The non-appearance of the pigment at temperatures higher than 30 °C has been a distinct characteristic for most isolates that produce a blue pigment not yet characterized and that has been studied since 2010 by researchers in different countries (Andreani et al., 2014; Chierici et al., 2016; Rossi et al., 2018).

The identical fingerprint pattern of the blue pigmenting isolates suggests that possibly a single strain was present and causing the blue color defects in processed cheese. More than half of the isolates were pigment producers (52%), which indicates two possibilities: either there was a contamination with a high concentration of this strain or it became the dominant during cheese storage. The first hypothesis would be the most accepted, since producers of this pigment have been reported as strong biofilm producers (Rossi et al., 2018). So, a post-processing

contamination probably occurred, since the heat treatment used in the production of processed cheese inactivates *Pseudomonas* (Buňková and Buňka, 2017).

Therefore, we randomly selected a pigment isolate for WGS and genome analysis. The goal of this study was to provide further insight about the genomic content of *Pseudomonas* causing blue discoloration in cheese. We obtained a draft genome using next-generation sequencing technology and, after assembly, it was possible to identify the bacterium under study as belonging to the *P. paracarnis* species. This species was described in January 2021 (Lick et al., 2021), in which the type strain was isolated in 2011 from a roast beef. We chose *P. fluorescens* A506, isolated from pear tree leaf (Loper et al., 2012), as refence to some analyses because this complete genome has shown high similarity with the *P. paracarnis* RQ057 draft genome. In further analyses, we identified that the strain A506 possibly belongs to the *P. paracarnis* species (Appendix S1). Due to the frequent discoveries of new *Pseudomonas* species, many genome sequences deposited in NCBI must have the taxonomy constantly revised.

The presence of two putative pyoverdine biosynthetic gene cluster demonstrates that this *P. paracarnis* strain RQ057 possibly produces this molecule, which is a well-described siderophore (Lamont and Martin, 2003; Meyer and Hornsperger, 1978; Ringel and Brüser, 2018). We identified a total of 13 genes related to the production of this pigment, which are divided into: regulation (*pvdS*), biosynthesis (*pvdG*, *pvdL*, *pvdH*, *pvdI*, *pvdD*, *pvdE*, *pvdF*, *pvdP*), ferripyoverdine receptor (*fpvA*), and modification (*pvdN*, *pvdM*, *pvdO*) (Lamont and Martin, 2003; Ringel and Brüser, 2018). As pyoverdine is a green-yellow molecule (Meyer and Hornsperger, 1978), it also has implications for the spoilage of some foods, because it leaves their surfaces with a green-yellow discoloration.

Most regions predicted as mobile genetic elements (MGE) exhibit GC content below the average GC content of the genome, as shown in Fig. 1. It is well known that MGE usually have a lower GC content than their bacterial hosts, and this can to be used with indicative of the presence of recently acquired mobile elements (Millan et al., 2015; Nishida, 2012). Two identified intact prophages showed the G+C content lower GC content of the draft genome (59.7%), region 2 and 3 with 56.28% and 57.03%, respectively (Table 1; Supplementary Table S1). Valot et al., (2015) studying the genome of *Pseudomonas aeruginosa* found that long-standing prophages show a greater GC content when compared to recently acquired ones, therefore, closer to the content of the host's central genome. The regions containing the ICE sequence and also the GI_3 and GI_4 (Supplementary Fig. S8 and S9), highlighted in Fig. 1, shows that the GC content is also lower than the draft genome average. We identified genes for transposase, type IV transport system (T4SS), and to pili construction proteins in this region (Fig. 7; Supplementary Fig. S9). The predicted proteins are involved with the F-pilus structure; and T4SS translocate DNA and protein substrates across prokaryotic cell envelopes using a mechanism that necessitates direct contact with the target cell, mediated generality by pilus (Alvarez-Martinez and Christie, 2009; Costa et al., 2015). In addition, this region contains genes involved in the tryptophan biosynthesis pathway. Andreani et al. (2019) and Reichler et al. (2019) also showed that *Pseudomonas* strains, capable of producing a distinct blue pigment in food, have an additional cluster of these genes. Here, we also showed that these genes have little or no similarity to the main ones of this biosynthesis pathway (Fig. 8). This could therefore be an indication that this region came from a phylogenetically distant species. Thus, we can propose that this is a region with plasmid-like conjugative properties, mainly because it contains potential conjugative-transposon-like mobile genetic elements.

5. Conclusion

This study isolated and identified a strain of *Pseudomonas* from processed cheese "requeijão em barra" as the cause of blue discoloration. Furthermore, we demonstrated that whole-genome sequencing is a very important method for the proper identification and study of this spoilage. Functional annotation, identification of mobile genetic elements and gene clusters clarified the particular features of this strain that are not species dependent, such as the presence of duplicated genes of tryptophan biosynthesis pathway (*trp*CBAFD). This can have important implications for understanding the particularity of blue discoloration-producing strains. Therefore, we are allowing future research to have the knowledge base to develop tools that can track these spoilers in the food industry and even develop methods that can inhibit the production of the blue pigment.

References

Alvarez-Martinez, C. E., & Christie, P. J. (2009). Biological Diversity of Prokaryotic Type IV Secretion Systems. *Microbiology and Molecular Biology Reviews*, *73*(4), 775–808. https://doi.org/10.1128/mmbr.00023-09

Andreani, N. A., Carraro, L., Zhang, L., Vos, M., & Cardazzo, B. (2019). Transposon mutagenesis in *Pseudomonas fluorescens* reveals genes involved in blue pigment production and antioxidant protection. *Food Microbiology*, *82*, 497–503. https://doi.org/10.1016/j.fm.2019.03.028

Andreani, N. A., Martino, M. E., Fasolato, L., Carraro, L., Montemurro, F., Mioni, R., …

Cardazzo, B. (2014). Tracking the blue: A MLST approach to characterise the *Pseudomonas fluorescens* group. *Food Microbiology*, *39*, 116–126. https://doi.org/10.1016/j.fm.2014.11.009

Andrews, S. (2010). FastQC - A quality control tool for high throughput sequence data. http://www.bioinformatics.babraham.ac.uk/projects/fastqc/. *Babraham Bioinformatics*.

Antipov, D., Hartwick, N., Shen, M., Raiko, M., Lapidus, A., & Pevzner, P. A. (2016). PlasmidSPAdes: Assembling plasmids from whole genome sequencing data. *Bioinformatics*, *32*(22), 3380–3387. https://doi.org/10.1093/bioinformatics/btw493

Arndt, D., Grant, J. R., Marcu, A., Sajed, T., Pon, A., Liang, Y., & Wishart, D. S. (2016). PHASTER: a better, faster version of the PHAST phage search tool. *Nucleic Acids Research*, *44*, W16–W21. https://doi.org/10.1093/nar/gkw387

Bankevich, A., Nurk, S., Antipov, D., Gurevich, A. A., Dvorkin, M., Kulikov, A. S., … Pevzner, P. A. (2012). SPAdes: A new genome assembly algorithm and its applications to single-cell sequencing. *Journal of Computational Biology*, *19*(5), 455–477. https://doi.org/10.1089/cmb.2012.0021

Bateman, A., Martin, M. J., Orchard, S., Magrane, M., Agivetova, R., Ahmad, S., … Zhang, J. (2021). UniProt: The universal protein knowledgebase in 2021. *Nucleic Acids Research*, *49*, D480–D489. https://doi.org/10.1093/nar/gkaa1100

Bertelli, C., Laird, M. R., Williams, K. P., Lau, B. Y., Hoad, G., Winsor, G. L., & Brinkman, F. S. L. (2017). IslandViewer 4: Expanded prediction of genomic islands for larger-scale datasets. *Nucleic Acids Research*, *45*(W1), W30–W35. https://doi.org/10.1093/nar/gkx343

Blin, K., Shaw, S., Steinke, K., Villebro, R., Ziemert, N., Lee, S. Y., … Weber, T. (2019). AntiSMASH 5.0: Updates to the secondary metabolite genome mining pipeline. *Nucleic Acids Research*, *47*, W81–W87. https://doi.org/10.1093/nar/gkz310

Bolger, A. M., Lohse, M., & Usadel, B. (2014). Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics*, *30*(15), 2114–2120. https://doi.org/10.1093/bioinformatics/btu170

Buňková, L., & Buňka, F. (2017). Microflora of processed cheese and the factors affecting it. *Critical Reviews in Food Science and Nutrition*, *57*(11), 2392–2403. https://doi.org/https://doi.org/10.1080/10408398.2015.1060939

Carattoli, A., Zankari, E., Garciá-Fernández, A., Larsen, M. V., Lund, O., Villa, L., … Hasman, H. (2014). In Silico detection and typing of plasmids using plasmidfinder and plasmid multilocus sequence typing. *Antimicrobial Agents and Chemotherapy*, *58*(7), 3895–3903. https://doi.org/10.1128/AAC.02412-14

Carminati, D., Bonvini, B., Rossetti, L., Zago, M., Tidona, F., & Giraffa, G. (2019). Investigation on the presence of blue pigment-producing *Pseudomonas* strains along a production line of fresh mozzarella cheese. *Food Control*, *100*, 321–328. https://doi.org/10.1016/j.foodcont.2019.02.009

Carrascosa, C., Martínez, R., Sanjuán, E., Millán, R., del Rosario-Quintana, C., Acosta, F., … Jaber, J. R. (2021). Identification of the *Pseudomonas fluorescens* group as being responsible for blue pigment on fresh cheese. *Journal of Dairy Science*, *104*(6). https://doi.org/10.3168/jds.2020-19517

Carrascosa, C., Millán, R., Jaber, J. R., Lupiola, P., del Rosario-Quintana, C., Mauricio, C., & Sanjuán, E. (2015). Blue pigment in fresh cheese produced by *Pseudomonas fluorescens*. *Food Control*, *54*, 95–102. https://doi.org/10.1016/j.foodcont.2014.12.039

Cenci-Goga, B. T., Karama, M., Sechi, P., Iulietto, M. F., Novelli, S., & Mattei, S. (2014). Evolution under different storage conditions of anomalous blue coloration of Mozzarella cheese intentionally contaminated with a pigment-producing strain of *Pseudomonas*

fluorescens. *Journal of Dairy Science*, *97*, 6708–6718. https://doi.org/10.3168/jds.2014-8611

Chen, H. (2018). *VennDiagram: Generate High-Resolution Venn and Euler Plots*. Retrieved from https://cran.r-project.org/package=VennDiagram

Chevreux, B., Wetter, T., & Suhai, S. (1999). Genome Sequence Assembly Using Trace Signals and Additional Sequence Information. *Computer Science and Biology: Proceedings of the German Conference on Bioinformatics (GCB) '99, GCB, Hannover, Germany.*, *99*, 45–56.

Chierici, M., Picozzi, C., La Spina, M. G., Orsi, C., Vigentini, I., Zambrini, V., & Foschino, R. (2016). Strain diversity of *Pseudomonas fluorescens* group with potential blue pigment phenotype isolated from dairy products. *Journal of Food Protection*, *79*(8), 1430–1435. https://doi.org/10.4315/0362-028X.JFP-15-589

Conesa, A., Götz, S., García-Gómez, J. M., Terol, J., Talón, M., & Robles, M. (2005). Blast2GO: A universal tool for annotation, visualization and analysis in functional genomics research. *Bioinformatics*, *21*(18), 3674–3676. https://doi.org/10.1093/bioinformatics/bti610

Costa, T. R. D., Felisberto-Rodrigues, C., Meir, A., Prevost, M. S., Redzej, A., Trokter, M., & Waksman, G. (2015). Secretion systems in Gram-negative bacteria: Structural and mechanistic insights. *Nature Reviews Microbiology*, *13*, 343–359. https://doi.org/10.1038/nrmicro3456

Ecale Zhou, C. L., Malfatti, S., Kimbrel, J., Philipson, C., McNair, K., Hamilton, T., … Souza, B. (2019). MultiPhATE: Bioinformatics pipeline for functional annotation of phage isolates. *Bioinformatics*, *35*(21), 4402–4404. https://doi.org/10.1093/bioinformatics/btz258

Garrido-Sanz, D., Meier-Kolthoff, J. P., Göker, M., Martín, M., Rivilla, R., & Redondo-Nieto, M. (2016). Genomic and genetic diversity within the *Pseudomonas fluorescens* complex. *PLoS ONE*, *11*(2). https://doi.org/10.1371/journal.pone.0150183

Glück, C., Rentschler, E., Krewinkel, M., Merz, M., von Neubeck, M., Wenning, M., … Fischer, L. (2016). Thermostability of peptidases secreted by microorganisms associated with raw milk. *International Dairy Journal*, *56*, 186–197. https://doi.org/10.1016/j.idairyj.2016.01.025

Gurevich, A., Saveliev, V., Vyahhi, N., & Tesler, G. (2013). QUAST: Quality assessment tool for genome assemblies. *Bioinformatics*. https://doi.org/10.1093/bioinformatics/btt086

Henry, L., Wickham, H., & Chang, W. (2020). *ggstance: Horizontal "ggplot2" Components*. Retrieved from https://cran.r-project.org/package=ggstance

Hyatt, D., Chen, G.-L., LoCascio, P. F., Land, M. L., Larimer, F. W., & Hauser, L. J. (2010). Prodigal: prokaryotic gene recognition and translation initiation site identification. *BMC Bioinformatics*, *11*(1), 119. https://doi.org/10.1186/1471-2105-11-119

Kanehisa, M., & Goto, S. (2000). KEGG: Kyoto Encyclopedia of Genes and Genomes. *Nucleic Acids Research*, *28*(1), 27–30. https://doi.org/10.1093/nar/28.1.27

Kirner, S., Krauss, S., Sury, G., Lam, S. T., Ligon, J. M., & Pee, K. Van. (1996). The non-haem chloroperoxidase from *Pseudomonas fluorescens* and its relationship to pyrrolnitrin biosynthesis. *Microbiology*, *142*(8), 2129–2135.

Kolbe, D. L., & Eddy, S. R. (2011). Fast filtering for RNA homology search. *Bioinformatics*, *27*(22), 3102–3109. https://doi.org/10.1093/bioinformatics/btr545

Krzywinski, M., Schein, J., Birol, I., Connors, J., Gascoyne, R., Horsman, D., … Marra, M. A. (2009). Circos: An information aesthetic for comparative genomics. *Genome Research*,

19(9), 1639–1645. https://doi.org/10.1101/gr.092759.109

Lamont, I. L., & Martin, L. W. (2003). Identification and characterization of novel pyoverdine synthesis genes in *Pseudomonas aeruginosa*. *Microbiology*, (149), 833–842. https://doi.org/10.1099/mic.0.26085-0

Langille, M. G. I., Hsiao, W. W. L., & Brinkman, F. S. L. (2010). Detecting genomic islands using bioinformatics approaches. *Nature Reviews Microbiology*, *8*, 373–382. https://doi.org/10.1038/nrmicro2350

Laslett, D., & Canback, B. (2004). ARAGORN, a program to detect tRNA genes and tmRNA genes in nucleotide sequences. *Nucleic Acids Research*, *32*(1), 11–16. https://doi.org/10.1093/nar/gkh152

Lee, I., Chalita, M., Ha, S. M., Na, S. I., Yoon, S. H., & Chun, J. (2017). ContEst16S: An algorithm that identifies contaminated prokaryotic genomes using 16S RNA gene sequences. *International Journal of Systematic and Evolutionary Microbiology*, *67*, 2053–2057. https://doi.org/10.1099/ijsem.0.001872

Letunic, I., & Bork, P. (2016). Interactive tree of life (iTOL) v3: an online tool for the display and annotation of phylogenetic and other trees. *Nucleic Acids Research*, *44*, W242–W245. https://doi.org/10.1093/nar/gkw290

Lick, S., Wibberg, D., Winkler, A., Blom, J., Grimmler, C., Goesmann, A., … Kröckel, L. (2021). *Pseudomonas paracarnis* sp. nov., isolated from refrigerated beef. *International Journal of Systematic and Evolutionary Microbiology*, *71*, 004652. https://doi.org/10.1099/ijsem.0.004652

Liu, M., Li, X., Xie, Y., Bi, D., Sun, J., Li, J., … Ou, H. Y. (2019). ICEberg 2.0: An updated database of bacterial integrative and conjugative elements. *Nucleic Acids Research*, *47*, D660–D665. https://doi.org/10.1093/nar/gky1123

Loper, J. E., Hassan, K. A., Mavrodi, D. V., Davis, E. W., Lim, C. K., Shaffer, B. T., … Paulsen, I. T. (2012). Comparative genomics of plant-associated *Pseudomonas* spp.: Insights into diversity and inheritance of traits involved in multitrophic interactions. *PLoS Genetics*, *8*(7), e1002784. https://doi.org/10.1371/journal.pgen.1002784

Lopes, A., Amarir-Bouhram, J., Faure, G., Petit, M. A., & Guerois, R. (2010). Detection of novel recombinases in bacteriophage genomes unveils Rad52, Rad51 and Gp2.5 remote homologs. *Nucleic Acids Research*, *38*(12), 3952–3962. https://doi.org/10.1093/nar/gkq096

Lopes, A., Tavares, P., Petit, M. A., Guérois, R., & Zinn-Justin, S. (2014). Automated classification of tailed bacteriophages according to their neck organization. *BMC Genomics*, *15*, 1–17. https://doi.org/10.1186/1471-2164-15-1027

Luo, R., Liu, B., Xie, Y., Li, Z., Huang, W., Yuan, J., … Wang, J. (2012). SOAPdenovo2: An empirically improved memory-efficient short-read de novo assembler. *GigaScience*, *1*(18). https://doi.org/10.1186/2047-217X-1-18

Machado, S. G., Baglinière, F., Marchand, S., Coillie, E. Van, Vanetti, M. C. D., Block, J. De, & Heyndrickx, M. (2017). The biodiversity of the microbiota producing heat-resistant enzymes responsible for spoilage in processed bovine milk and dairy products. *Frontiers in Microbiology*, *8*(302), 1–22. https://doi.org/10.3389/fmicb.2017.00302

Machado, S. G., da Silva, F. L., Bazzolli, D. M. S., Heyndrickx, M., Costa, P. M. d. A., & Vanetti, M. C. D. (2015). *Pseudomonas* spp. and *Serratia liquefaciens* as predominant spoilers in cold raw milk. *Journal of Food Science*, *80*(8), M1842–M1849. https://doi.org/10.1111/1750-3841.12957

Maier, C., Huptas, C., von Neubeck, M., Scherer, S., Wenning, M., & Lücking, G. (2020). Genetic organization of the *aprX-lipA2* operon affects the proteolytic potential of

Pseudomonas species in milk. *Frontiers in Microbiology*, *11*(June), 1–13. https://doi.org/10.3389/fmicb.2020.01190

Marchler-Bauer, A., & Bryant, S. H. (2004). CD-Search: Protein domain annotations on the fly. *Nucleic Acids Research*, *32*, W327–W331. https://doi.org/10.1093/nar/gkh454

Martin, N. H., Murphy, S. C., Ralyea, R. D., Wiedmann, M., & Boor, K. J. (2011). When cheese gets the blues: *Pseudomonas fluorescens* as the causative agent of cheese spoilage. *Journal of Dairy Science*, *94*, 3176–3183. https://doi.org/10.3168/jds.2011-4312

Mcnair, K., Zhou, C., Dinsdale, E. A., Souza, B., & Edwards, R. A. (2019). PHANOTATE: A novel approach to gene identification in phage genomes. *Bioinformatics*, *35*(22), 4537–4542. https://doi.org/10.1093/bioinformatics/btz265

Meier-Kolthoff, J. P., Auch, A. F., Klenk, H. P., & Göker, M. (2013). Genome sequence-based species delimitation with confidence intervals and improved distance functions. *BMC Bioinformatics*, *14*(60). https://doi.org/10.1186/1471-2105-14-60

Meier-Kolthoff, J. P., & Göker, M. (2019). TYGS is an automated high-throughput platform for state-of-the-art genome-based taxonomy. *Nature Communications*, *10*, 2182. https://doi.org/10.1038/s41467-019-10210-3

Meyer, J. M., & Hornsperger, J. M. (1978). Role of pyoverdine(pf), the iron-binding fluorescent pigment of *Pseudomonas fluorescens*, in iron transport. *Journal of General Microbiology*, *107*, 329–331. https://doi.org/10.1099/00221287-107-2-329

Millan, A. S., Toll-Riera, M., Qi, Q., & MacLean, R. C. (2015). Interactions between horizontally acquired genes create a fitness cost in *Pseudomonas aeruginosa*. *Nature Communications*, *6*, 4–11. https://doi.org/10.1038/ncomms7845

Mistry, J., Chuguransky, S., Williams, L., Qureshi, M., Salazar, G. A., Sonnhammer, E. L. L.,

… Bateman, A. (2021). Pfam: The protein families database in 2021. *Nucleic Acids Research*, *49*, D412–D419. https://doi.org/10.1093/nar/gkaa913

Moriya, Y., Itoh, M., Okuda, S., Yoshizawa, A. C., & Kanehisa, M. (2007). KAAS: An automatic genome annotation and pathway reconstruction server. *Nucleic Acids Research*, *35*, W182–W185. https://doi.org/10.1093/nar/gkm321

Nishida, H. (2012). Comparative analyses of base compositions, DNA sizes, and dinucleotide frequency profiles in archaeal and bacterial chromosomes and plasmids. *International Journal of Evolutionary Biology*, *2012*, 1–5. https://doi.org/10.1155/2012/342482

Nishimura, Y., Yoshida, T., Kuronishi, M., Uehara, H., Ogata, H., & Goto, S. (2017). ViPTree: The viral proteomic tree server. *Bioinformatics*, *33*(15), 2379–2380. https://doi.org/10.1093/bioinformatics/btx157

Oliveros, J. C. (2007). Venny 2.1. An interactive tool for comparing lists with Venn's diagrams. Retrieved March 1, 2021, from https://bioinfogp.cnb.csic.es/tools/venny/index.html

Palleroni, N. J. (2005). Genus I *Pseudomonas* Migula 1894. In D. J. Brenner, N. R. Krieg, & J. T. Staley (Eds.), *Bergey's Manual of Systematic Bacteriology, The Proteobacteria Part B The Gammaproteobacteria. Vol. 2* (2nd ed., pp. 323–379). East Lansing, Michigan: Springer.

Quinlan, A. R., & Hall, I. M. (2010). BEDTools: A flexible suite of utilities for comparing genomic features. *Bioinformatics*, *26*(6), 841–842. https://doi.org/10.1093/bioinformatics/btq033

Quintieri, L., Caputo, L., De Angelis, M., & Fanelli, F. (2020). Genomic analysis of three cheese-borne *Pseudomonas lactis* with biofilm and spoilage-associated behavior. *Microorganisms*, *8*(8), 1208. https://doi.org/10.3390/microorganisms8081208

R Core Team. (2020). *R: A Language and Environment for Statistical Computing*. Retrieved from https://www.r-project.org/

Reichler, S. J., Martin, N. H., Evanowski, R. L., Kovac, J., Wiedmann, M., & Orsi, R. H. (2019). A century of gray: A genomic locus found in 2 distinct *Pseudomonas* spp. is associated with historical and contemporary color defects in dairy products worldwide. *Journal of Dairy Science*, *102*, 5979–6000. https://doi.org/10.3168/jds.2018-16192

Richter, M., Rosselló-Móra, R., Oliver Glöckner, F., & Peplies, J. (2016). JSpeciesWS: A web server for prokaryotic species circumscription based on pairwise genome comparison. *Bioinformatics*, *32*(6), 929–931. https://doi.org/10.1093/bioinformatics/btv681

Ringel, M. T., & Brüser, T. (2018). The biosynthesis of pyoverdines. *Microbial Cell*, *5*(10), 424–437. https://doi.org/10.15698/mic2018.10.649

Rodrigues, R. da S., Machado, S. G., de Carvalho, A. F., & Nero, L. A. (2021). *Pseudomonas* sp. as the causative agent of anomalous blue discoloration in Brazilian fresh soft cheese (Minas Frescal). *International Dairy Journal*, *117*, 105020. https://doi.org/10.1016/j.idairyj.2021.105020

Rossi, C., Serio, A., Chaves-López, C., Anniballi, F., Auricchio, B., Goffredo, E., … Paparella, A. (2018). Biofilm formation, pigment production and motility in *Pseudomonas* spp. isolated from the dairy industry. *Food Control*, *86*, 241–248. https://doi.org/10.1016/j.foodcont.2017.11.018

RStudio Team. (2020). *RStudio: Integrated Development Environment for R*. Retrieved from http://www.rstudio.com/

Scatamburlo, T. M., Yamazi, A. K., Cavicchioli, V. Q., Pieri, F. A., & Nero, L. A. (2015). Spoilage potential of *Pseudomonas* species isolated from goat milk. *Journal of Dairy Science*, *98*, 759–764. https://doi.org/10.3168/jds.2014-8747

Seemann, T. (2014). Prokka: Rapid prokaryotic genome annotation. *Bioinformatics*, *30*(14), 2068–2069. https://doi.org/10.1093/bioinformatics/btu153

Shen, W., Le, S., Li, Y., & Hu, F. (2016). SeqKit: A cross-platform and ultrafast toolkit for FASTA/Q file manipulation. *PLoS ONE*, *11*(10), e0163962. https://doi.org/10.1371/journal.pone.0163962

Spilker, T., Coenye, T., Vandamme, P., & LiPuma, J. J. (2004). PCR-based assay for differentiation of *Pseudomonas aeruginosa* from other *Pseudomonas* species recovered from cystic fibrosis patients. *Journal of Clinical Microbiology*, *45*(5), 2074–2079. https://doi.org/10.1128/JCM.42.5.2074-2079.2004

Sullivan, M. J., Petty, N. K., & Beatson, S. A. (2011). Easyfig: A genome comparison visualizer. *Bioinformatics*, *27*(7), 1009–1010. https://doi.org/10.1093/bioinformatics/btr039

Valot, B., Guyeux, C., Rolland, J. Y., Mazouzi, K., Bertrand, X., & Hocquet, D. (2015). What it takes to be a *Pseudomonas aeruginosa*? The core genome of the opportunistic pathogen updated. *PLoS ONE*, *10*(5), 1–15. https://doi.org/10.1371/journal.pone.0126468

Versalovic, J., Schneider, M., De Bruijn, F. J., & Lupski, J. R. (1994). Genomic fingerprinting of bacteria using repetitive sequence-based polymerase chain reaction. *Methods in Molecular and Cellular Biology*, *5*, 25–40.

Wickham, H. (2011). The split-apply-combine strategy for data analysis. *Journal of Statistical Software*, *40*(1), 1–29. Retrieved from http://www.jstatsoft.org/v40/i01/

Wickham, H. (2016). *ggplot2: Elegant Graphics for Data Analysis*. Retrieved from https://ggplot2.tidyverse.org

Wickham, H., François, R., Henry, L., & Müller, K. (2021). *dplyr: A Grammar of Data*

Manipulation. Retrieved from https://cran.r-project.org/web/packages/dplyr/index.html

Wilkins, D. (2020). *gggenes: Draw Gene Arrow Maps in "ggplot2."* Retrieved from https://cran.r-project.org/package=gggenes

Ye, J., Fang, L., Zheng, H., Zhang, Y., Chen, J., Zhang, Z., … Wang, J. (2006). WEGO: A web tool for plotting GO annotations. *Nucleic Acids Research*, *34*, W293–W297. https://doi.org/10.1093/nar/gkl031

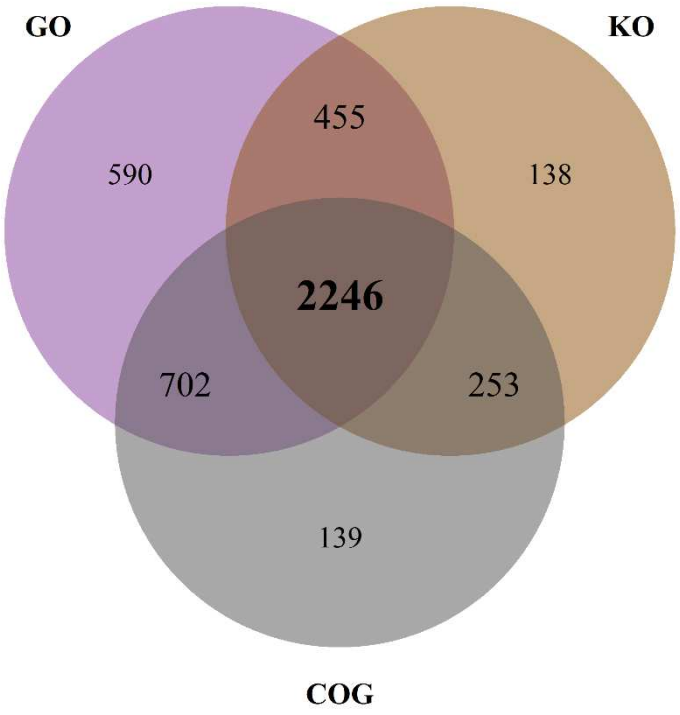
Ye, J., Zhang, Y., Cui, H., Liu, J., Wu, Y., Cheng, Y., … Shi, C. (2018). WEGO 2.0: A web tool for analyzing and plotting GO annotations, 2018 update. *Nucleic Acids Research*, *46*, W71–W75. https://doi.org/10.1093/nar/gky400

Zerbino, D. R., & Birney, E. (2008). Velvet: Algorithms for de novo short read assembly using de Bruijn graphs. *Genome Research*, *18*, 821–829. https://doi.org/10.1101/gr.074492.107

Supplementary Material

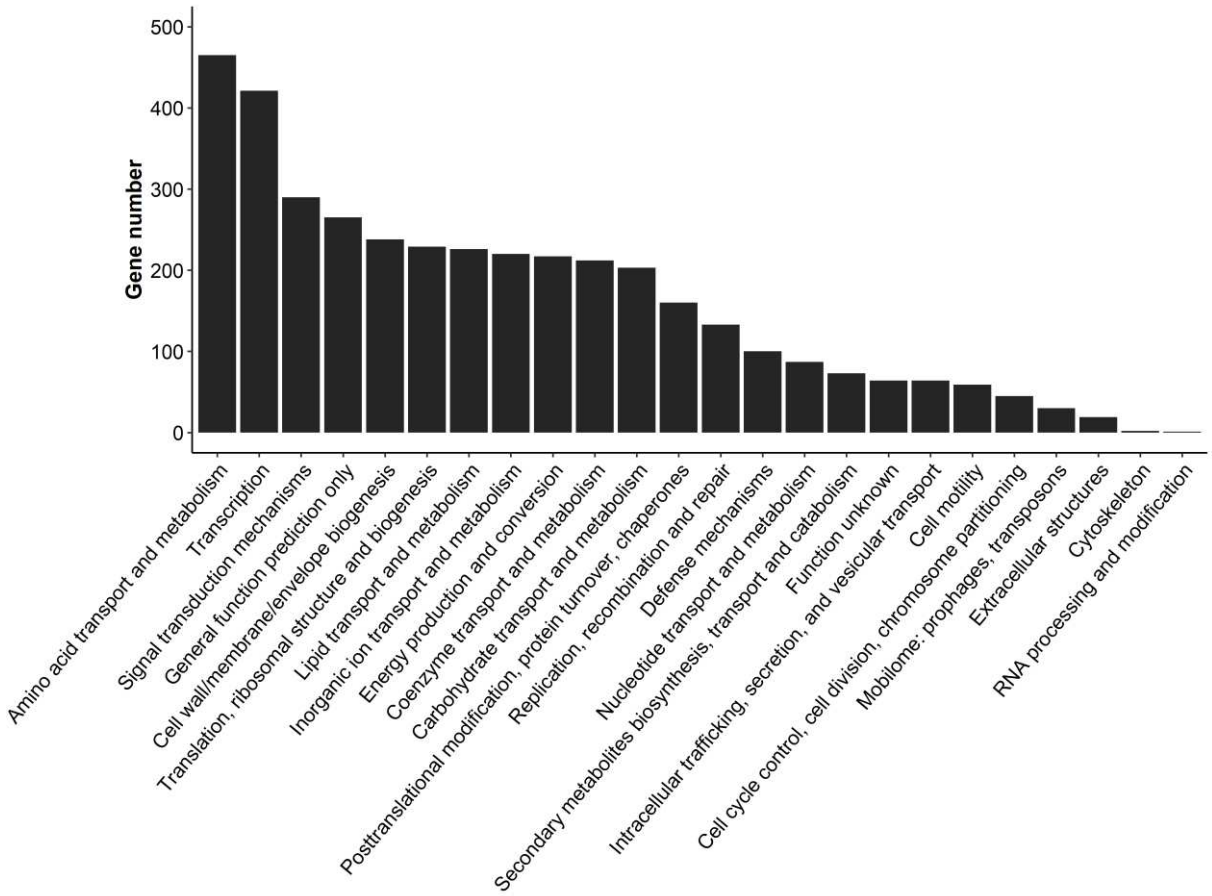
Supplementary Fig. S1.

Dendrogram of rep-PCR $(GTG)_5$ fingerprints constructed by BioNumerics software version 6.6 using UPGMA and the Dice coefficient (optimization 5%, tolerance 1.5%). Node values indicate the percentage of similarity.



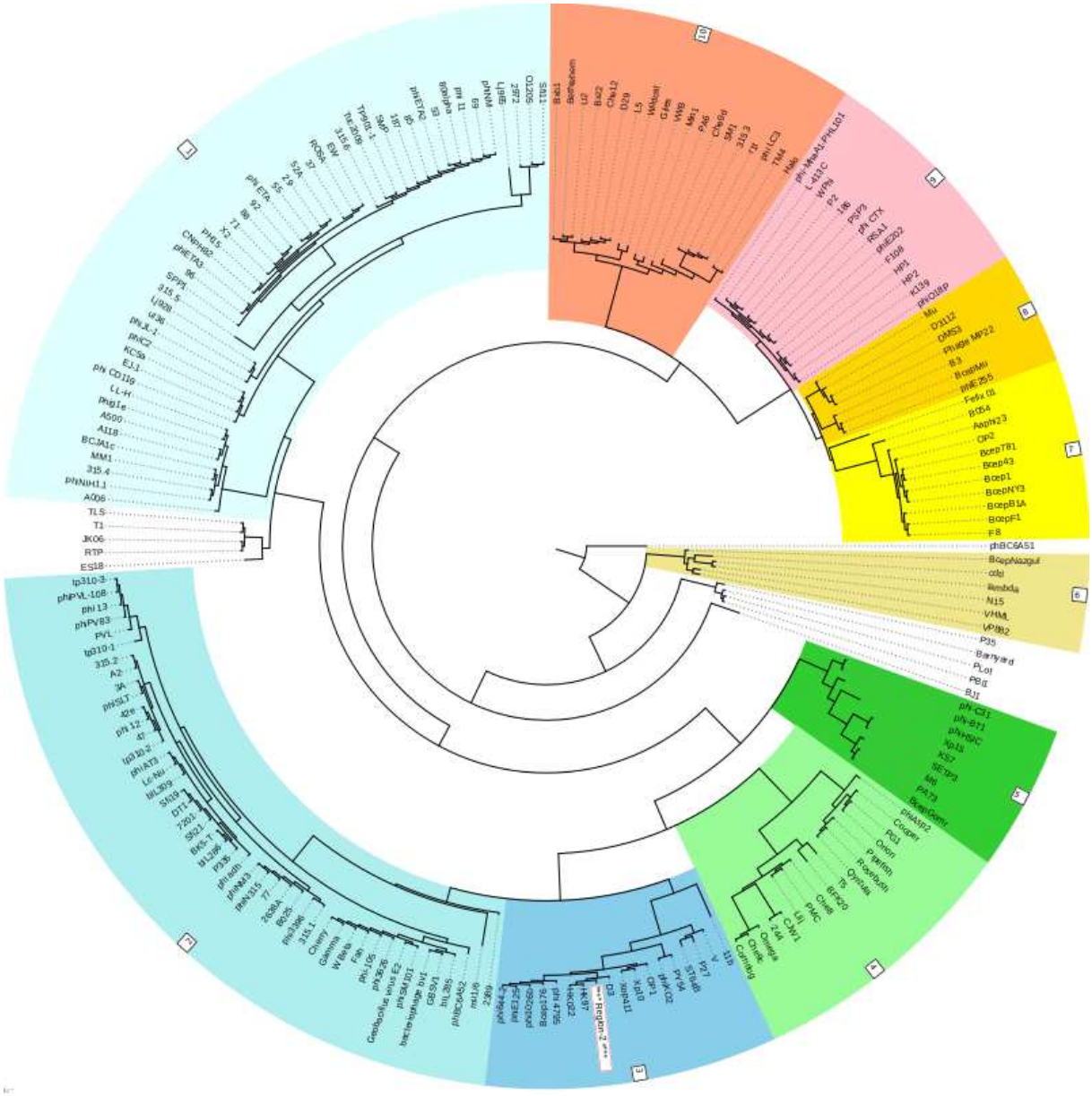
Supplementary Fig. S2.

A Venn diagram illustrating shared predicted protein-coding genes annotated in Clusters of Orthologous Genes (COG), KEGG Orthology (KO), and Gene Ontology (GO) databases. A total of 2,246 genes between the databases were identified.



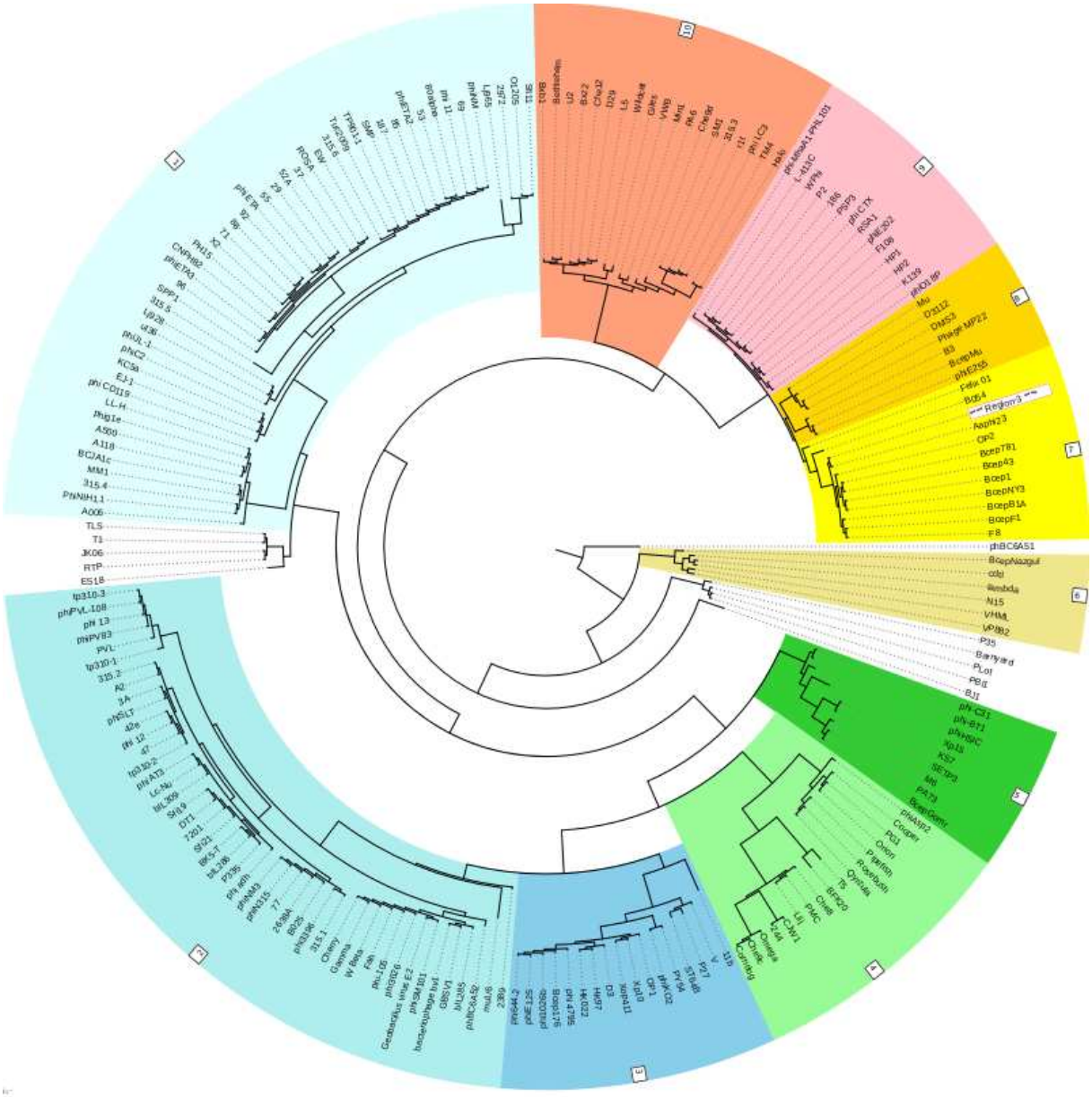
Supplementary Fig. S3.

Coding-sequences functional classification in Clusters of Orthologous Genes (COG) of the *P. paracarnis* RQ057 draft genome, summarized in 24 categories. The x-axis represents the categories and the y-axis represents the number of genes in each categories.



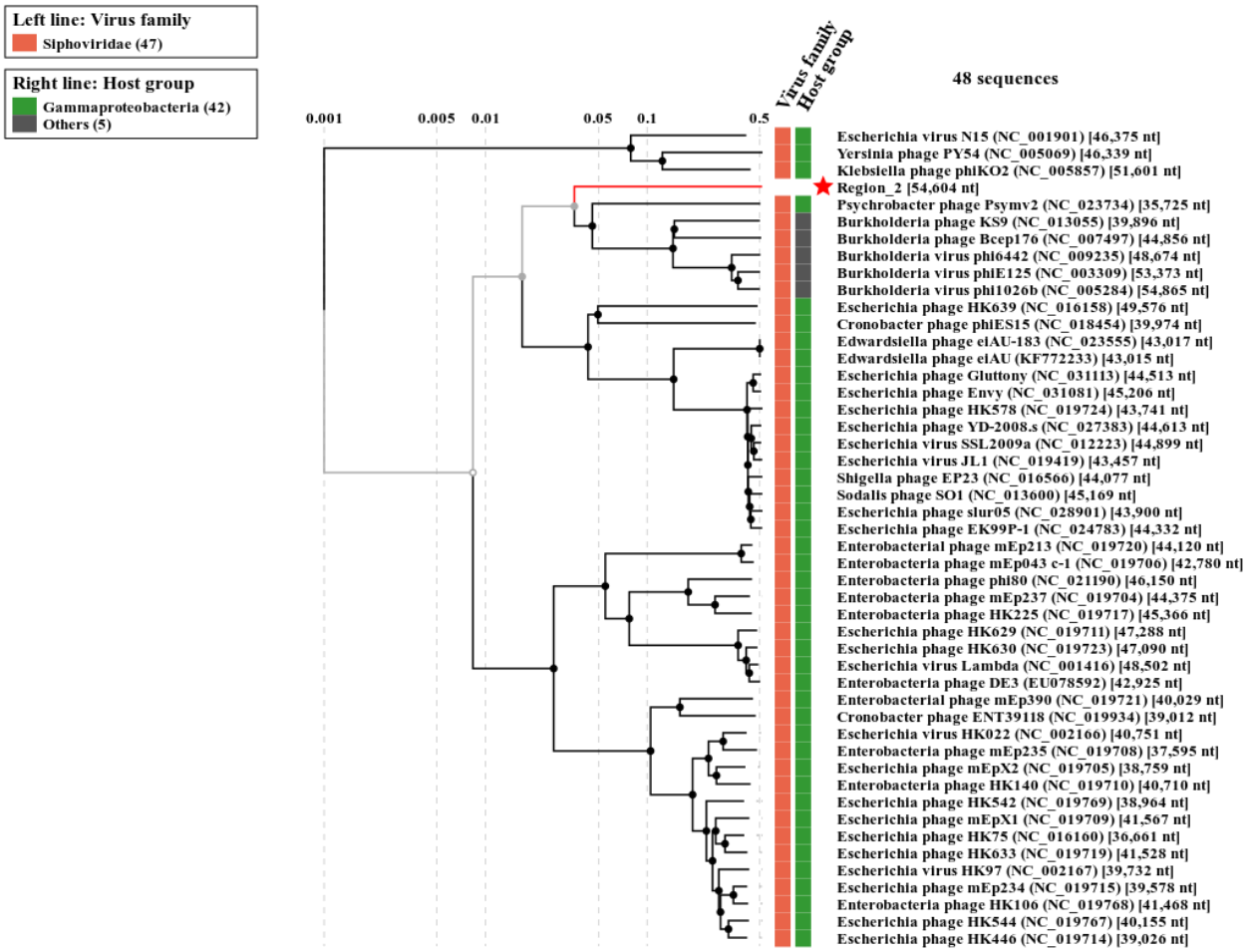
Supplementary Fig. S4.

Circular tree built automatically in Virfam with Aclame from the sequence of predicted prophage named Region 2. The different branches of the tree are sorted into 10 clusters, highlighted by different background colors. The prophage under study is in the dark blue cluster, in a red box with white filling.



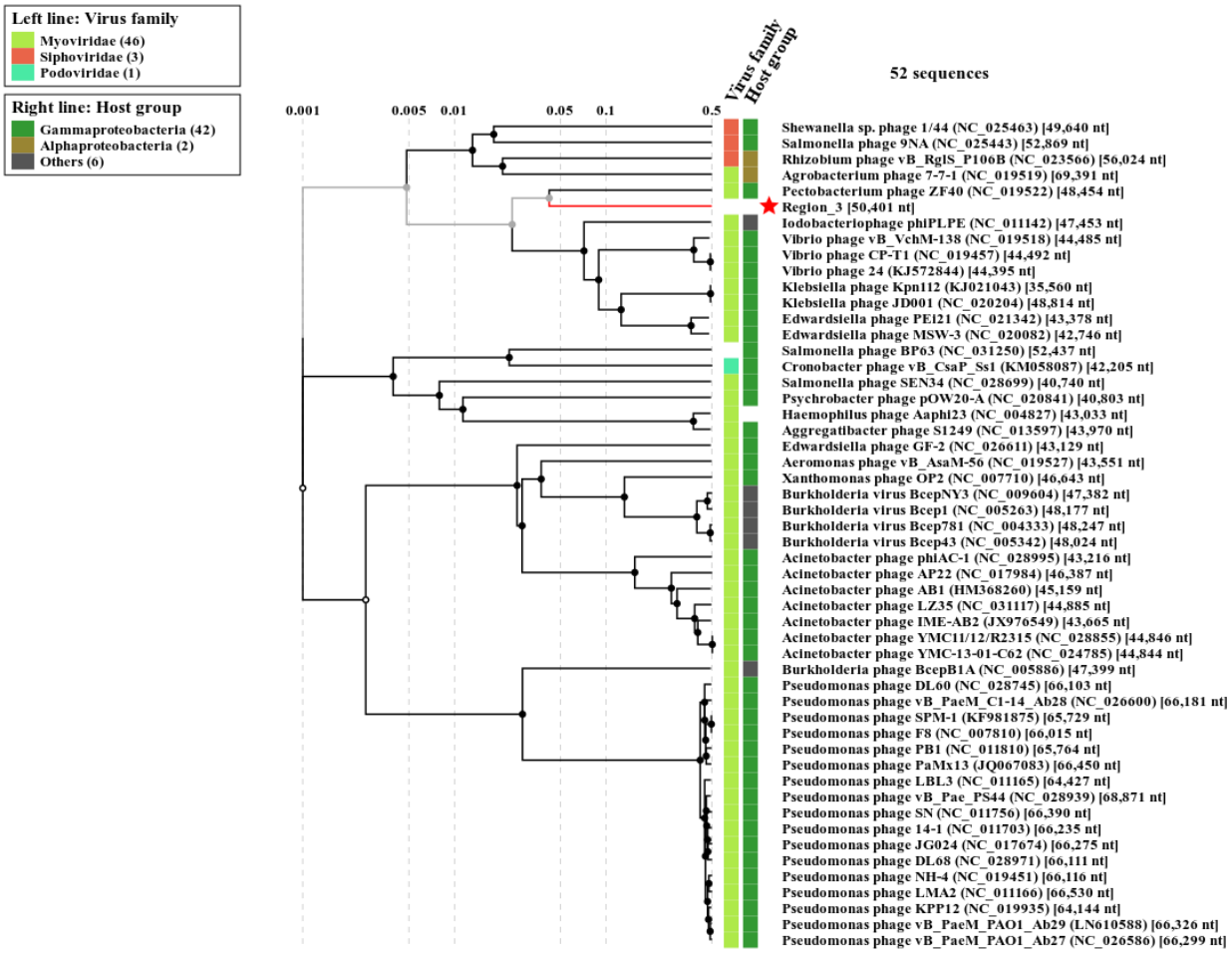
Supplementary Fig. S5.

Circular tree built automatically in Virfam with Aclame from the sequence of predicted prophage named Region 3. The different branches of the tree are sorted into 10 clusters, highlighted by different background colors. The prophage under study is in the gold yellow cluster, in a red box with white filling.



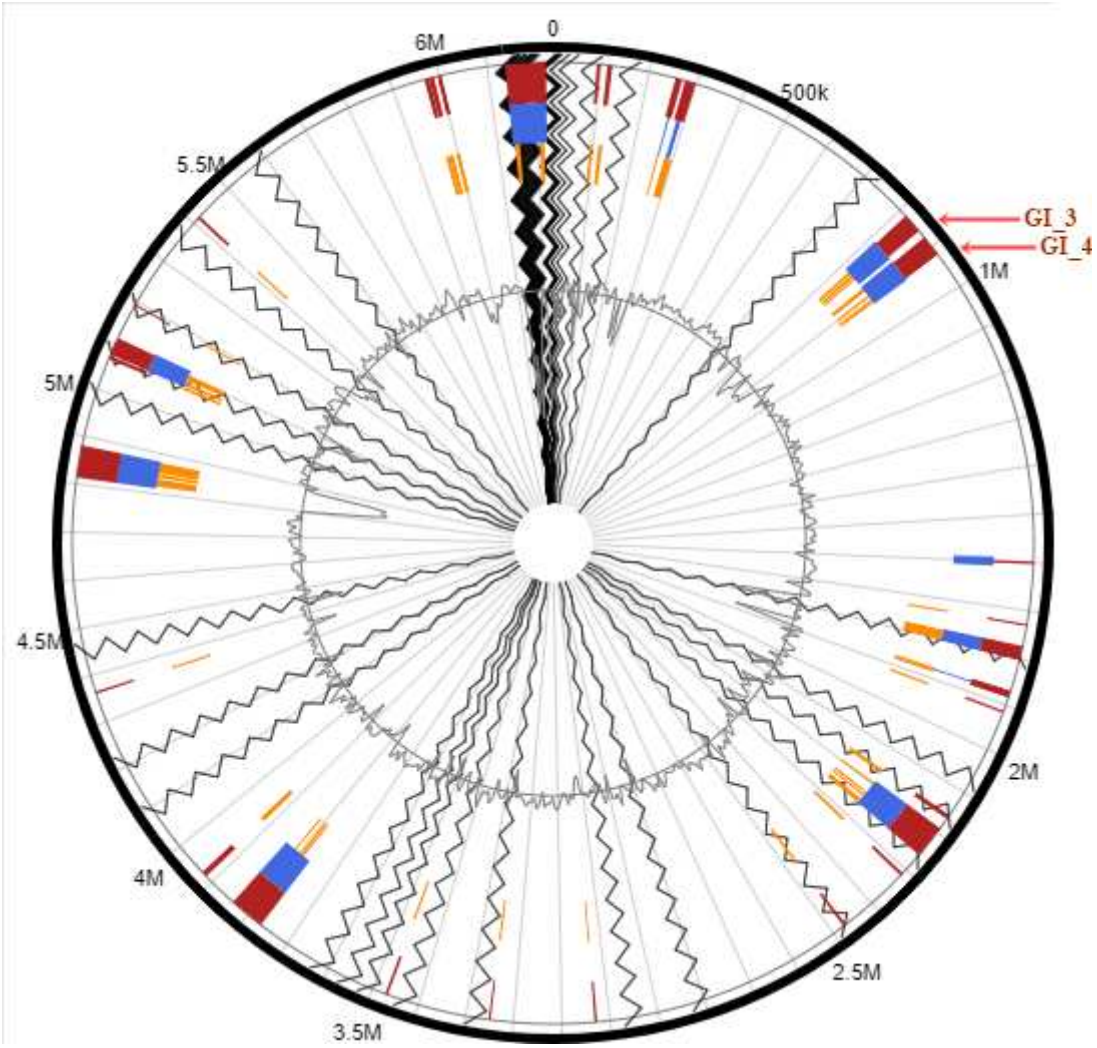
Supplementary Fig. S6.

The clade of the proteomic tree plotted by ViPTree including the Region 2 prophage (represented by red star) and other closest relative phages, based on genome-wide sequence similarities computed by tBLASTx.



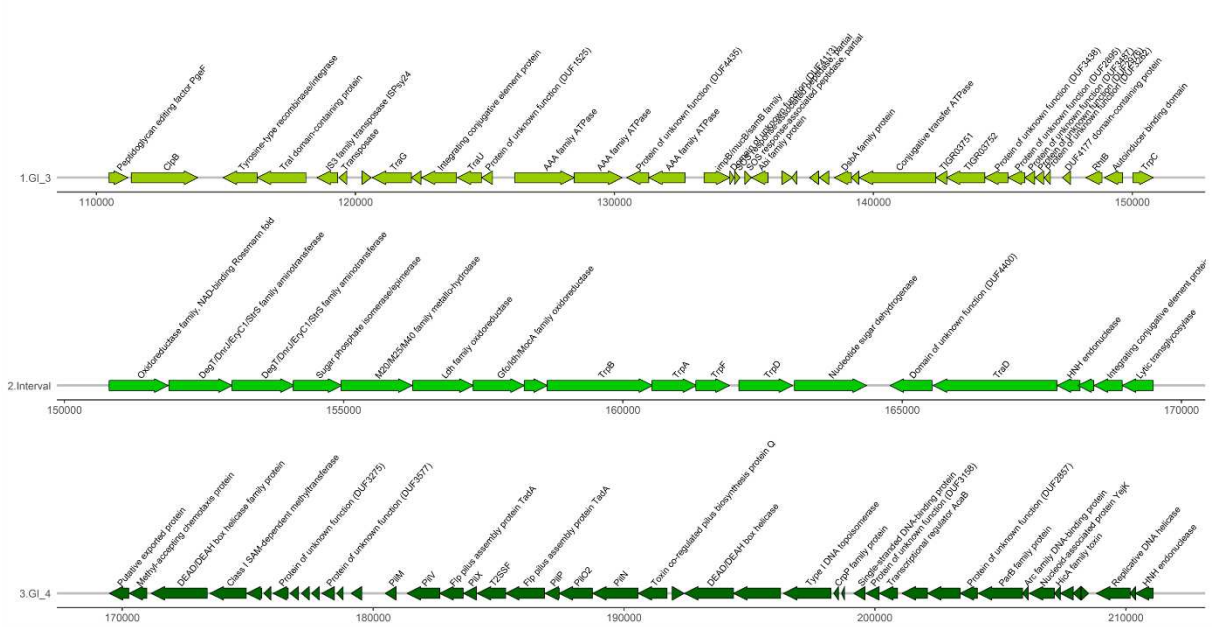
Supplementary Fig. S7.

The clade of the proteomic tree plotted by ViPTree including the Region 3 prophage (represented by red star) and other closest relative phages, based on genome-wide sequence similarities computed by tBLASTx.



Supplementary Fig. S8.

Predicted genomic islands (GIs) of *Pseudomonas paracarnis* RQ057 draft genome identified using IslandViewer. The yellow color displays predictions made from SIGI-HMM method, while the blue color displays predictions made IslandPath-DIMOB method. The red color displays the GIs predicted using at least one method. The two GIs that surround the exclusive region of blue discoloration are indicated (GI_3 and GI_4).



Supplementary Fig. S9.

Schematic representation of protein-coding genes organization in the genomics islands (GI_3 and GI_4) and in the interval between them (Interval), obtained from the *P. paracarnis* QR057 draft genome. The direction of the arrows represents the coding strand of the coding sequence. Name above the arrows correspond to annotation of genes.

Supplementary Table S1.

Intact prophage regions identified in the *Pseudomonas paracarnis* RQ057 draft genome by the PHASTER web server. For each region, the estimated completeness, score, specific phage-related keyword found in protein name, number of ORF, phage hit proteins, presence or absence of phage attachment site, percentage of GC, length, and most common phage are indicated.

Region	Completeness	Score [a]	Specific Keyword [b]	N ORF	Phage Hit Proteins	Attachment Sited [c]	GC %	Length (Kb)	Most Common Phage
1	Intact	140	Tail, plate, protease	27	21	No	59.34%	18.2	PHAGE_Vibrio_VP882_NC_009016(14)
2	Intact	150	Transposase, head, tail, capsid, terminase, lysis, integrase	70	47	yes	56.28%	54.6	PHAGE_Entero_HK022_NC_002166(8)
3	Intact	150	Integrase, portal, tail, lysin, head, capsid, plate, terminase	73	52	yes	57.03%	50.4	PHAGE_Iodoba_phiPLPE_NC_011142(14)

[a] Intact (score > 90); Questionable (score 70-90); Incomplete (score < 70)

[b] The specific phage-related keyword(s) found in protein name(s) in the region

ORF: Open Reading Frame

[c] The putative phage attachment site

Appendix S1

Taxonomic review of *Pseudomonas fluorescens* A506

The complete genome sequence of *P. fluorescens* A506 (CP003041) was uploaded to the Type (Strain) Genome Server (TYGS), and the result shows that this strain belongs to the species of *P. paracarnis* (Fig. 1). Furthermore, the calculation of dDDH using 14 genomes selected from the results provided by TYGS and *P. fluorescens* ATCC 13525^T and *P. paracarnis* RQ057 indicated that *P. fluorescens* A506 is related to the *P. paracarnis* V5/DAB/2/5^T (94.10%) and *P. paracarnis* RQ057 (92.70%), however it is distantly related to *P. fluorescens* ATCC 13525^T (32.30%). Same pattern was observed by calculating ANIb: *P. fluorescens* A506 is related to the *P. paracarnis* V5/DAB/2/5^T (99.09%) and *P. paracarnis* RQ057 (98.95%), but is distantly related to *P. fluorescens* ATCC 13525^T (85.84%). Therefore, by the genome-based taxonomy obtained from TYGS and the use of established cutoff for each algorithm (dDDH: 70%; ANIb: 95%), we can say that strain A506 belongs the *P. paracarnis* species.

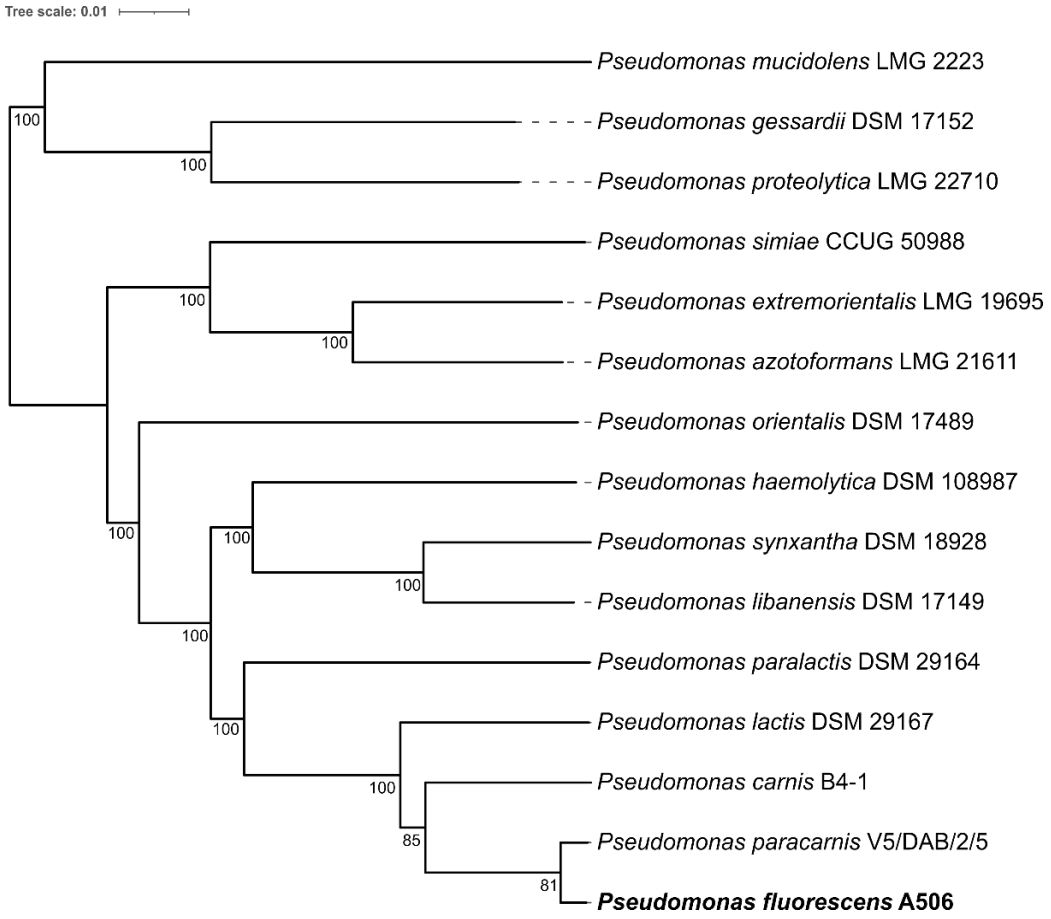


Fig. 1. Tree inferred with FastME 2.1.6.1 (Lefort et al., 2015) from GBDP distances calculated from genome sequences. The branch lengths are scaled in terms of GBDP distance formula d_5. The numbers above branches are GBDP pseudo-bootstrap support values > 60 % from 100 replications, with an average branch support of 97.2 %. The tree was rooted at the midpoint (Farris, 1972).

Table 1. Relatedness of the complete genome sequence of *P. fluorescens* A506 to whole genome sequences of *Pseudomonas* strains based on ANIb and dDDH.

Strains [a]	***Pseudomonas fluorescens* A506**	
	dDDH	**ANIb**
RQ057 (This study)	**92.70**	**98.95**
P. paracarnis V5/DAB/2/5^T (CAJFCN000000000)	**94.10**	**99.09**
P. fluorescens ATCC 13525^T (LT907842)	32.30	85.84
P. carnis B4-1^T (CABIVL000000000)	62.90	94.93
P. lactis DSM 29167^T (JYLO00000000)	60.00	94.57
P. paralactis DSM 29164^T (JYLN00000000)	41.70	90.08
P. libanensis DSM 17149^T (JYLH00000000)	39.90	89.35
P. haemolytica DSM 108987^T (VOIW00000000)	39.00	88.98
P. synxantha DSM 18928^T (JYLJ00000000)	38.20	88.63
P. orientalis DSM_17489^T (JYLM00000000)	34.20	86.92
P. extremorientalis LMG 19695^T (MDGK00000000)	32.90	86.13
P. azotoformans LMG 21611^T (LT629702)	32.90	86.26
P. simiae CCUG 50988^T (FOKB00000000)	32.30	85.98
P. gessardii DSM 17152^T (VFEW00000000)	29.40	83.85
P. proteolytica DSM 15321^T (VFEV00000000)	29.20	83.73
P. mucidolens LMG 2223^T (LT629802)	28.00	82.58

[a] In brackets, the accession numbers (GenBank) of the whole genome sequences used are indicated.

[b] Type strain of *P. proteolytica* is LMG 22710 = DSM 15321 = MTCC 4994 = CCUG 51515.

Reference

Farris, J. S. (1972). Estimating Phylogenetic Trees from Distance Matrices. *The American Naturalist*, *106*(951), 645–668. https://doi.org/10.1086/282802

Lefort, V., Desper, R., & Gascuel, O. (2015). FastME 2.0: A comprehensive, accurate, and fast distance-based phylogeny inference program. *Molecular Biology and Evolution*, *32*(10), 2798–2800. https://doi.org/10.1093/molbev/msv150

CHAPTER IV: Short communication: Draft genome sequence of *Pseudomonas carnis* strain isolated from blue discolored fresh cheese

Rafaela da Silva Rodrigues et al.

Manuscript prepared for submission to *Brazilian Journal of Microbiology*

Title page

Rafaela da Silva Rodrigues[1], Solimar Gonçalves Machado[1], Antônio Fernandes de Carvalho[1], Luís Augusto Nero[2]*

Short communication: Draft genome sequence of *Pseudomonas carnis* strain isolated from blue discolored fresh cheese

[1] Department of Food Technology, Federal University of Viçosa, Viçosa, MG, Brazil

[2] Department of Veterinary Medicine, Federal University of Viçosa, Viçosa, MG, Brazil

* Address correspondence to Luís Augusto Nero, nero@ufv.br, tel. +55 31 3612 5615

Abstract

Most *Pseudomonas* spp. are responsible for spoilage in refrigerated foods such as alteration in flavor, texture and appearance. Two samples of Minas Frescal cheese with blue discoloration were analysed and contained a high *Pseudomonas* concentration (7.72±0.36 log CFU/g). During our studies, from the 26 *Pseudomonas* isolates analysed, 19 were capable of producing diffusible dark pigment. So, a pigment-producing isolate (C020) was selected by rep-PCR fingerprinting and had the whole genome sequenced. The draft genome assembled comprises 42 contigs totaling 6,366,75 bp with an average G+C content of 59.97%, and the species prediction performed by TYGS, based on the draft genome sequence, identified the C020 as *Pseudomonas carnis*. In order to investigate the phylogenetic relationships of this isolate with strains already identified of this species, we performed an analysis based on complete genomic sequences. First, an analysis of all *P. carnis* genomes deposited in GenBank to date shows that 11% (4/37) are misidentified, and belong to the *Pseudomonas paracarnis* species. A comparative analysis based on the phylogenomic analysis showed that there is no evolutionary relationship between *P. carnis* strains carrying second copies of *trp* genes related to blue discoloration (*trpABCDF*), which reinforces the assertion that these genes are contained in a mobile element. But all of these have been isolated from foods. Therefore, possibly the current issues observed in different countries with pigment producers of this species were not caused by one spreading strain.

Keywords: Genomic analysis, Phylogeny, Spoilage, Bacterial pigment.

1. Introduction

Pseudomonas is a bacterial genus commonly found in raw or processed foods and is associated with economic losses due to texture, flavor, odor and appearance defects. It is heat-sensitive bacteria, however, the biofilm-forming ability of many members of this genus allow them to persist on surfaces commonly used in the food industry and thus contaminate products after processing [1]. In addition, many are psychotropic, which ensures their multiplication in refrigerated foods, reducing their shelf-life [2]. Psychotropic *Pseudomonas* can produce hydrolytic enzymes such as peptidases, lipases, and phospholipases, but also biogenic amines (such as cadaverine and putrescine), and pigments – which also impair the food quality [3–5].

Studies investigating the occurrence of dark pigments caused by *Pseudomonas* in food (reported as blue, gray or black pigments) have intensified in recent years mainly in uncultured cheeses and refrigerated fluid milk [6–9]. This spoilage has also been affecting meat products and, as a result, two species described and related to this phenotype were named *Pseudomonas carnis* and *Pseudomonas paracarnis*: *carnis* is a Latin noun, from meat, referring to the source of isolation of the species [10,11].

Nonetheless, ensuring a correct identification of *Pseudomonas* is a challenge, because strains belonging to the *Pseudomonas fluorescens* group, within *P. fluorescens* species complex, are very close genetically, even though they are taxonomically assigned to different species[12]. Therefore, sequencing several genes or even complete genomes facilitates identification, when correct tools are employed, including subsequent comparisons with complete sequences of type-strains [13].

In this study, we report the draft genome of *Pseudomonas carnis* isolated from blue discolored Minas Frescal cheese. In addition, we performed a comparative genomic analysis

among all genomes of this species deposited in the GenBank to verify the existence of an evolutionary relationship of blue pigment producers and if there is a clustering within this species of pigment producers.

2. Material and methods

2.1 Isolation of *Pseudomonas*

Two samples of Minas Frescal cheese with an apparent spoilage, blue discoloration, were analyzed in June 2020. *Pseudomonas* agar base (PAB; Oxoid Ltd., Basingstoke, UK) supplemented with CFC (cetrimide, fucidin and cephaloridine – Oxoid) and bromothymol blue (0.02 g L^{-1}) was used for isolation of *Pseudomonas*, with incubation at 25±1 °C for 48h. Then, purified isolates were tested for oxidase, catalase, and Gram staining. Longwave ultraviolet light was used to check the ability of isolates in producing fluorescence pigments [14]. The presumptive *Pseudomonas* isolates were stored at -80 °C in Brain Heart Infusion (BHI – Kasvi, São José dos Pinhais, PR, Brazil) with 20% v/v glycerol.

2.2 Blue pigment test

Presumptive isolates to genus *Pseudomonas* were tested for pigment production in plate count agar (PCA) and potato dextrose agar (PDA), at 25 °C, and also in 10% (w/v) reconstituted whole milk (w/v), at 10, 25 and 32 °C, incubated for 5 d.

2.3 Identification and typing

An identification by PCR amplification of the genus-specific region of 16S rRNA for *Pseudomonas* was performed using primers PA-GS-F (5'-GACGGGTGAGTAATGCCTA-3') and PA-GS-R (5'-CACTGGTGTTCCTTCCTATA-3')[15]. Then the isolates were analyzed by repetitive extragenic palindromic PCR (Rep-PCR) assay using single primer $(GTG)_5$ (5′-GTGGTGGTGGTGGTG-3′)[16]. BioNumerics version 6.6 (Applied Maths, Kortrijk, Belgium) was used to analyze the fingerprints, with Dice coefficient for similarity analysis and UPGMA (5% optimization and 1.5% tolerance) to create a dendrogram. The DNA fingerprint patterns were compared using a cutoff value of 90% similarity.

2.4 Whole-genome sequencing and bioinformatic analyzes

One blue pigment-producing isolate identified as C020 was selected for whole-genome sequencing, performed by Neoprospecta Microbiome Technologies (Florianopolis, SC, Brazil) using the Illumina Miseq platform. The programs FastQC v. 0.11.9 [17], Trimmomatic v.0.39 [18], and Mira Assembler v.4.9.6 [19] were used for check, trimming and assembly, respectively. Contigs less than 500 bp in length were excluded and ContEst16S (Contamination Estimator by 16S) algorithm was utilized to check for potential contamination in the draft genome sequence [20]. QUAST v.5.0.2 tool was used to assess the quality of assembled draft genome and genes were annotated with Prokka v.1.14.5 [21] software with default settings. Finally, the genome sequence data were uploaded to the Type (Strain) Genome Server (TYGS), a free bioinformatics platform available under https://tygs.dsmz.de, for a whole genome-based taxonomic analysis [22]. The circular representation of the draft genome was constructed using Proksee (https://proksee.ca/).

2.5 Whole-genome phylogenies

All genomes of *P. carnis* were retrieved from the NCBI GenBank (https://www.ncbi.nlm.nih.gov/data-hub/genome/) and checked for species using average nucleotide identity (ANI) based on BLASTN algorithm (ANIb) calculated with the web service (https://jspecies.ribohost.com/jspeciesws/). When the species was misidentified, genomes were submitted to TYGS for a whole genome-based taxonomic analysis, and ANIb was repeated with the type strain of the identified species.

A phylogenomic tree was inferred with the *P. carnis* genomes based on Genome BLAST Distance Phylogeny method (GBDP) in the TYGS platform (https://tygs.dsmz.de/) [22]. The iTOL web-based software v. 6 (https://itol.embl.de/) was used to visualize phylogenetic tree. Genomes were also analysed for the presence of second copies of *trp* genes (*trpABCDF)*, identified in a region exclusive in the genome of pigmenting strains. A comparison was performed with predicted proteins obtained from *P. carnis* B4-1^{T} (unpublished data). For this, all genomes were annotated with Prokka v.1.14.5 [21] and search was performed using the Basic Local Alignment Search Tool (BLAST) program for protein (BLASTp) available from the National Center for Biotechnology Information (NCBI) (https://blast.ncbi.nlm.nih.gov/Blast.cgi).

2.6 Data availability

This Whole Genome Shotgun project has been deposited at DDBJ/ENA/GenBank under the accession JAPJDQ000000000. The version described in this paper is version JAPJDQ010000000.

3. Results and discussion

3.1 Isolation and identification of *Pseudomonas*

The spoiled samples of Minas Frescal cheese (blue discoloration) contained a *Pseudomonas* spp. concentration of approximately 7.72±0.36 log CFU/g, from which we obtained 26 presumptive isolates for this genus. PCR amplification of the genus-specific region of 16S rRNA confirmed all isolates as *Pseudomonas*, of which 21 produced fluorescence under ultraviolet light. In addition, 19 fluorescent isolates were capable of producing diffusible dark pigment on PCA and PDA after 1d incubation at 25 °C, and on reconstituted whole milk (10%, w/v) at 10 and 25 °C, after 4d and 1d, respectively – but pigment did not arise at 32 °C. Furthermore, pigmenting isolates were clustered together by the cluster analysis of rep-PCR (GTG_5) fingerprints, which showed 2 different clusters with similarity indices of at least 90% and 3 solitary fingerprints (Supplementary Figure S1). Consequently, a pigment-producing isolate identified as C020 was randomly selected for whole-genome sequencing.

The draft genome consists of 42 contigs totaling 6,366,753 bp with an N_{50} value of 494,884 and average G+C content of 59.97% (Table 1). A total of 5,714 predicted CDSs, 65 tRNAs, 11 rRNAs, and1 tmRNA were identified after annotation. Figure 1 shows the genome map generated in Proksee Server.

Table 1. Genome assembly metrics.

Feature	Value
Contigs	42
Total length (bp)	6,366,753
Largest contig (bp)	746,652
N_{50}	494,884
N_{75}	222,898
GC%	59.97
N's	7
CDS	5,714
Gene	5,791
rRNA	11
tRNA	65
tmRNA	1

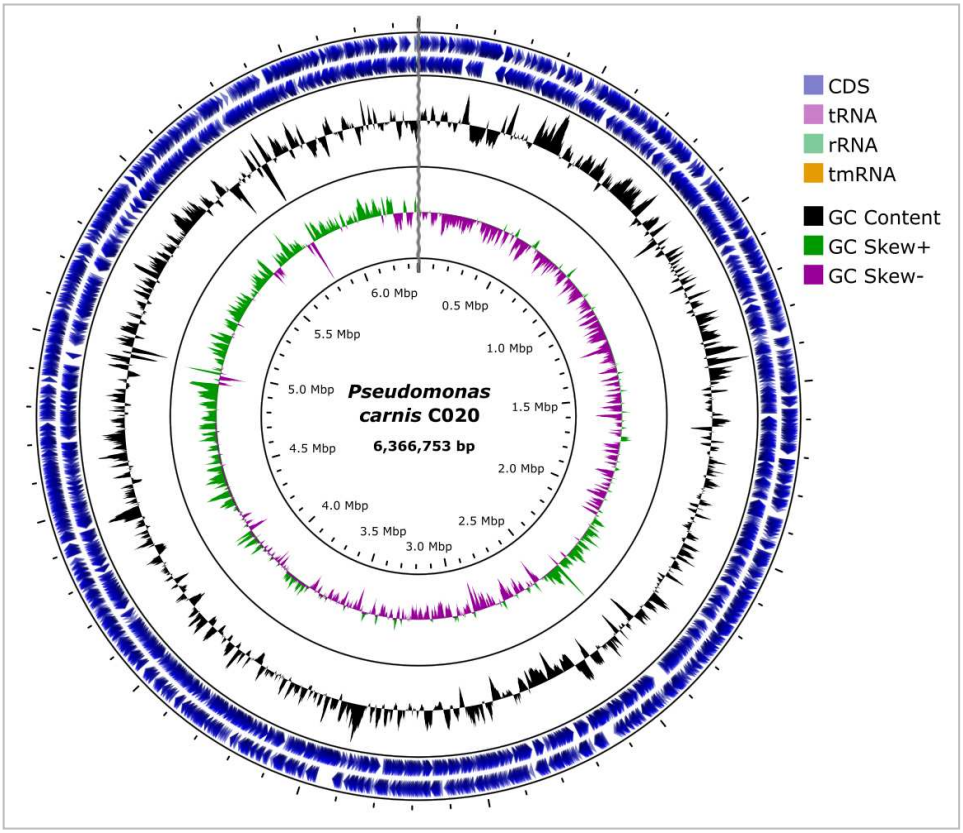


Figure 1. Graphical circular maps of the draft genomes of *Pseudomonas carnis* C0020, performed with Proksee server (https://proksee.ca/). The circles indicate contigs concatenated. From outside to center: ring 1 and 2 display genes on both the forward and reverse strand; ring 3 displays G + C content (deviation from the average); and ring 4 displays G + C skew.

The species prediction based on the draft genome sequence, performed in the TYGS web server, identified the strain as *Pseudomonas carnis*. In pairwise comparisons of genome versus type-strain genomes, the draft genome of C020 showed 87.0 % of dDDH (> 70% value) with *Pseudomonas carnis* B4-1^T by formula d_4, – a robust formula against the use of incomplete draft genomes, and also known as GGDC formula 2 [22]. The description of this species was published in 2020 [10], isolated from meat. However, it has also been identified in other food products such as cheese [23], soymilk and tofu [24]. Figure 2 shows the phylogenetic tree generated with the closest-type strains genomes.

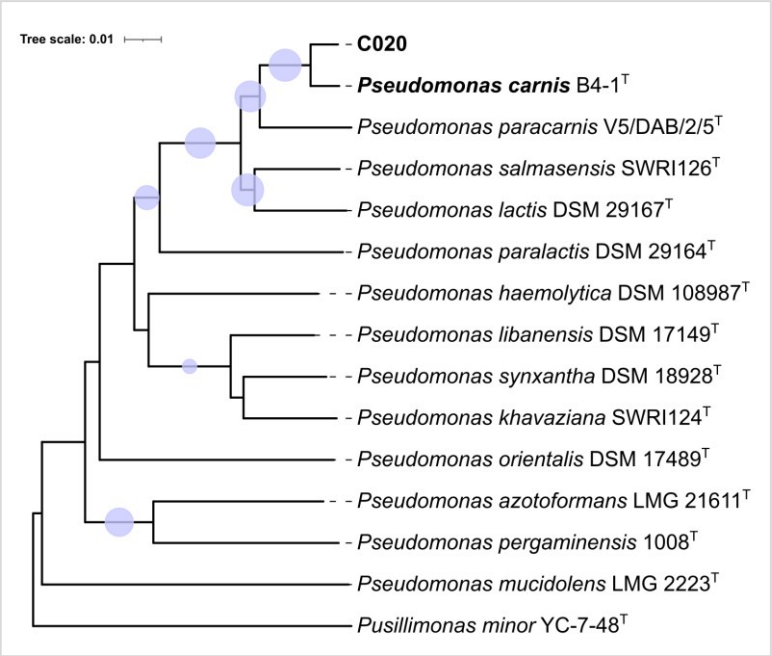


Figure 2. Tree inferred with FastME 2.1.6.1 [25] from GBDP distances calculated from genome sequences. The branch lengths are scaled in terms of GBDP distance formula d_5. The symbols on the branches represent GBDP pseudo-bootstrap support values > 60 % from 100 replications, with an average branch support of 64.3 %. The tree was rooted at the midpoint [26].

3.2 Whole-genome phylogenies

We retrieved 37 genomes scientifically named as belonging *Pseudomonas carnis* from NCBI GenBank (www.ncbi.nlm.nih.gov/data-hub/genome/) (Supplementary Table S1). However, 4 of 37 genomes deposited in GenBank were misidentified at the species level according to the ANIb results, below the established cutoff (<95%), and also by pairwise comparisons using TYGS server with type-strain genomes (Supplementary table S2, S3). The dDDH from analysis by TYGS constated they are *Pseudomonas paracarnis* (>70% value), and ANIb also demonstrated it (Supplementary table S3). Unfortunately, misidentifications like these could be influence subsequent researches and also spread to other databases. Strains of *P. paracarnis* have also been identified and associated with food spoilage by discoloration [11]. It is a species recently described, isolated from meat (refrigerated dry aged beef: roast beef) and is phylogenetically close to *P. carnis* [11], which could affect identification when using less robust methods.

Thus, 34 genomes (33 from GenBank and C020) were analyzed by generation of phylogenomic tree, with *Pseudomonas paracarnis* V5/DAB/2/5^{T} as outgroup (Figure 3). This phylogenetic evaluation inferred with genome sequences showed that there was neither clustering nor evolutionary relationships for strains containing second copies of *trp* genes related to blue discoloration (Figure 3). The genes, involved in the tryptophan biosynthesis pathway (*trpABCDF*), have been investigated in genomes of strains related to blue discoloration of dairy products [6,8,9]. In addition to containing these marker genes, all were previously confirmed to produce a dark pigment by phenotypic testing [10,23,24,27]. However, all identified here as carriers of these genes were isolated from food (Figure 3).

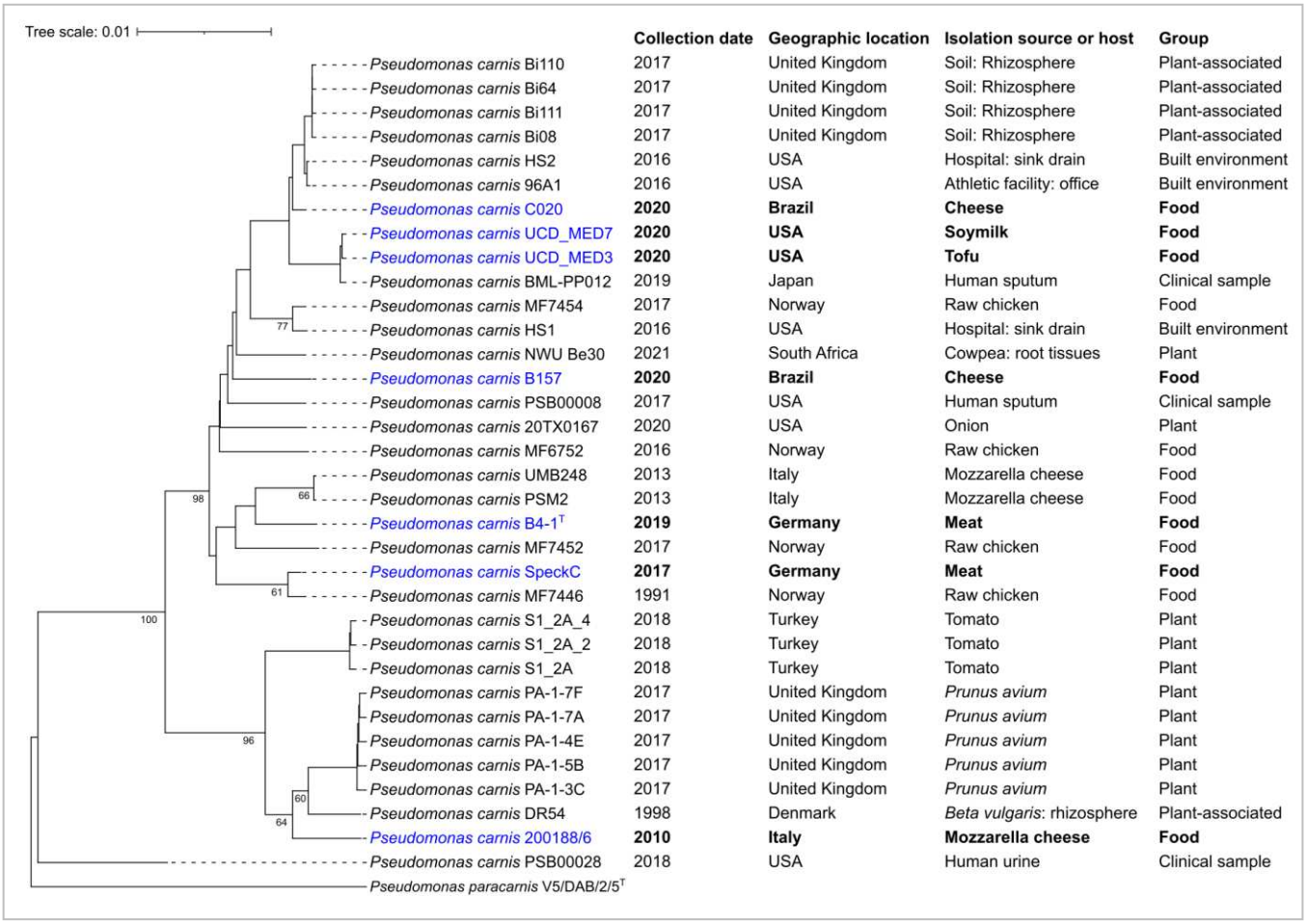


Figure 3. Tree inferred with FastME 2.1.6.1[25] from GBDP distances calculated from genome sequences. The branch lengths are scaled in terms of GBDP distance formula d_5. The numbers above branches are GBDP pseudo-bootstrap support values > 60 % from 100 replications, with an average branch support of 37.3 %. The tree was rooted at the midpoint [26]. *Pseudomonas paracarnis* V5/DAB/2/5^T (GenBank: CAJFCN000000000.1) was used as outgroup. Highlighted strains contain second copies of *trp* genes related to blue discoloration (*trpABCDF*).

So far, the investigation of the presence of *trpABCDF* has been studied only in strains isolated from food, as it leads to commercial losses. However, it is not a species-dependent characteristic and its absence in strains from other sources (soil, plants), until now, had not been observed. Furthermore, there is no obvious evolutionary relationship between carriers of specific genes. These two observations suggest that at some point there was or is a horizontal transfer of these genes to some members of this species (as well as to others, such as *P. paracarnis*). But it is still not possible to justify that a pigment-producing strain of *P. carnis*

entered a specific food processing plant, which allowed its spread to others. However, there are other factors that allow these isolates to be less detected in other environments and rather to be highlighted in food production: the cold storage allows most *Pseudomonas* species to become one of the most predominant genera in the microbiota of many foods [28–31]; as well as being better biofilm formers, which facilitates their permanence in these environments compared to other organisms, mainly these deteriorators, as it has already been demonstrated that pigment-producing *Pseudomonas* form stronger biofilms than non-pigment producing ones [1].

4. Conclusion

Here we announce the draft genome of a dark pigment producer that also carries duplicated genes for the tryptophan biosynthesis pathway (*trpABCDF*). Moreover, the present manuscript highlights the importance of the correct identification of *Pseudomonas* isolates. This is still a challenge, because even with the complete genomic sequence there are misidentifications, which could allow the propagation of information that affects studies. In addition, by genomic analysis, we realized that there is no evolutionary relationship between isolates obtained in different countries that were responsible for the blue discoloration of foods, therefore possibly these spoilages have not been caused by the recent spread of a single strain.

References

1. Rossi C, Serio A, Chaves-López C, et al. Biofilm formation, pigment production and motility in *Pseudomonas* spp. isolated from the dairy industry. *Food Control*.

2018;86:241-248. doi:10.1016/j.foodcont.2017.11.018

2. Cenci-Goga BT, Karama M, Sechi P, Iulietto MF, Novelli S, Mattei S. Evolution under different storage conditions of anomalous blue coloration of Mozzarella cheese intentionally contaminated with a pigment-producing strain of *Pseudomonas fluorescens*. *J Dairy Sci*. 2014;97:6708-6718. doi:10.3168/jds.2014-8611

3. Machado SG, Baglinière F, Marchand S, et al. The biodiversity of the microbiota producing heat-resistant enzymes responsible for spoilage in processed bovine milk and dairy products. *Front Microbiol*. 2017;8(302):1-22. doi:10.3389/fmicb.2017.00302

4. del Olmo A, Calzada J, Nuñez M. The blue discoloration of fresh cheeses: A worldwide defect associated to specific contamination by *Pseudomonas fluorescens*. *Food Control*. 2018;86:359-366. doi:10.1016/j.foodcont.2017.12.001

5. Scatamburlo TM, Yamazi AK, Cavicchioli VQ, Pieri FA, Nero LA. Spoilage potential of *Pseudomonas* species isolated from goat milk. *J Dairy Sci*. 2015;98:759-764. doi:10.3168/jds.2014-8747

6. Andreani NA, Carraro L, Martino ME, et al. A genomic and transcriptomic approach to investigate the blue pigment phenotype in *Pseudomonas fluorescens*. *Int J Food Microbiol*. 2015;213:88-98. doi:10.1016/j.ijfoodmicro.2015.05.024

7. Andreani NA, Martino ME, Fasolato L, et al. Tracking the blue: A MLST approach to characterise the *Pseudomonas fluorescens* group. *Food Microbiol*. 2014;39:116-126. doi:10.1016/j.fm.2014.11.009

8. Andreani NA, Carraro L, Zhang L, Vos M, Cardazzo B. Transposon mutagenesis in *Pseudomonas fluorescens* reveals genes involved in blue pigment production and antioxidant protection. *Food Microbiol*. 2019;82:497-503. doi:10.1016/j.fm.2019.03.028

9. Reichler SJ, Martin NH, Evanowski RL, Kovac J, Wiedmann M, Orsi RH. A century of gray: A genomic locus found in 2 distinct *Pseudomonas* spp. is associated with historical and contemporary color defects in dairy products worldwide. *J Dairy Sci*. 2019;102:5979-6000. doi:10.3168/jds.2018-16192

10. Lick S, Kröckel L, Wibberg D, et al. *Pseudomonas carnis* sp. nov., isolated from meat. *Int J Syst Evol Microbiol*. 2020. doi:10.1099/ijsem.0.003928

11. Lick S, Wibberg D, Winkler A, et al. *Pseudomonas paracarnis* sp. nov., isolated from refrigerated beef. *Int J Syst Evol Microbiol*. 2021;71:004652. doi:10.1099/ijsem.0.004652

12. Garrido-Sanz D, Meier-Kolthoff JP, Göker M, Martín M, Rivilla R, Redondo-Nieto M. Genomic and genetic diversity within the *Pseudomonas fluorescens* complex. *PLoS One*. 2016;11(2). doi:10.1371/journal.pone.0150183

13. Gomila M, Peña A, Mulet M, Lalucat J, García-Valdés E. Phylogenomics and systematics in *Pseudomonas*. *Front Microbiol*. 2015;6(Article 214). doi:10.3389/fmicb.2015.00214

14. Martin NH, Murphy SC, Ralyea RD, Wiedmann M, Boor KJ. When cheese gets the blues: *Pseudomonas fluorescens* as the causative agent of cheese spoilage. *J Dairy Sci*. 2011;94:3176-3183. doi:10.3168/jds.2011-4312

15. Spilker T, Coenye T, Vandamme P, LiPuma JJ. PCR-Based Assay for Differentiation of *Pseudomonas aeruginosa* from Other *Pseudomonas* Species Recovered from Cystic Fibrosis Patients. *J Clin Microbiol*. 2004;45(5):2074-2079. doi:10.1128/JCM.42.5.2074-2079.2004

16. Versalovic J, Schneider M, De Bruijn FJ, Lupski JR. Genomic fingerprinting of bacteria using repetitive sequence-based polymerase chain reaction. *Methods Mol Cell Biol*.

1994;5:25-40.

17. Andrews S. FastQC - A quality control tool for high throughput sequence data. http://www.bioinformatics.babraham.ac.uk/projects/fastqc/. *Babraham Bioinforma*. 2010.

18. Bolger AM, Lohse M, Usadel B. Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics*. 2014;30(15):2114-2120. doi:10.1093/bioinformatics/btu170

19. Chevreux B, Wetter T, Suhai S. Genome Sequence Assembly Using Trace Signals and Additional Sequence Information. *Comput Sci Biol Proc Ger Conf Bioinforma '99, GCB, Hann Ger*. 1999;99:45-56.

20. Lee I, Chalita M, Ha SM, Na SI, Yoon SH, Chun J. ContEst16S: An algorithm that identifies contaminated prokaryotic genomes using 16S RNA gene sequences. *Int J Syst Evol Microbiol*. 2017;67:2053-2057. doi:10.1099/ijsem.0.001872

21. Seemann T. Prokka: Rapid prokaryotic genome annotation. *Bioinformatics*. 2014;30(14):2068-2069. doi:10.1093/bioinformatics/btu153

22. Meier-Kolthoff JP, Göker M. TYGS is an automated high-throughput platform for state-of-the-art genome-based taxonomy. *Nat Commun*. 2019;10:2182. doi:10.1038/s41467-019-10210-3

23. Rodrigues R da S, Machado SG, de Carvalho AF, Nero LA. *Pseudomonas* sp. as the causative agent of anomalous blue discoloration in Brazilian fresh soft cheese (Minas Frescal). *Int Dairy J*. 2021;117:105020. doi:10.1016/j.idairyj.2021.105020

24. De León ME, Wilson HS, Jospin G, Eisen JA. Draft genome sequences and genomic analysis for pigment production in bacteria isolated from blue discolored Soymilk and Tofu. *J Genomics*. 2021;9:55-67. doi:10.7150/jgen.65500

25. Lefort V, Desper R, Gascuel O. FastME 2.0: A comprehensive, accurate, and fast distance-based phylogeny inference program. *Mol Biol Evol*. 2015;32(10):2798-2800. doi:10.1093/molbev/msv150

26. Farris JS. Estimating Phylogenetic Trees from Distance Matrices. *Am Nat*. 1972;106(951):645-668. doi:10.1086/282802

27. Chierici M, Picozzi C, La Spina MG, et al. Strain diversity of *Pseudomonas fluorescens* group with potential blue pigment phenotype isolated from dairy products. *J Food Prot*. 2016;79(8):1430-1435. doi:10.4315/0362-028X.JFP-15-589

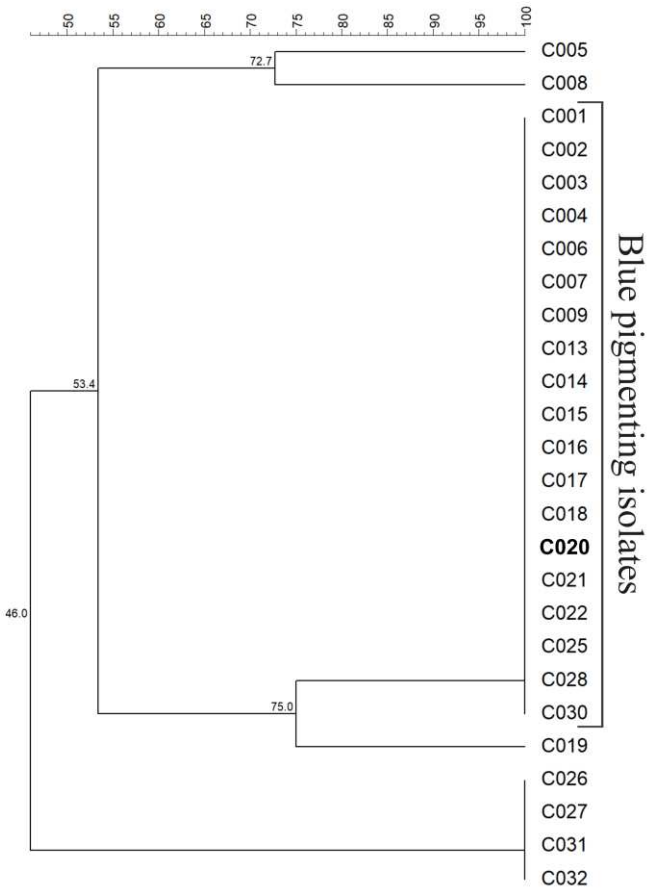
28. von Neubeck M, Baur C, Krewinkel M, et al. Biodiversity of refrigerated raw milk microbiota and their enzymatic spoilage potential. *Int J Food Microbiol*. 2015. doi:10.1016/j.ijfoodmicro.2015.07.001

29. Hahne J, Isele D, Berning J, Lipski A. The contribution of fast growing, psychrotrophic microorganisms on biodiversity of refrigerated raw cow's milk with high bacterial counts and their food spoilage potential. *Food Microbiol*. 2019;79:11-19. doi:10.1016/j.fm.2018.10.019

30. Lee HS, Kwon M, Heo S, Kim MG, Kim G-B. Characterization of the Biodiversity of the Spoilage Microbiota in Chicken Meat Using Next Generation Sequencing and Culture Dependent Approach. *Korean J Food Sci Anim Resour*. 2017;37(4):535-541. doi:10.5851/kosfa.2017.37.4.535

31. Coton M, Joffraud JJ, Mekhtiche L, Leroi F, Coton E. Biodiversity and dynamics of the bacterial community of packaged king scallop (Pecten maximus) meat during cold storage. *Food Microbiol*. 2013;35(2):99-107. doi:10.1016/j.fm.2013.02.011

Supplementary material



Supplementary Figure S1. Dendrogram obtained by cluster analysis of rep-PCR (GTG_5) fingerprints, constructed by BioNumerics 6.6 software. Similarity analysis was performed using the Dice coefficient, and clustering was by the UPGMA (optimization 5%, tolerance 1.5%). Node values represent percentages of similarity.

Supplementary Table S1. Description obtained from NCBI's BioSample database of all genomes named as *Pseudomonas carnis*.

Organism Scientific Name	BioSample	WGS project	Level	Size (bp)	Submission Date	Geographic location
Pseudomonas carnis HS2	SAMN14085467	JAAMRI01	Contig	6516451	23/07/2020	USA: Chicago
Pseudomonas carnis HS1	SAMN14085466	JAAMRH01	Contig	6519674	23/07/2020	USA: Chicago
Pseudomonas carnis 96A1	SAMN14085459	JAAMRA01	Contig	6350780	23/07/2020	USA:Springfield
Pseudomonas carnis PSB00028	SAMN12220705	JADUCE01	Contig	6329332	13/12/2020	USA: Pittsburgh, Pennsylvania
Pseudomonas carnis PSB00008	SAMN10436423	JADTYU01	Contig	5916413	13/12/2020	USA: Pittsburgh, Pennsylvania
Pseudomonas carnis S1_2A_2	SAMN16885832	JAEIKR01	Scaffold	5957443	22/12/2020	Turkey: Mersin
Pseudomonas carnis S1_2A_4	SAMN16885833	JAEIKS01	Scaffold	5968881	22/12/2020	Turkey: Mersin
Pseudomonas carnis S1_2A	SAMN16885831	JAEIKQ01	Scaffold	5987354	22/12/2020	Turkey: Mersin
Pseudomonas carnis MF7446	SAMN17076052	JAEKCR01	Contig	6162157	26/12/2020	Norway
Pseudomonas carnis MF7452	SAMN17076057	JAEKCO01	Contig	6430424	26/12/2020	Norway
Pseudomonas carnis MF7454	SAMN17076059	JAEKCM01	Contig	6808852	26/12/2020	Norway
Pseudomonas carnis MF6752	SAMN17076037	JAENSU01	Contig	6100244	11/01/2021	Norway
Pseudomonas carnis UMB248	SAMN05188743	LYXI01	Contig	5474606	26/02/2021	Italy: Milano
Pseudomonas carnis 200188/6	SAMN03268414	JXLJ01	Contig	5719207	11/05/2021	Italy: Brescia
Pseudomonas carnis PSM2	SAMN03284985	JXMI01	Contig	5721670	11/05/2021	Italy: Milano
Pseudomonas carnis UCD_MED7	SAMN18848723	JAGTYE01	Contig	6311992	02/07/2021	USA: Davis
Pseudomonas carnis UCD_MED3	SAMN18848722	JAGTYD01	Contig	6316412	02/07/2021	USA: Davis
Pseudomonas carnis B157	SAMN16988608	JAEFBE01	Contig	6567469	11/08/2021	Brazil: Minas Gerais
Pseudomonas carnis DR54	SAMN18094175	JAFLXE01	Scaffold	6084452	07/09/2021	Denmark: Holeby
Pseudomonas carnis BML-PP035	SAMD00422628	BQIC01	Contig	6066608	22/11/2021	Japan:Saitama, Kawagoe
Pseudomonas carnis BML-PP038	SAMD00422631	BQIF01	Contig	6082891	22/11/2021	Japan:Saitama, Kawagoe
Pseudomonas carnis BML-PP010	SAMD00422604	BQHE01	Contig	6197743	22/11/2021	Japan:Saitama, Kawagoe
Pseudomonas carnis BML-PP012	SAMD00422606	BQHG01	Contig	6357705	22/11/2021	Japan:Saitama, Kawagoe
Pseudomonas carnis BML-PP016	SAMD00422610	BQHK01	Contig	6501702	22/11/2021	Japan:Saitama, Kawagoe
Pseudomonas carnis PA-1-5B	SAMN13195904	WJZK01	Contig	5989441	02/02/2022	United Kingdom: Kent
Pseudomonas carnis PA-1-4E	SAMN13195900	WJZG01	Contig	5955252	02/02/2022	United Kingdom: Kent
Pseudomonas carnis PA-1-7F	SAMN13195910	WJZQ01	Contig	6018355	02/02/2022	United Kingdom: Kent
Pseudomonas carnis PA-1-7A	SAMN13195908	WJZO01	Contig	6013967	02/02/2022	United Kingdom: Kent
Pseudomonas carnis PA-1-3C	SAMN13195896	WJZC01	Contig	5963276	02/02/2022	United Kingdom: Kent
Pseudomonas carnis NWU Be30	SAMN24505123	JAMKPY01	Scaffold	5901107	30/06/2022	South Africa
Pseudomonas carnis 20TX0167	SAMN30285734	JANQAQ01	Scaffold	6009690	20/08/2022	USA
Pseudomonas carnis SpeckC	SAMEA5093986	UYJC01	Contig	6217594	19/12/2018	Germany
Pseudomonas carnis B4-1^{T}	SAMEA5093985	CABIVL01	Contig	6194570	07/08/2019	Germany
Pseudomonas carnis Bi110	SAMEA7892418	CAKKLO01	Contig	6696265	20/11/2021	United Kingdom
Pseudomonas carnis Bi64	SAMEA7892352	CAKKMP01	Contig	6761930	20/11/2021	United Kingdom
Pseudomonas carnis Bi111	SAMEA7892419	CAKKME01	Contig	6695877	20/11/2021	United Kingdom
Pseudomonas carnis Bi08	SAMEA7892338	CAKKMA01	Contig	6697465	20/11/2021	United Kingdom

Supplementary Table S2. ANIb (by JSpeciesWS) results of *Pseudomonas carnis* genomes compared to *Pseudomonas carnis* type strain (B4-1^T).

Genomes	ANIb [%]
Pseudomonas carnis PSM2	**98.69**
Pseudomonas carnis UMB248	**98.46**
Pseudomonas carnis SpeckC	**98.41**
Pseudomonas carnis MF7446	**98.36**
Pseudomonas carnis C020	**98.30**
Pseudomonas carnis 20TX0167	**98.30**
Pseudomonas carnis PSB00008	**98.30**
Pseudomonas carnis B157	**98.28**
Pseudomonas carnis MF7452	**98.28**
Pseudomonas carnis NWU_Be30	**98.24**
Pseudomonas carnis HS1	**98.20**
Pseudomonas carnis Bi08	**98.18**
Pseudomonas carnis Bi111	**98.18**
Pseudomonas carnis MF7454	**98.18**
Pseudomonas carnis Bi110	**98.17**
Pseudomonas carnis Bi64	**98.17**
Pseudomonas carnis MF6752	**98.14**
Pseudomonas carnis UCD_MED3	**98.11**
Pseudomonas carnis UCD_MED7	**98.10**
Pseudomonas carnis 96A1	**98.08**
Pseudomonas carnis HS2	**98.04**
Pseudomonas carnis BML-PP012	**97.95**
Pseudomonas carnis S1_2A	**97.30**
Pseudomonas carnis S1_2A_4	**97.29**
Pseudomonas carnis S1_2A_2	**97.28**
Pseudomonas carnis PA-1-7F	**97.18**
Pseudomonas carnis PA-1-5B	**97.16**
Pseudomonas carnis PA-1-7A	**97.16**
Pseudomonas carnis PA-1-3C	**97.14**
Pseudomonas carnis PA-1-4E	**97.14**
Pseudomonas carnis DR54	**97.02**
Pseudomonas carnis 200188/6	**97.00**
Pseudomonas carnis PSB00028	**96.62**

Supplementary Table S3. ANIb (by JSpeciesWS) and dDDH (by TYGS web server) results of misidentified *Pseudomonas carnis* compared to *Pseudomonas carnis* and *Pseudomonas paracarnis* type strains.

Genomes	***P. carnis* B4-1 ^{T}**		***P. paracarnis* V5/DAB/2/5^{T}**	
	ANIb [%]	**dDDH [%] ***	**ANIb [%]**	**dDDH [%] ***
Pseudomonas carnis BML-PP038	94.73	63.1	**98.82**	**92.7**
Pseudomonas carnis BML-PP010	94.70	62.8	**98.73**	**93.0**
Pseudomonas carnis BML-PP035	94.72	62.9	**98.66**	**92.4**
Pseudomonas carnis BML-PP016	94.79	62.8	**98.56**	**91.9**

*Formula d_4 (also known as GGDC formula 2) by TYGS.

GENERAL CONCLUSIONS

- ✓ The blue discoloration of cheeses in Brazil is caused by *Pseudomonas*, which produce a blue pigment mainly at refrigeration temperatures and at 25 °C.
- ✓ Multi-locus sequence typing (MLST) of the *P. fluorescens* group showed that blue pigment producers isolated from Brazil may also share an evolutionary development with isolates from other countries.
- ✓ A exclusive genomic region in the genome of these spoilers contains not only tryptophan biosynthesis pathway genes, but a total of 74 genes, including genes related to horizontal transfer.
- ✓ The blue discoloration of Minas Frescal cheese and processed cheese "requeijão em barra" was caused by strains of recently described species, *Pseudomonas carnis* and *Pseudomonas paracarnis*, which are also related to the spoilage of meat products.
- ✓ There is no evolutionary relationship between the pigment strains within the *P. carnis* species deposited in the NCBI public databases. However, all *P. carnis* pigment strains identified to date have been isolated from food.
- ✓ From this study, the best assembled genome, for use in further studies such as the development of markers for the exclusive genomic region, was that of the *P. paracarnis* strain RQ057, because this whole exclusive region is contained in one contig.

ATTACHMENTS - Licenses to use the content of published articles

CCC Marketplace

This is a License Agreement between Rafaela da Silva Rodrigues ("User") and Copyright Clearance Center, Inc. ("CCC") on behalf of the Rightsholder identified in the order details below. The license consists of the order details, the Marketplace Permissions General Terms and Conditions below, and any Rightsholder Terms and Conditions which are included below.

All payments must be made in full to CCC in accordance with the Marketplace Permissions General Terms and Conditions below.

Order Date	17-Feb-2023	Type of Use	Republish in a thesis/dissertation
Order License ID	1325369-1	Publisher	ELSEVIER BV
ISSN	0958-6946	Portion	Chapter/article

LICENSED CONTENT

Publication Title	International dairy journal	Country	Netherlands
Article Title	Pseudomonas sp. as the causative agent of anomalous blue discoloration in Brazilian fresh soft cheese (Minas Frescal)	Rightsholder	Elsevier Science & Technology Journals
		Publication Type	Journal
		Start Page	105020
Author/Editor	INTERNATIONAL DAIRY FEDERATION.	Volume	117
Date	01/01/1991		
Language	English		

REQUEST DETAILS

Portion Type	Chapter/article	Distribution	Other territories and/or countries
Page Range(s)	7	Enter Territories/Countries	Brazil
Total Number of Pages	7	Translation	Original language of publication
Format (select all that apply)	Print, Electronic	Copies for the Disabled?	No
Who Will Republish the Content?	Academic institution	Minor Editing Privileges?	No
Duration of Use	Life of current edition	Incidental Promotional Use?	No
Lifetime Unit Quantity	Up to 499	Currency	USD
Rights Requested	Main product		

NEW WORK DETAILS

Title	Pseudomonas spp. causing spoilage in Brazilian cheeses by pigment production	Institution Name	Universidade Federal de Viçosa
		Expected Presentation Date	2023-02-22
Instructor Name	Luís Augusto Nero		

ADDITIONAL DETAILS

Order Reference Number	N/A	The Requesting Person/Organization to Appear on the License	Rafaela da Silva Rodrigues

REQUESTED CONTENT DETAILS

Title, Description or Numeric Reference of the Portion(s)	Pseudomonas sp. as the causative agent of anomalous blue discoloration in Brazilian fresh soft cheese (Minas Frescal)	Title of the Article/Chapter the Portion Is From	Pseudomonas sp. as the causative agent of anomalous blue discoloration in Brazilian fresh soft cheese (Minas Frescal)
Editor of Portion(s)	Rodrigues, Rafaela da Silva; Machado, Solimar Gonçalves; Carvalho, Antônio Fernandes de; Nero, Luís Augusto	Author of Portion(s)	Rodrigues, Rafaela da Silva; Machado, Solimar Gonçalves; Carvalho, Antônio Fernandes de; Nero, Luís Augusto
Volume of Serial or Monograph	117	Issue, if Republishing an Article From a Serial	N/A
Page or Page Range of Portion	105020	Publication Date of Portion	2021-06-01

RIGHTSHOLDER TERMS AND CONDITIONS

Elsevier publishes Open Access articles in both its Open Access journals and via its Open Access articles option in subscription journals, for which an author selects a user license permitting certain types of reuse without permission. Before proceeding please check if the article is Open Access on http://www.sciencedirect.com and refer to the user license for the individual article. Any reuse not included in the user license terms will require permission. You must always fully and appropriately credit the author and source. If any part of the material to be used (for example, figures) has appeared in the Elsevier publication for which you are seeking permission, with credit or acknowledgement to another source it is the responsibility of the user to ensure their reuse complies with the terms and conditions determined by the rights holder. Please contact permissions@elsevier.com with any queries.

Marketplace Permissions General Terms and Conditions

The following terms and conditions ("General Terms"), together with any applicable Publisher Terms and Conditions, govern User's use of Works pursuant to the Licenses granted by Copyright Clearance Center, Inc. ("CCC") on behalf of the applicable Rightsholders of such Works through CCC's applicable Marketplace transactional licensing services (each, a "Service").

1) **Definitions.** For purposes of these General Terms, the following definitions apply:

"License" is the licensed use the User obtains via the Marketplace platform in a particular licensing transaction, as set forth in the Order Confirmation.

"Order Confirmation" is the confirmation CCC provides to the User at the conclusion of each Marketplace transaction. "Order Confirmation Terms" are additional terms set forth on specific Order Confirmations not set forth in the General Terms that can include terms applicable to a particular CCC transactional licensing service and/or any Rightsholder-specific terms.

CCC Marketplace

Order Date	17-Feb-2023	**Type of Use**	Republish in a thesis/dissertation
Order License ID	1325369-2	**Publisher**	ELSEVIER BV
ISSN	0958-6946	**Portion**	Chapter/article

LICENSED CONTENT

Publication Title	International dairy journal	**Country**	Netherlands
Article Title	Comparative genomic and functional annotation of Pseudomonas spp. genomes responsible for blue discoloration of Brazilian fresh soft cheese	**Rightsholder**	Elsevier Science & Technology Journals
		Publication Type	Journal
		Start Page	105605
		Volume	140
Author/Editor	INTERNATIONAL DAIRY FEDERATION.		
Date	01/01/1991		
Language	English		

REQUEST DETAILS

Portion Type	Chapter/article	**Distribution**	Other territories and/or countries
Page Range(s)	10	**Enter Territories/Countries**	Brazil
Total Number of Pages	10	**Translation**	Original language of publication
Format (select all that apply)	Print, Electronic	**Copies for the Disabled?**	No
Who Will Republish the Content?	Academic institution	**Minor Editing Privileges?**	No
Duration of Use	Life of current edition	**Incidental Promotional Use?**	No
Lifetime Unit Quantity	Up to 499	**Currency**	USD
Rights Requested	Main product		

NEW WORK DETAILS

Title	Pseudomonas spp. causing spoilage in Brazilian cheeses by pigment production	**Institution Name**	Universidade Federal de Viçosa
		Expected Presentation Date	2023-02-22
Instructor Name	Luís Augusto Nero		

ADDITIONAL DETAILS

Order Reference Number	N/A	**The Requesting Person/Organization to Appear on the License**	Rafaela da Silva Rodrigues

REQUESTED CONTENT DETAILS

Title, Description or Numeric Reference of the Portion(s)	Comparative genomic and functional annotation of Pseudomonas spp. genomes responsible for blue discoloration of Brazilian fresh soft cheese	**Title of the Article/Chapter the Portion Is From**	Comparative genomic and functional annotation of Pseudomonas spp. genomes responsible for blue discoloration of Brazilian fresh soft cheese
Editor of Portion(s)	Rodrigues, Rafaela da Silva; Machado, Solimar Gonçalves; Carvalho, Antônio Fernandes de; Nero, Luís Augusto	**Author of Portion(s)**	Rodrigues, Rafaela da Silva; Machado, Solimar Gonçalves; Carvalho, Antônio Fernandes de; Nero, Luís Augusto
Volume of Serial or Monograph	140	**Issue, if Republishing an Article From a Serial**	N/A
Page or Page Range of Portion	105605	**Publication Date of Portion**	2023-05-01

RIGHTSHOLDER TERMS AND CONDITIONS

Marketplace Permissions General Terms and Conditions

SCIENTIFIC PRODUCTS DURING THE DOCTORATE

Journal papers

1. **Rodrigues, R. S**., Machado, S. G., Carvalho, A. F., Nero, L. A. *Pseudomonas* sp. as the causative agent of anomalous blue discoloration in Brazilian fresh soft cheese (Minas Frescal). *International Dairy Journal*, v. 117, 105020, 2021. https://doi.org/10.1016/j.idairyj.2021.105020

2. Oliveira, F. S., **Rodrigues, R. S**., Carvalho, A. F., Nero, L. A. Genomic analyses of *Pediococcus pentosaceus* ST65ACC, a bacteriocinogenic strain isolated from artisanal raw-milk cheese. *Probiotics and Antimicrobial Proteins*, 2022. https://doi.org/10.1007/s12602-021-09894-1

3. Teixeira, C. G., **Rodrigues, R. S.**, Yamatogi, R. S., Lucau-Danila, A., Drider, D., Nero, L. A., Carvalho, A. F. Genomic analyses of *Weissella cibaria* W25, a potential bacteriocin-producing strain isolated from pasture in Campos das Vertentes, Minas Gerais, Brazil. *Microorganisms*, v. 10, n. 2, 314, 2022. https://doi.org/10.3390/microorganisms10020314

4. Fusieger, A., da Silva, R. R., Cavicchioli, V. Q., **Rodrigues, R. S.**, Honorato, J. A., Silva, S. R. J., Pena, M. L., Caggia, C., Nero, L. A., Carvalho, A. F. Influence of emulsifying salts on the growth of *Bacillus thuringiensis* CFBP 3476 and *Clostridium perfringens* ATCC 13124 in processed cheese. *Foods*, v. 11, n. 20, 3217, 2022. https://doi.org/10.3390/foods11203217

5. Costa, N. A. S., **Rodrigues, R. S.**, Carvalho, A. F., & Machado, S. G. Proteases e lipases produzidas por *Pseudomonas*: um desafio na indústria de lácteos. *The Journal of Engineering and Exact Sciences*, v. 8, n. 9, 14900-01e, 2022. https://doi.org/10.18540/jcecvl8iss9pp14900-01e

6. Costa, N. A. S., **Rodrigues, R. S.**, Carvalho, A. F., & Machado, S. G. Formação de biofilme e produção de pigmento azul por *Pseudomonas*: o impacto na indústria de lácteos. *The Journal of Engineering and Exact Sciences*, v. 8, n. 8, 14898-01e, 2022. https://doi.org/10.18540/jcecvl8iss8pp14898-01e

7. Souza, L. V., **Rodrigues, R. S.**, Fusieger, A., da Silva, R. R., Silva, S. R. J., Martins, E., Machado, S. G., Caggia, C., Randazzo, C. L., Carvalho, A. F. Diversity of filamentous fungi associated with dairy processing environments and spoiled products in Brazil. *Foods*, v. 12, n. 1, 153, 2023. https://doi.org/10.3390/foods12010153

8. **Rodrigues, R S.**, Machado, S. G., Carvalho, A. F, Nero, L. A. Comparative genomic and functional annotation of *Pseudomonas* spp. genomes responsible for blue discoloration of Brazilian fresh soft cheese. *International Dairy Journal*, 2023. https://doi.org/10.1016/j.idairyj.2023.105605

9. Teixeira, C. G., **Rodrigues, R. S.**, Lucau-Danila, A., Nero, L. A., Carvalho, A. F., Drider, D. Genome analyses of *Weissella* strains isolated from Campos das Vertentes, Minas Gerais, Brazil revealed new bacteriocins with a large spectrum of activity. *Food Bioscience*, v. 52, 102421, 2023. https://doi.org/10.1016/j.fbio.2023.102421

Abstracts

1. **Rodrigues, R. S**., Machado, S. G., Carvalho, A. F., Nero, L. A. *Pseudomonas* spp. as the responsible agent for blue discoloration spoilage in fresh cheese (Minas Frescal). In: 31º Congresso Brasileiro de Microbiologia 2021. *Anais do 31º Congresso Brasileiro de Microbiologia 2021*, 2021.

2. Oliveira, F. S., **Rodrigues, R. S**., Carvalho, A. F., Nero, L. A. Complete genome sequence of *Pediococcus pentosaceus* ST65ACC isolated from artisanal raw-milk cheese. In: 31º Congresso Brasileiro de Microbiologia 2021. *Anais do 31º Congresso Brasileiro de Microbiologia 2021*, 2021.

Other scientific contributions

Antônio Fernandes, Luís Nero, Solimar Gonçalves Machado, Andressa Falqueto, **Rafaela Rodrigues**. Manchas azuis em queijos frescos. *Profissão Queijeira*. N° 11, março-abril 2022.