Isolation and Characterization of Carotenoids in *Pedobacter* and *Roseopedobacter* gen. nov.

Presented to the faculty of Lycoming College in partial fulfillment of the requirements for Departmental Honors in Biology

by
Kiyah Bell
Lycoming College
May 4, 2022

Approved by:

(Signature)

(Signature)

(Signature)

(Signature)

Isolation and Characterization of Carotenoids in *Pedobacter* and *Roseopedobacter* gen. nov.

A thesis presented by

Kiyah Bell

to

The Joanne and Arthur Haberberger Fellowship

and

the Biology Department

in partial fulfillment of the requirements for departmental honors in Biology

Lycoming College

Williamsport, PA

May 2022

Table of Contents

Abstract

Pedobacter is a genus in the *Bacteroidota* that usually produce pink or yellow pigments, typically assumed to be carotenoids. Carotenoids are the most diverse and ubiquitous pigments found in nature, synthesized by plants, fungi, algae, and bacteria. They are a class of terpenoids with a linear, isoprene backbone, and their structure consists of an extended conjugated system.

In this work, five *Pedobacter* strains were recovered from the Lycoming College Culture Collection (LCCC). Phylogenetic analysis using the Genome Taxonomy Database (GTDB) and the RNA polymerase β-subunit (rpoB) tree revealed that the yellow *Pedobacter* strains and the pink *Pedobacter* strains each shared a common ancestor and formed separate clades on long branches. A computational genomic study, the Average Amino Acid Identity (AAI), suggested that organisms of different colors should be in distinct genera. After mining the genome for carotenoid biosynthetic genes using the Rapid Annotation using Subsystem Technology (RAST) website, all of the pink species encode β-carotene ketolase, which the yellow species lack. The presence of the β-carotene ketolase enzyme should allow the pink species to produce canthaxanthin and astaxanthin, pink carotenoids, while in absence of this enzyme the yellow species should produce β-carotene and zeaxanthin, yellow and orange carotenoids. The pigments were extracted with methanol and separated by HPLC to obtain UV-Vis spectroscopic characteristics with a diode array detector (DAD) and mass data from quadrupole-time of flight (Q-TOF) mass spectrometer. Compounds with a three-peaked spectrum, typical of carotenoids, plus a maximum absorbance around 480 nm was commonly found in the pink strains, and 450 nm was the peak wavelength found in the yellow strains. The yellow species presented a pigment profile composed of zeaxanthin and lutein from mass spectrum data. In conclusion, the pink

strains should be reclassified into the new proposed genus, *Roseopedobacter,* due to their synthesis of pink carotenoids.

Introduction

Pigments are used as colorants and antioxidants in the cosmetics, pharmaceuticals, food, and textile industries. Since prehistoric times, humans have extracted natural pigments from plants, animals, and insects. Synthetic dyes have been created as an alternative to natural colorants to protect the organisms from which natural pigments are extracted. However, these dyes can be carcinogenic, lack bio-degradability, and build up in habitats of living organisms.[1] Yet, they are still used in large amounts due to their low cost, stability, and bright pigmentation.[2] In the search for safer alternatives, interest in pigment extraction from bacteria has increased because they are widely accessible in soil, water, and air as well as easy to isolate and cultivate.[3] Additionally, the pigments in bacteria are biodegradable and non-toxic, unlike synthetic dyes.[2,4]

Carotenoids are the most diverse and ubiquitous pigments found in nature, synthesized by plants, fungi, algae, and bacteria.[6] They are a class of terpenoids with a linear, isoprene backbone, producing yellow, orange, and red pigments.[6] Additionally, their structure consists of an extended conjugated system. C40 chains are the most common carotenoids, and these include carotenes such as β-carotene, zeaxanthin, astaxanthin, and canthaxanthin.[7] C30 and C50 carbon chains are less common and typically found in photosynthetic bacteria and archaea.[7]

In animals, carotenoids prove to be biologically and physiologically beneficial. Animals cannot synthesize carotenoids *de novo*, therefore they must obtain them from their diet.[7] This diet can include fruits and vegetables, which are traditional sources of carotenoids.[7] Moreover, carotenoids are involved in the treatment and prevention of diseases such as macular

degeneration and some kinds of cancer.[7] β-Carotene and lycopene protect against prostate cancer, while lutein and zeaxanthin treat macular degeneration by protecting the eye from blue light.[7,8]

Some bacterial species have the potential to produce carotenoids. These carotenoids are considered natural secondary metabolites necessary for survival within heterotrophic bacterial cells .[6,7] This pigment absorbs excessive light to protect bacteria from UV radiation, and act as singlet molecular oxygen and peroxyl radicals for photo-oxidative damage protection.[7] Also, this pigment functions in membrane fluidity to regulate the transport of nutrients and allow for growth at low temperatures.[6]

A cluster of six *crt* genes, *crtW*, *crtY*, *crtI*, *crtB*, *crtE*, and *crtZ* code for enzymes that catalyze carotenoid biosynthesis in *Brevunidomas* sp. SD212 such that *crtW* codes for β-carotene ketolase, *crtY* codes for lycopene cyclase, *crtI* codes for phyotene desaturase, *crtB* codes for phytoene synthase, *crtE* codes for geranyl geranyl pyrophosphate (GGPP) synthase, and *crtZ* codes for β-carotene hydroxylase.[8,16] Spectroscopic analysis, high performance liquid chromatography (HPLC) and mass spectrometry (MS), of extracted carotenoid pigments helped to identify carotenoid structures. The carotenoid structures and biosynthetic enzymes were assembled into a biosynthetic pathway (Figure 1). The initial step of carotenoid synthesis includes the derivation of two isoprenoid precursors, isopentenyl diphosphate (IPP) and dimethylallyl diphosphate (DMAPP), from the Methylerythritol Phosphate Pathway (MEP) or Non-mevalonate pathway.[9] Three IPP (5C) molecules and one DMAPP (5C) molecule are condensed to produce Geranylgeranyl Pyrophosphate (GGPP) by the enzyme GGPP synthase. Two GGPP (20C) molecules are converted to phytoene, C40 chain, by phytoene synthase.[10]

Other C40 carotenoids are derived from phytoene. Two enzymes, β-carotene hydroxylase and β-carotene ketolase, catalyze the reactions to convert phytoene into other carotenoids.[10]

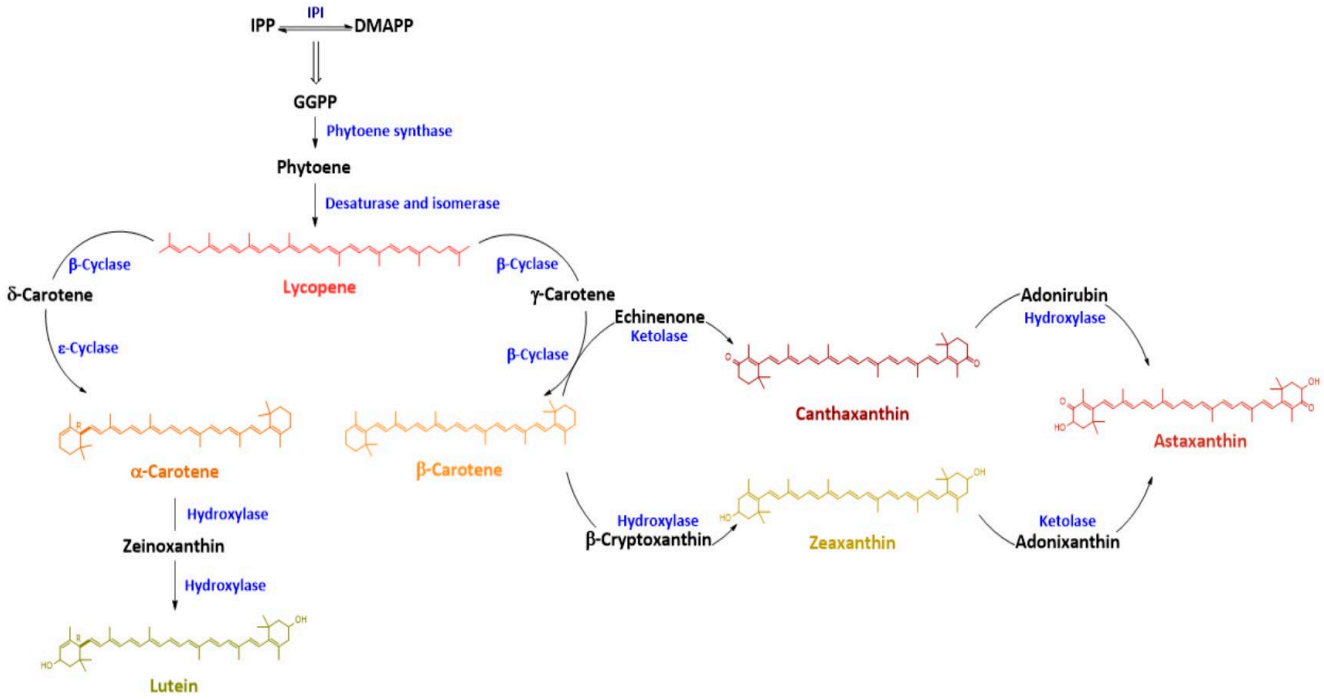

Figure 1. Proposed carotenoid biosynthetic pathway of C40 carotenoids derived from phytoene. The enzymes are highlighted in blue, and the arrows represent a catalyzed reaction.[8]

HPLC with ultraviolet-visible (UV-vis) detection and mass spectrometry (LC-MS) detection is commonly used to analyze carotenoids. Researchers prefer reversed-phase HPLC (RP-HPLC) containing a C18 column and a binary elution gradient.[15] Atmospheric pressure chemical ionization (APCI) is often chosen as the ionization source for carotenoids because of their low polarity and small molecular weight.[15] APCI with positive ionization produces ions via protonation of molecules. First, the HPLC performs a chromatographic of the compounds based on polarity. UV-vis spectra are acquired of compounds that elute. Then, compounds are passed

into the MS, ionized, and separated based on molecular weight. Elucidation of the chemical structure is completed by comparison of the data with standards and literature values.[7]

A recent study performed HPLC analysis of carotenoid standards, which included β-carotene, zeaxanthin, canthaxanthin, and astaxanthin, resulting in UV-vis spectra (Figure 2).[17] β-Carotene and zeaxanthin's UV-vis spectra are nearly identical. Similarly, the UV-vis spectra of canthaxanthin and astaxanthin are also quite similar. β-Carotene and zeaxanthin exhibit compounds with three maximum peaks around 420 nm, 450 nm, and 480 nm (Figure 2). Canthaxanthin and astaxanthin present one maximum peak at 480 nm (Figure 2). Following HPLC analysis, researchers published APCI (+) mass spectra for β-carotene, zeaxanthin, canthaxanthin, and astaxanthin extracted from varying sources. The m/z for each of the carotenoids shown below is: β-carotene 536.4 g/mol, zeaxanthin 564.4 g/mol, canthaxanthin 565.4 g/mol, and astaxanthin 596.4 g/mol (Figure 3).[18,19,20]

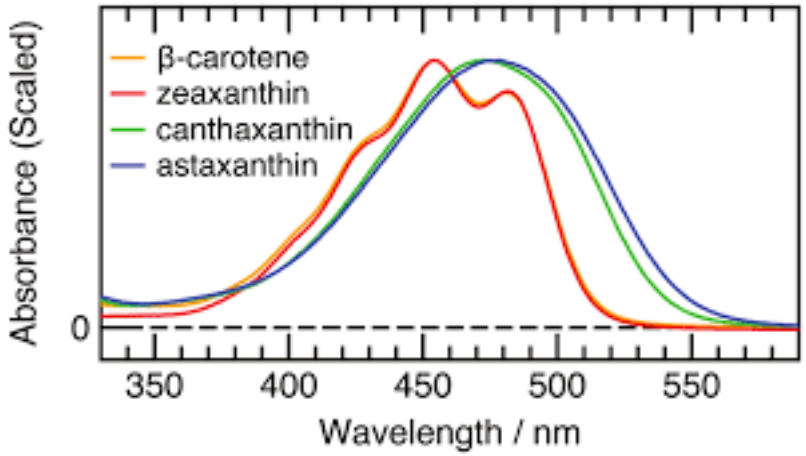

Figure 2. UV-vis spectra of carotenoid standards, β-carotene, zeaxanthin, canthaxanthin, and astaxanthin, analyzed by HPLC.[17]

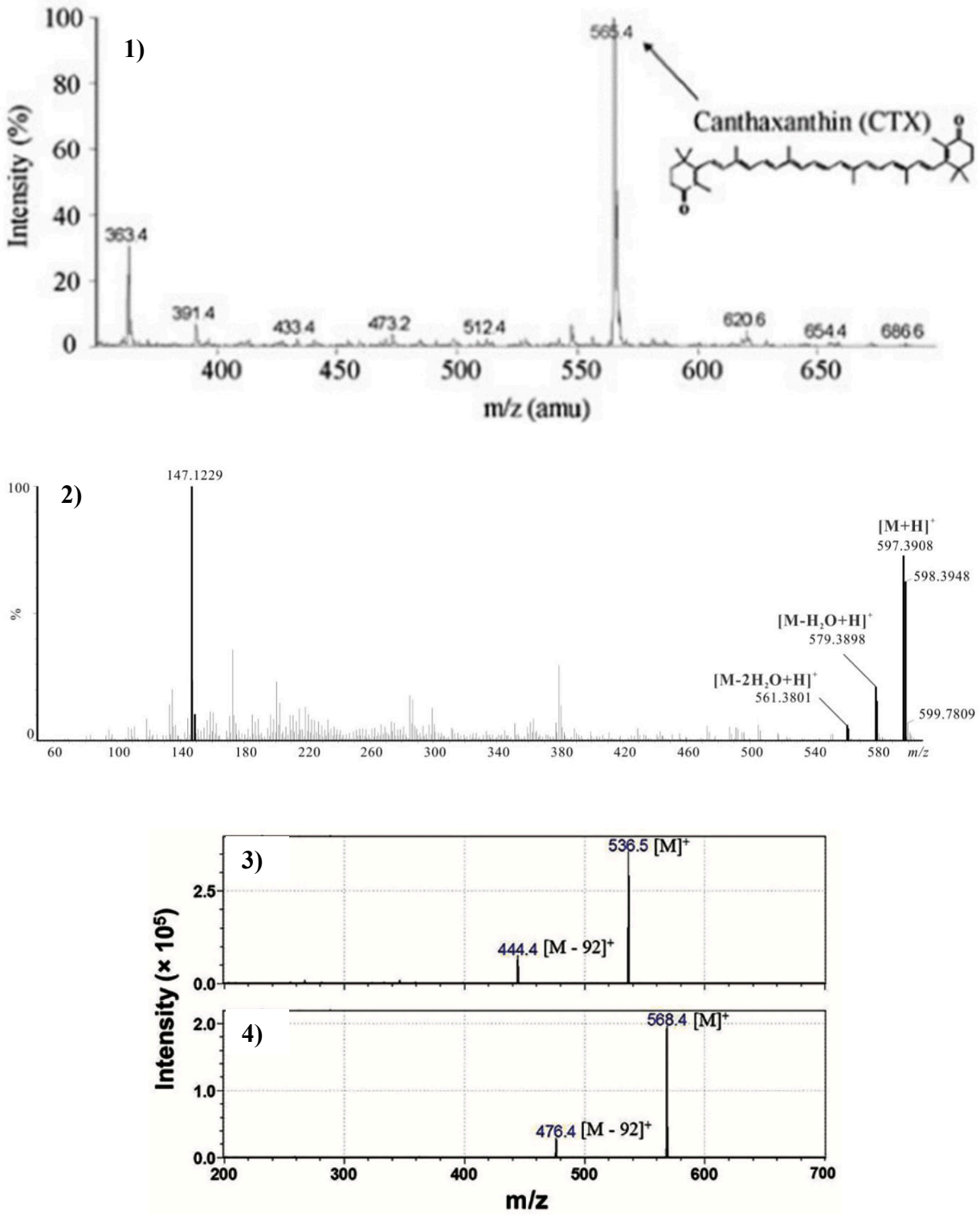

Figure 3. MS- APCI (+) spectra of (1) canthaxanthin, (2) astaxanthin, (3) β-carotene, and (4) zeaxanthin.[18,19,20] Canthaxanthin was extracted from *Dietzia* natronolimnaea HS-1 and

astaxanthin was a standard.[18,19] Extraction of β-carotene and zeaxanthin originated from *Pandalus* borealis. [20]

In the present study, pigmented strains were recovered and identified by 16S rRNA gene sequencing. Examination of the genomes sequenced by prior students allowed for a more accurate classification of the organisms. Reported here are the genes responsible for carotenoid biosynthesis and the carotenoid profile in *Pedobacter* strains, and the proposal of a new genus called *Roseopedobacter*. The study of pigments in *Pedobacter* are limited, therefore carotenoid characterization in this genus was based on the UV-vis and mass spectrum. In contrast with other published work, pigments were analyzed from many organisms, not just one, from a taxonomic and evolutionary perspective. Taxonomic analysis of *Pedobacter* species was conducted through phylogenetic trees, Genome Taxonomy Database (GTDB) tree and RNA polymerase β-subunit (rpoB) tree, and genomic calculations of AAI and ANI. The function of genes involved in carotenoid formation was analyzed through genome annotation in Rapid Annotation using Subsystem Technology (RAST) and National Center for Biotechnology Information (NCBI). Identification and characterization of carotenoids in *Pedobacter* extracts was carried out by HPLC-MS.

Materials and Methods

Bacterial recovery and growth

Bacterial strains were obtained from the German Collection of Microorganism and Cell Cultures or isolated from a nearby creek in Pennsylvania by Dr. Jeff Newman (Table 1). The bacteria

were recovered from the -80˚C Lycoming College Culture Collection and plated on Reasoner's 2A agar (R2A) (Table 1). Plates were generally incubated at 22˚C for 2 days.

Genus	Species	Strain	LCCC #	Source	Year	Color	Genome Accession
Pedobacter	alluvionis	DSM 19624^T	1354	DSMZ	2013	pink	RCCK01
Pedobacter	heparinus	DSM 2366^T	1440	DSMZ	2014	yellow	NC_013061
Pedobacter	sp.	MC2016-24	1733	Mill Creek, PA	2016	yellow	JADFBX01
Pedobacter	sp.	MR2016-19	1724	Miller's Run, PA	2016	pink	JADCNK01
Pedobacter	sp.	R20-19	1073	Larry's Creek, PA	2013	pink	JCKI01

Table 1. The LCCC #, origin, color, source, and year of the strains recovered from the Lycoming College Culture Collection. LCCC # refers to the number that the organism is assigned in the Lycoming College Culture Collection. DSMZ is the German Collection of Microorganism and Cell Cultures GmbH (https://www.dsmz.de).

Amplification of the 16S rRNA gene and sequencing

A single colony was inoculated into 100 µL of deionized water in a 0.5 mL centrifuge tube. The cells were subjected to two freeze (-70˚C) and thaw (80˚C) cycles for two minutes each. Reaction

mixtures for PCR contained 12.5 μL of PCR Master Mix (Thermo Fisher), 12.5 μL of 2x 16S rRNA primer mix (27f and 1492r at 1.5 μM each), and 1 μL of genomic DNA in 0.2 mL microcentrifuge tubes. Amplification of the 16S rRNA gene fragments were done in a BIO-RAD T100^{TM} Thermal Cycler. The three cycles occurred as followed: denaturation for 3.5 min at 95˚C, then 34 cycles of 30s at 95˚C, 30s at 50˚C, 60s at 72˚C, and a final extension step of 5 min at 72°C. The PCR products were analyzed by gel electrophoresis in 1% agarose gel at 100 volts for 20 minutes. Agarose gel consisted of 40 mL of deionized water, 0.4g of molecular grade agarose, 800 μL of 50X TAE buffer, and 10 μL of ethidium bromide (2 mg/mL). Then, the PCR products were cleaned by the addition of 2 μL of Exosap (Thermo Fisher), 3 μL of deionized water, and 20 μL of PCR product into a 0.5 mL microcentrifuge tube and incubated at 37 °C for 30 minutes, followed by 15 min at 85˚C to inactivate the enzyme. The PCR products were sent to Genewiz for Sanger Sequencing. The RNA sequences were evaluated by the Nucleotide Basic Local Alignment Tool (BLASTn) and compared to the bacterial and archaeal 16S rRNA gene sequences of the National Centre for Biotechnology Information database (NCBI) (https://blast.ncbi.nlm.nih.gov/Blast.cgi).[12] This procedure ensured that the organism recovered matched the expected strain.

Pigment extraction

Strains were streaked on two R2A plates. One plate was grown in the presence of light and the other plate in a dark incubator at 22 °C to determine if synthesis of the pigment is induced by exposure to light. After 24 hours of growth, approximately 15-70 mg of bacteria was extracted with LC-MS grade methanol in a 1.5 mL microcentrifuge tube. The extracted pigment was

placed in an orbital shaker at 25°C and 200 rpm for an hour. The supernatant was decanted into a 2 mL glass vial and stored in a freezer at -18 °C.

Carotenoid identification

The carotenoids from the bacterial extracts were identified by chromatographic and spectroscopic attributes analyzed by HPLC-MS and the data was compared to literature values. Analysis was performed with Agilent Technologies 1200 series HPLC Zorbax Eclipse Plus C18 (2.1x50 mm) 1.8-Micron column fitted with a guard column of the same material coupled to an Agilent 6545 Q-TOF mass spectrometer. The APCI was applied in the positive ionization mode due to the biochemical composition of carotenoids. Absorbance detection was with a diode array detector (DAD) at 450 nm and 480 nm, and the spectra were recorded in the λ 250-750 nm range. The HPLC method used was previously stated.[22] Reversed phase carotenoid separation was carried out at 10 °C with a sample volume of 0.5 μL. Separation occurred at a flow rate of 0.4 ml/min and elution with 0.1% acetic acid in methanol (solvent A) and Methyl tert-Butyl Ether (MTBE) (Solvent B). The short gradient selected was 100% solvent B at 0 min, 50% solvent B at 10 min, 50% solvent B at 12 min, and 0% solvent B at 15 min.

Results and Discussion

1. Recovery and Growth of Strains

Three novel strains, *Pedobacter* sp. MC2016-24, *Pedobacter* sp. MR2016-19, and *Pedobacter* sp. R20-19, and two related type strains, *Pedobacter alluvionis* and *Pedobacter heparinus*, were recovered and examined from a taxonomic and evolutionary perspective. One of the organisms, *Pedobacter* alluvionis, was not recovered from the LCCC during early experiments, therefore the

pigmentation of that strain is not described here. The four strains appeared pink or yellow, suggesting they are carotenoid producers (Figure 4).

Each of the four of the organisms were grown on two separate petri plates, with one plate grown in the presence of light and the other plate grown in a dark incubator (Figure 4). Traditionally the plates are grown in the dark, however comparing the growth in the dark versus the light helped to visually assess if carotenoid synthesis was induced by exposure to light. In each strain, more pigment was produced in the light, therefore carotenoid synthesis may be induced by light exposure (Figure 4).

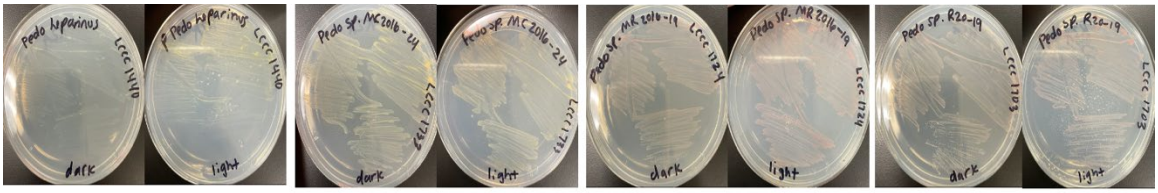

Figure 4. Yellow and pink *Pedobacter* strains analyzed in this study, grown in a dark incubator versus in the presence of light: (A) *P*. heparinus (dark), (B) *P*. *heparinus* (dark), (C) *P*. sp. MC2016-24 (dark), (D) *P*. sp. MC2016-24 (light), (E) *P*. sp. MR2016-19 (dark), (F) *P*. sp. MR2016-19 (light), (G) *P*. sp. R20-19 (dark), and (H) *P*. sp. R20-19 (light). Bacterial growth plates A, C, E, and G were incubated in the dark. Plates B, D, F, and H were grown in the light.

Later all *Pedobacter* organisms in the LCCC were recovered, which showed the varying pink and yellow pigments produced. *P*. *alluvionis* was recovered and grown displaying its bright pink pigment (Figure 5).

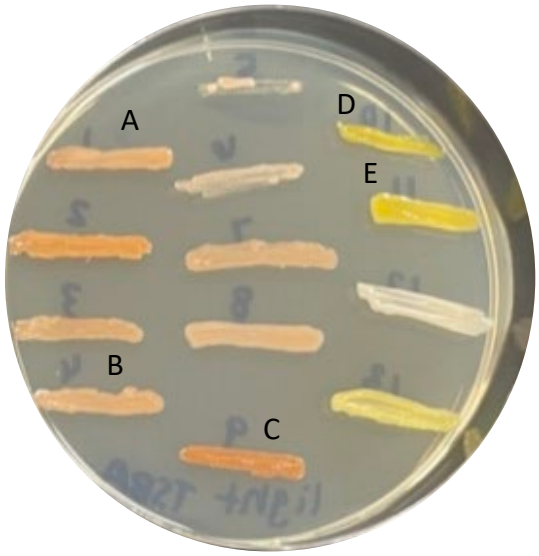

Figure 5. All *Pedobacter* organisms in the LCCC recovered on Tryptic Soy Blood Agar (TSBA) for two days of growth. The five organisms in this study are labeled: (A) *P*. sp. R20-19, (B) *P*. sp. MR2016-19, (C) *P*. *alluvionis*, (D) *P*. *heparinus*, and (E) *P*. sp. MC2016-24.

2. Phylogenomic Studies

Taxonomic analysis of a Genome Taxonomy Database Tree (GTDB) was constructed in Mega 7 to solely display the *Pedobacter* genus.[14,26,27,28] This tree illustrates the evolutionary relationships between organisms based on the alignment of 120 concatenated protein sequences. *P*. sp. R20-19 clustered most closely with *P. alluvionis* and other pink *Pedobacter* strains (Figure 6). Also, the pink strains shared a single common ancestor indicating that they are closely related. The non-pink *Pedobacter* species cluster into several groups, one of which includes a bacterium classified into the genus *Nubsella*. The GTDB tree is the best method for determining phylogenetic relationships between organisms because it is assembled from a large dataset, however only organisms that are deposited into GenBank are listed in the tree and a new tree is only released yearly. *P*. sp. MC2016-24 and *P*. sp. MR2016-19 had not been deposited into GenBank at the

time the GTDB tree was constructed, therefore these organisms are not listed in the tree (Figure 6).

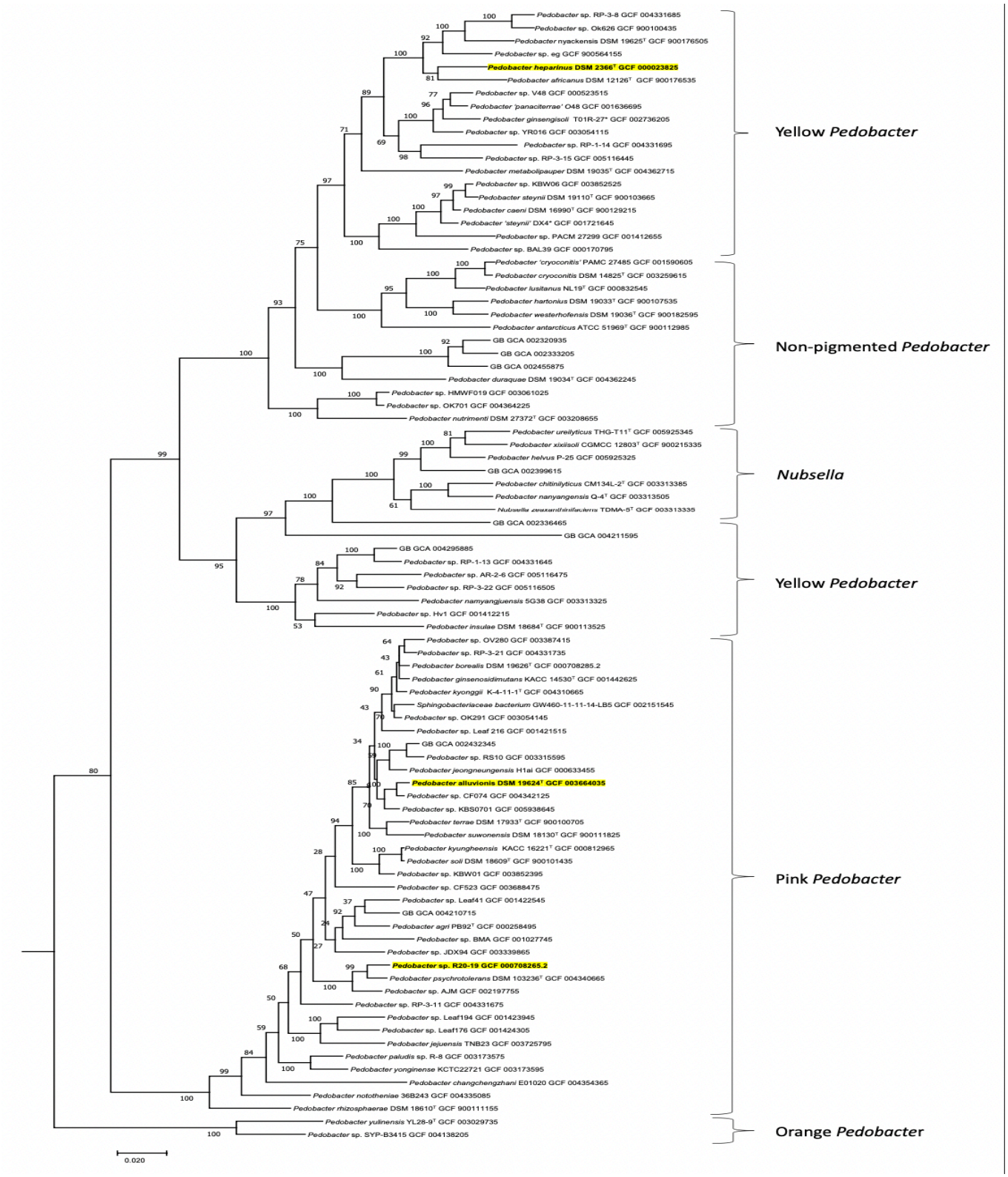

Figure 6. Genome Taxonomy Database (GTDB) neighbor-joining phylogenetic tree displays the *Pedobacter* branch.[14,26,27,28] The organisms analyzed within this study are highlighted in yellow.

Similarly to the GTDB tree, the RNA-polymerase β-subunit (rpoB) tree allows for phylogenetic studies of the *Pedobacter* genus. The rpoB tree determines evolutionary relationships between organisms based on the RNA polymerase β-subunit, generated in MEGA 7. [14,26,27,28] These data suggest that *P*. heparinus and *P*. sp. MC2016-24 clustered most closely with other yellow *Pedobacter* strains and shared a common ancestor, appearing to be monophyletic (Figure 7). While *P*. *alluvionis*, *P*. sp. MR2016-19, and *P*. sp. R20-19 clustered most closely with the other pink *Pedobacter* strains and shared a common ancestor; indicating that the pink strains are monophyletic as well (Figure 7). Conducting phylogenetic analysis through a rpoB tree is an easy process, can be used as a substitute for the GTDB tree, and has good resolution.

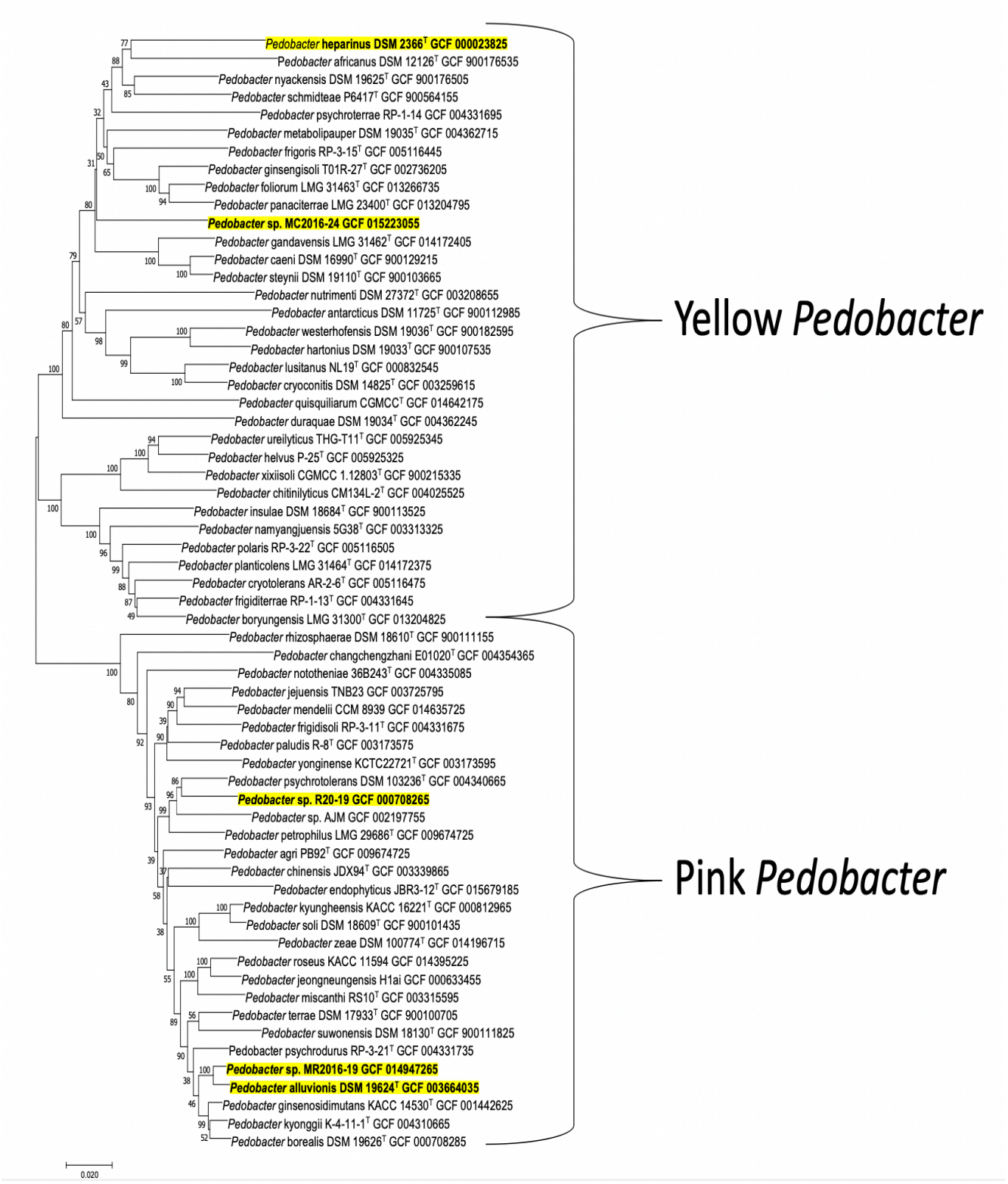

Figure 7. Neighbor-joining phylogenetic tree constructed from the RNA polymerase β-subunit.

The organisms analyzed within this study are highlighted in bold. [14,26,27,28]

3. Genome Relatedness and Comparisons

Genomic studies were performed through the calculation of Average Nucleotide Identity (ANI) and Average Amino Acid Identity (AAI). ANI is the average percent similarity of the genes shared between organisms. Calculations were completed by downloading the FASTA genome sequences and then uploading the sequences to the OrthoANI tool.[23] An ANI value below 95 means that the organisms being compared are different species, and a value above 95 means that the organisms are the same species. *P*. sp. MC2016-24 is a separate species from *P. heparinus* because they share an ANI percent of 74.62% (Figure 8). *P*. sp. MR2016-19 is a separate species from *P*. alluvionis because they share an ANI percent of 93.93% (Figure 8).

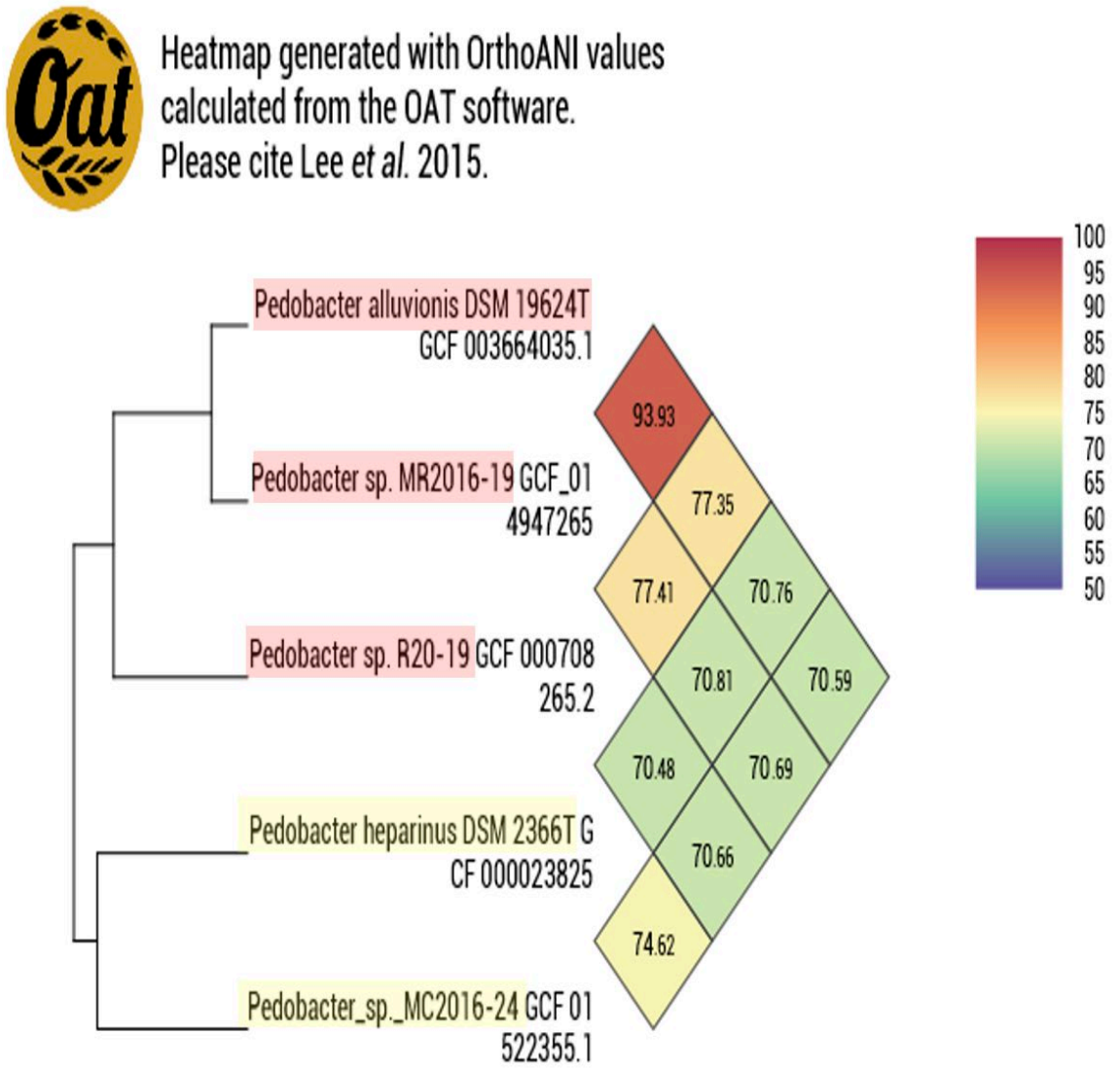

Figure 8. Calculated Average Nucleotide Identity (ANI) shared between the five *Pedobacter* strains. The pink *Pedobacter* strains are highlighted in pink, while the yellow *Pedobacter* strains are highlighted in yellow.

Average Amino Acid Identity (AAI) is a complementary technique to ANI that examines genus membership. AAI is the percentage of amino acids shared between orthologous proteins (those

that perform the same function) in different organisms. The AAI values were generated by a custom AAI calculator available at http://lycofs01.lycoming.edu/~newman /ROSA.html. An AAI value below 70 suggests that the organisms being compared are in a different genus, while a value above 75 suggests the same genus.[24] The pink strains have AAI values above 75, therefore they should be characterized in the same genus, *Roseopedobacter* (Table 2). The yellow strains have AAI percentages between 70-75%, therefore referring to other criteria such as the GTDB tree and rpoB tree will solidify the genus determination (Table 2). The GTDB and rpoB trees displayed that the yellow strains, *P. heparinus* and *P.* sp. MC2016-24 were monophyletic, so the yellow strains should be characterized in the same genus, *Pedobacter*. The pink and yellow strains should belong to distinct genera with shared AAI percentages around the mid to high 60s.

Pedobacter	sp. MR2016-19	alluvionis	sp. R20-19	heparinus	sp. MC2016-24
sp. MR2016-19		94.7	80.0	65.7	65.4
alluvionis	94.6		79.9	66.3	65.4
sp. R20-19	80.0	79.8		65.6	65.4
heparinus	65.8	66.3	65.5		73.4
sp. MC2016-24	65.5	65.3	65.4	73.5	

Table 2. Calculated Average Amino Acid Identity (AAI) shared between the five *Pedobacter* strains. The pink *Pedobacter* strains are highlighted in pink, while the yellow *Pedobacter* strains are highlighted in yellow. The top left box in bold shows the AAI values shared between the pink strains, and the bottom right box in bold shows the AAI values shared between the yellow strains.

Additional evidence for genomic similarity shared between the pink strains versus the yellow strains was apparent in the number of shared genes. The Venn Diagram, generated from a custom computational tool, displays unique genes within an organism and shared genes between organisms to help determine biochemical and molecular differences.[25] The total number of genes shared between the five *Pedobacter* organisms was 2067 (Figure 9). The closely related yellow strains shared 568 genes, while the pink strains shared 543 genes (Figure 9). However, the number of genes shared lowers significantly when comparing pink and yellow strains. For example, two yellow species, *P*. *heparinus* and *P*. sp. MC2016-24, and two pink species, *P*. sp. MR2016-19 and *P*. sp. R20-19 share only 29 unique genes (Figure 9). Another example includes the comparison of three pink species, *P*. sp. MR2016-19, *P*. sp. R20-19, and *P*. *alluvionis*, to one yellow species, *P*. *heparinus,* with 149 shared genes (Figure 9).

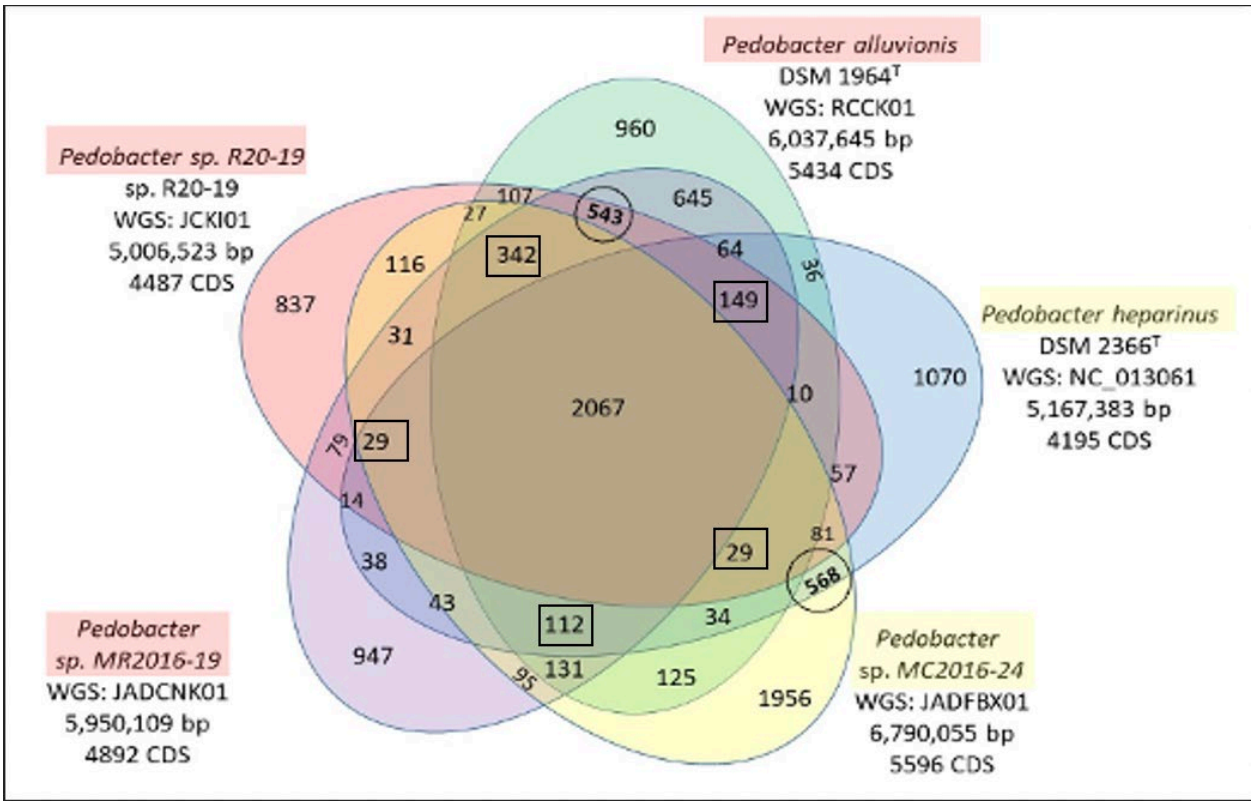

Figure 9. Venn Diagram comparing unique genes shared amongst these microbes. Each colored oval is indicative of the unique genes in the *Pedobacter* strain, and the overlap of the ovals indicates the shared genes between the organisms. The circled numbers represent the shared genes between *Pedobacter* species of the same color. The boxed numbers represent shared genes between pink and yellow strains.

4. Enzymatic and Biosynthetic Mechanism of Carotenoids

To further support the division of the *Pedobacter* genus, enzymes involved in carotenoid biosynthesis were discovered through genome mining in RAST. All the organisms evaluated contained phytoene synthase, phytoene dehydrogenase/ phytoene desaturase, lycopene cyclase, and β-carotene hydroxylase (Table 3). The pink strains also contained β-carotene ketolase, while

the yellow strains did not (Table 3). The difference in this one enzyme may allow the pink strains to produce pink pigments. Phytoene dehydrogenase and phytoene desaturase catalyze the same basic reaction, however RAST differentiates the two enzymes within their system suggesting that there may be some subtle differences. The yellow strains contained phytoene dehydrogenase, while the pink strains contained phytoene desaturase (Table 3).

Organism	Color	Phytoene synthase (*crtB*)	Phytoene dehydrogenase/ desaturase* (*crtI*)	Lycopene cyclase (*crtY*)	β-carotene hydroxylase (*crtZ*)	β-carotene ketolase (*crtW*)
Pedobacter heparinus	Yellow	+	+	+	+	-
Pedobacter sp. MC2016-24	Yellow	+	+	+	+	-
Pedobacter alluvionis	Pink	+	+*	+	+	+
Pedobacter sp. MR2016-19	Pink	+	+*	+	+	+
Pedobacter sp. R20-19	Pink	+	+*	+	+	+

Table 3. Five *crt* genes coding for enzymes that catalyze carotenoid synthesis in five *Pedobacter* species. The (+) signifies the presence of a gene and the (-) denotes the absence of a gene. The "*" represents phytoene desaturase.

After determining the enzymes present in each strain, predictions were made as to which carotenoids will be produced by the pink strains versus the yellow strains. Pink *Pedobacter* strains should produce canthaxanthin and astaxanthin, pink and red carotenoids, because those

organisms contain β-carotene ketolase (Figure 10). Yellow *Pedobacter* strains should produce β-carotene and zeaxanthin, orange and yellow pigments, due to the lack of β-carotene ketolase (Figure 10).

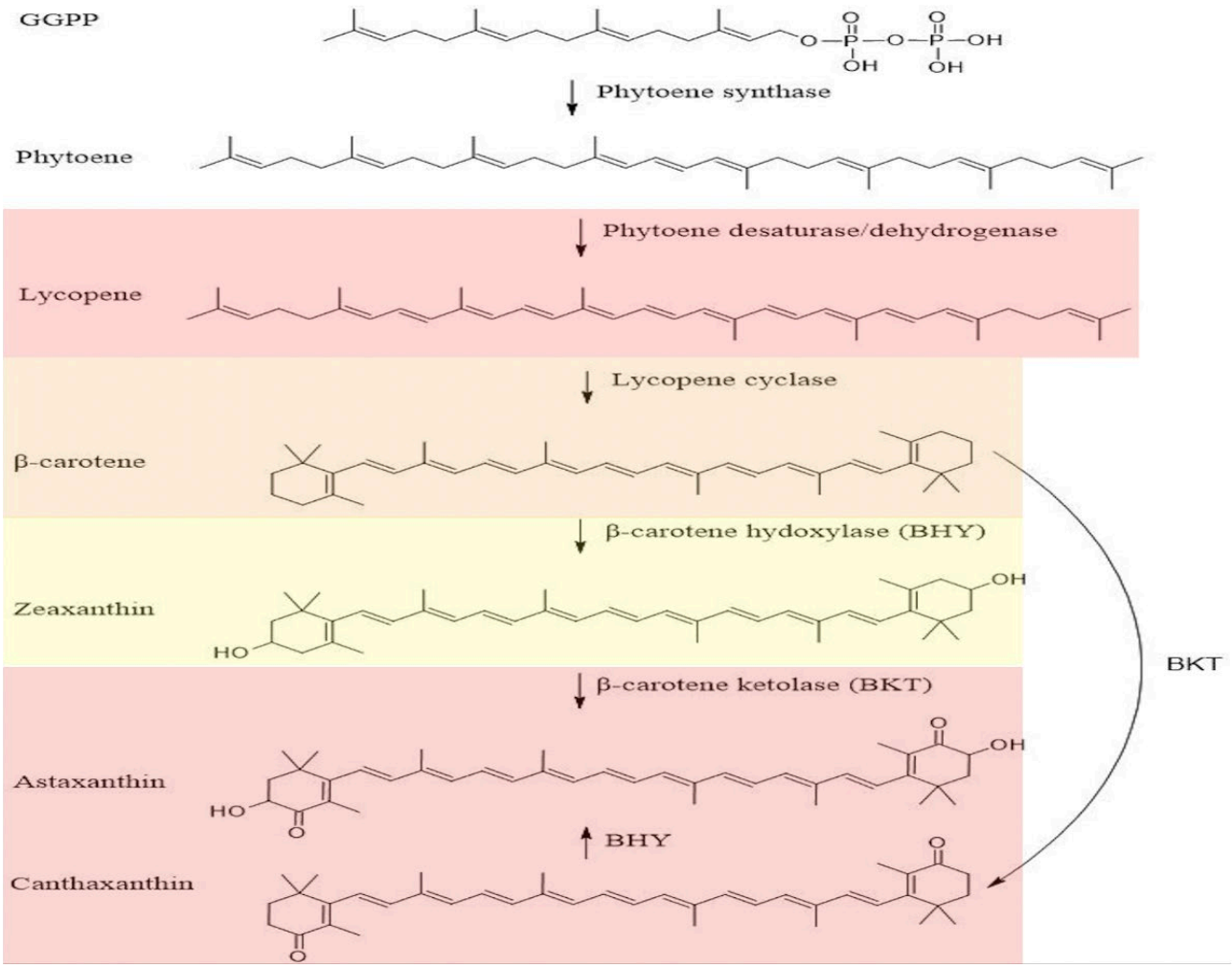

Figure 10. Carotenoid biosynthetic pathway proposed by Rebelo *et al*.[8] The highlighted boxes represent pigmentation of the carotenoids. GGPP refers to Geranylgeranyl pyrophosphate.

5. Characterization of carotenoids

To identify and characterize the carotenoid profile, chromatographs, UV-vis spectra, and mass spectroscopy properties were analyzed (Table 4). Individual carotenoid pigments were separated by HPLC before identification by MS occurred.

Pedobacter strain	Color	Peak	Elution time (min)	Peak wavelength (λ_{max} nm)
heparinus	Yellow	1, 2, 3	0.7, 1.25, 1.7	426, 450, 478
sp. R20-19	Pink	1, 2, 3, 4, 5	0.72, 0.88, 0.98, 1.42, 1.91	450, 480, 502
alluvionis	Pink	—	—	454, 478, 506
sp. MR2016-19	Pink	1, 2, 3, 4, 5	0.72, 0.88, 0.92, 1.3, 1.74	450, 480, 502
sp. MC2016-24	Yellow	1, 2, 3	0.7, 1.25, 1.7	426, 450, 478

Table 4. Identification of carotenoids in the bacterial extracts through performance of HPLC-MS. Analysis based on chromatographs, UV-vis spectra, and mass spectroscopy data shown in Figure 14, 15, 17, 18. "—" represents no data acquired for the organism.

The 3D plots and UV-Vis spectra obtained from the recovered strains, presented three peaks between 400-500 nm which is characteristic of carotenoids. The 3D plots were examined from HPLC analyses by former students but did not include all the organisms in this study. Within the 3D plots, the x-axis represents elution time, y-axis represents wavelength, and z-axis represents absorbance. The injection volume, growth media, number of peaks, peak wavelength, and elution time at the peak wavelength are listed above the plots for each organism. P. heparinus displayed a maximum absorbance at 450 nm with an elution time of 32.38 min (Figure 11). P. heparinus absorbs light in the indigo and blue region rendering a yellow substance. A peak wavelength at 480 nm with an elution time of 33.84 min was found in P. sp. R20-19 (Figure 12). P. alluvionis displayed a maximum absorbance at 478 nm eluting at 33.47 min (Figure 13). Differently from the yellow strain, the pink strains, P. sp R20-19 and P. alluvionis, absorbs light in the blue and green region reflecting red/orange pigment.

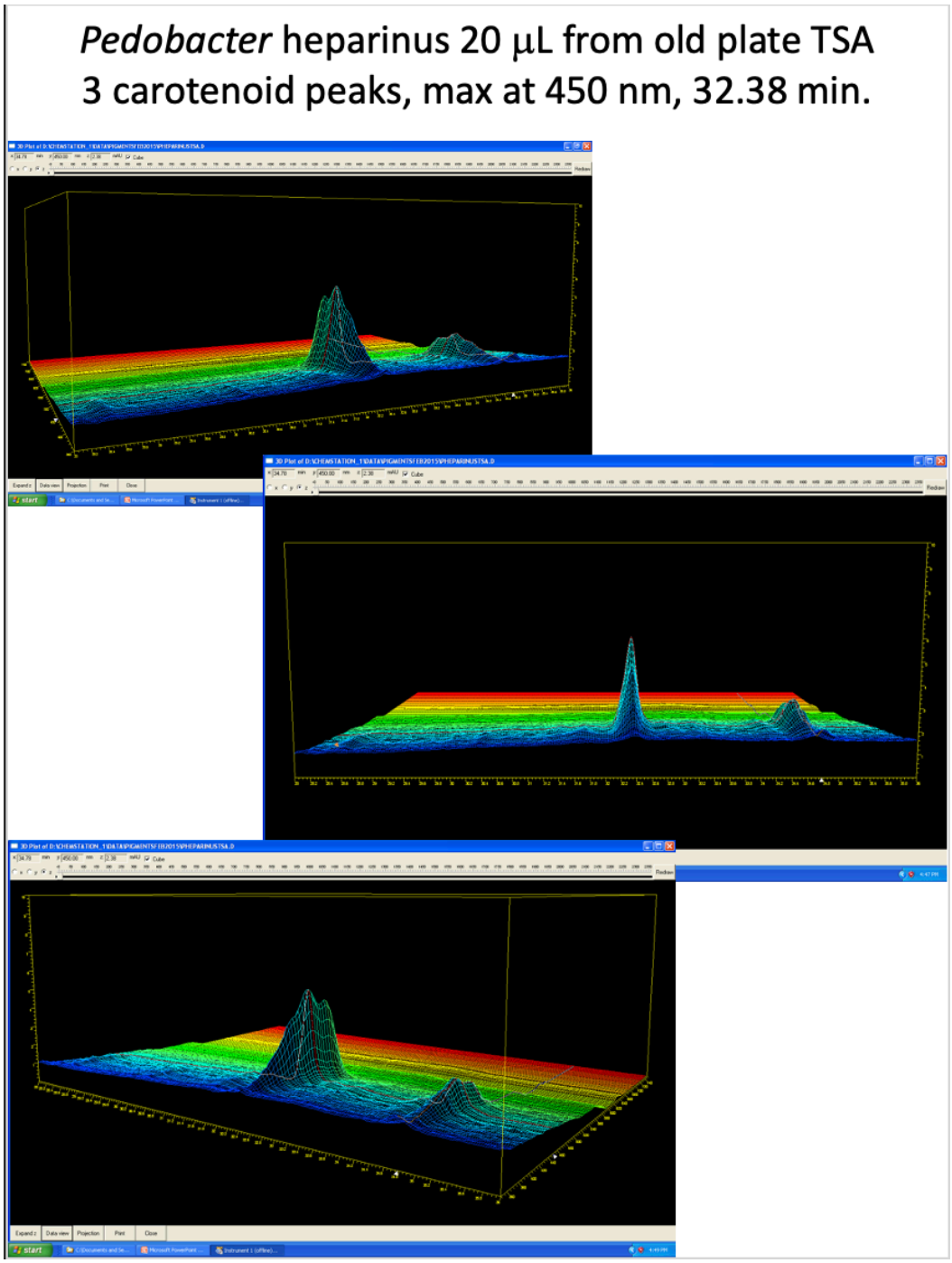

Figure 11. 3D plot of *P. heparinus* extracts in three different angles.

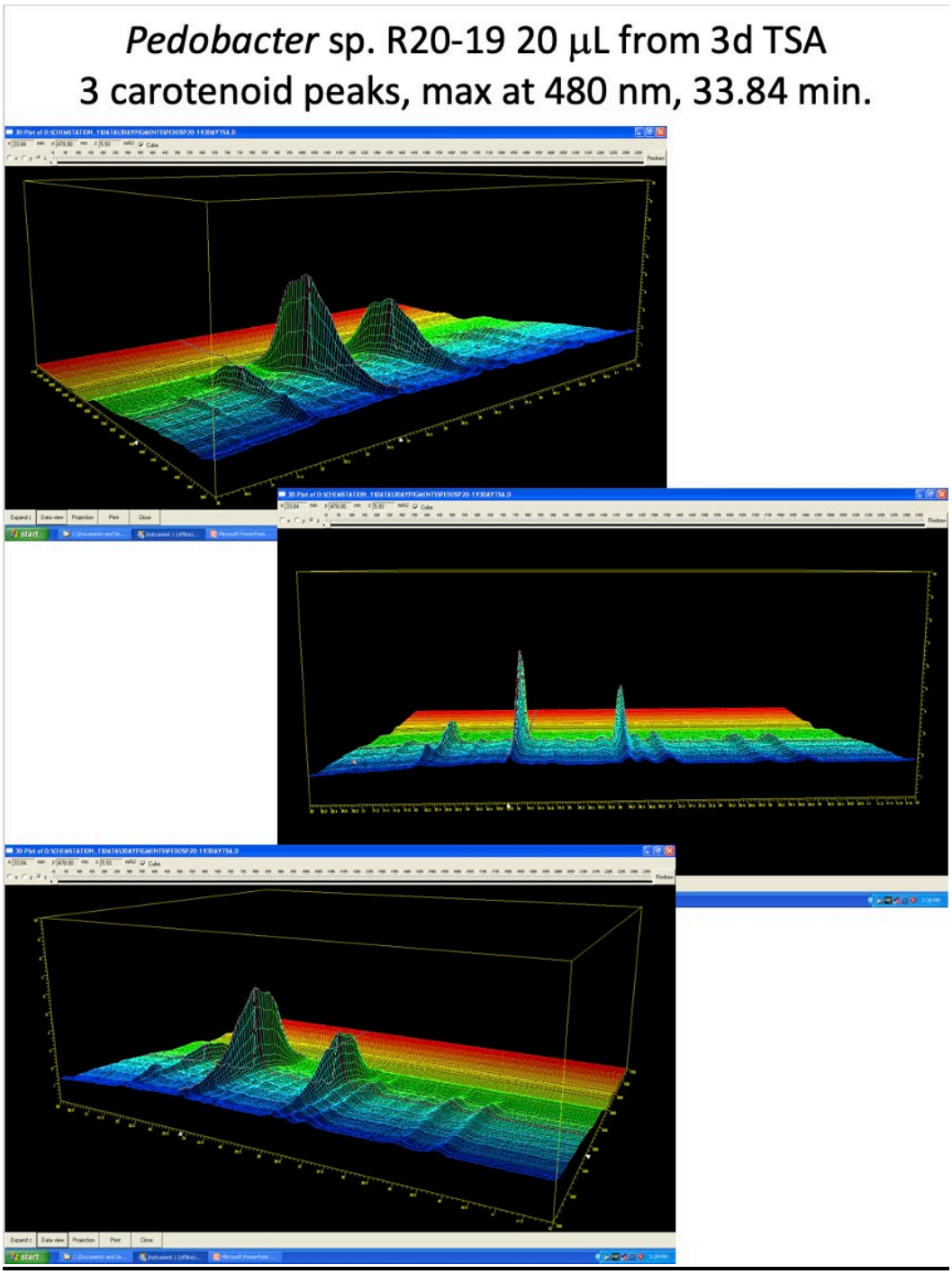

Figure 12. 3D plot of *P.* sp. R20-19 extracts in three different angles.

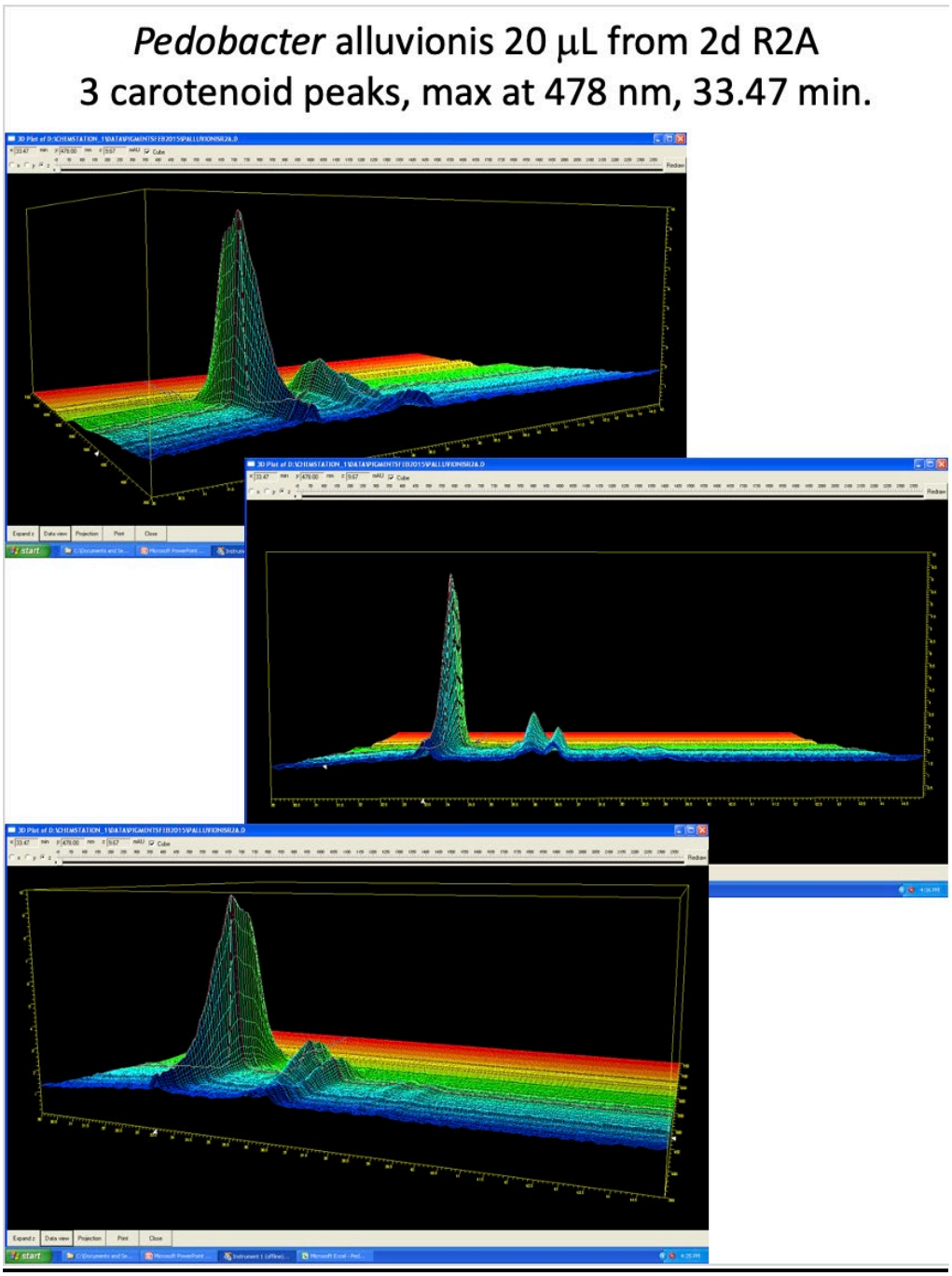

Figure 13. 3D plot of *P. alluvionis* extracts in three different angles.

Data acquisition by HPLC-MS generated chromatograms, UV-vis spectra, and mass spectrum data. For the chromatogram, acquisition time was trimmed to 0-3 min because the compounds eluted from the column within that time frame (Figure 14). The yellow species had three peaks eluting at approximately 0.7 min, 1.25 min, and 1.7 min with almost identically shaped chromatograms (Figure 14A and Table 4). The pink species displayed 5 peaks eluting at 0.72 min and 0.88 min (Figure 14B and Table 4). The compounds in P. sp. MR2016-19 began to elude faster than in P. sp. R20-19, therefore their elution times at peaks 3, 4, and 5 differ. For P. sp. R20-19, the elution time for peaks 3, 4, and 5 are: 0.98 min, 1.42 min, and 1.91 min (Figure 14B and Table 4). The elution time for peaks 3, 4, and 5 for P. sp. MR2016-19 are: 0.92 min, 1.3 min, 1.74 min (Figure 14B and Table 4). With differences in elution time between the two pink species, the chromatograms are still identical in shape. Distinctions in the chromatograms portray more pigments eluted from the pink species compared to the yellow species.

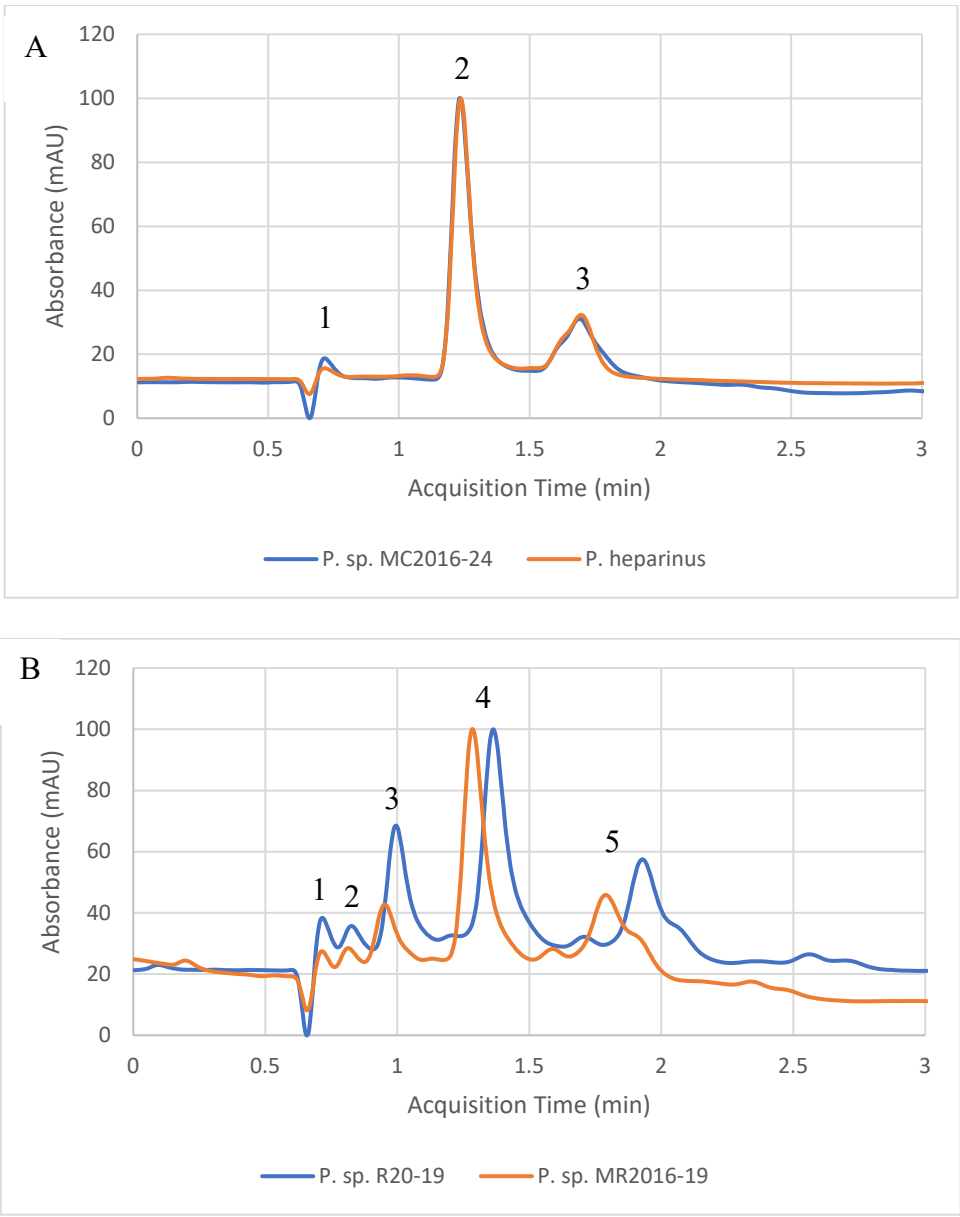

Figure 14. HPLC-DAD chromatogram at (A) 450 nm for yellow *Pedobacter* extracts and (B) at 480 nm for pink *Pedobacter* extracts. Peak numbers in the elution profile correspond to those in Table 4.

The absorbances obtained from the UV-vis spectra are similar to the absorbances in the 3D plots (Figure 11-13, 15). Maximum absorbances at 426 nm, 450 nm, and 478 nm was seen in the yellow strains (Figure 15A). The absorption spectra for the yellow strains resemble the absorption spectra and peak absorbances of β-carotene and zeaxanthin, therefore we hypothesize the mass spectrum will contain these two pigments (Figure 2 and 15A). For the pink strains, maximum wavelengths of 450 nm, 480 nm, and 502 nm was displayed (Figure 15B). The absorption spectra for the pink strains does not resemble astaxanthin, canthaxanthin, or any other carotenoid pigment, therefore the m/z in the MS should solidify the type of carotenoids. Both colors have the characteristic of carotenoids with three peaks present in the UV-Vis spectra, however the differences in the peak wavelengths and shape of the UV-Vis spectra show the distinct types of carotenoids possibly produced by the pink and yellow species (Figure 16).

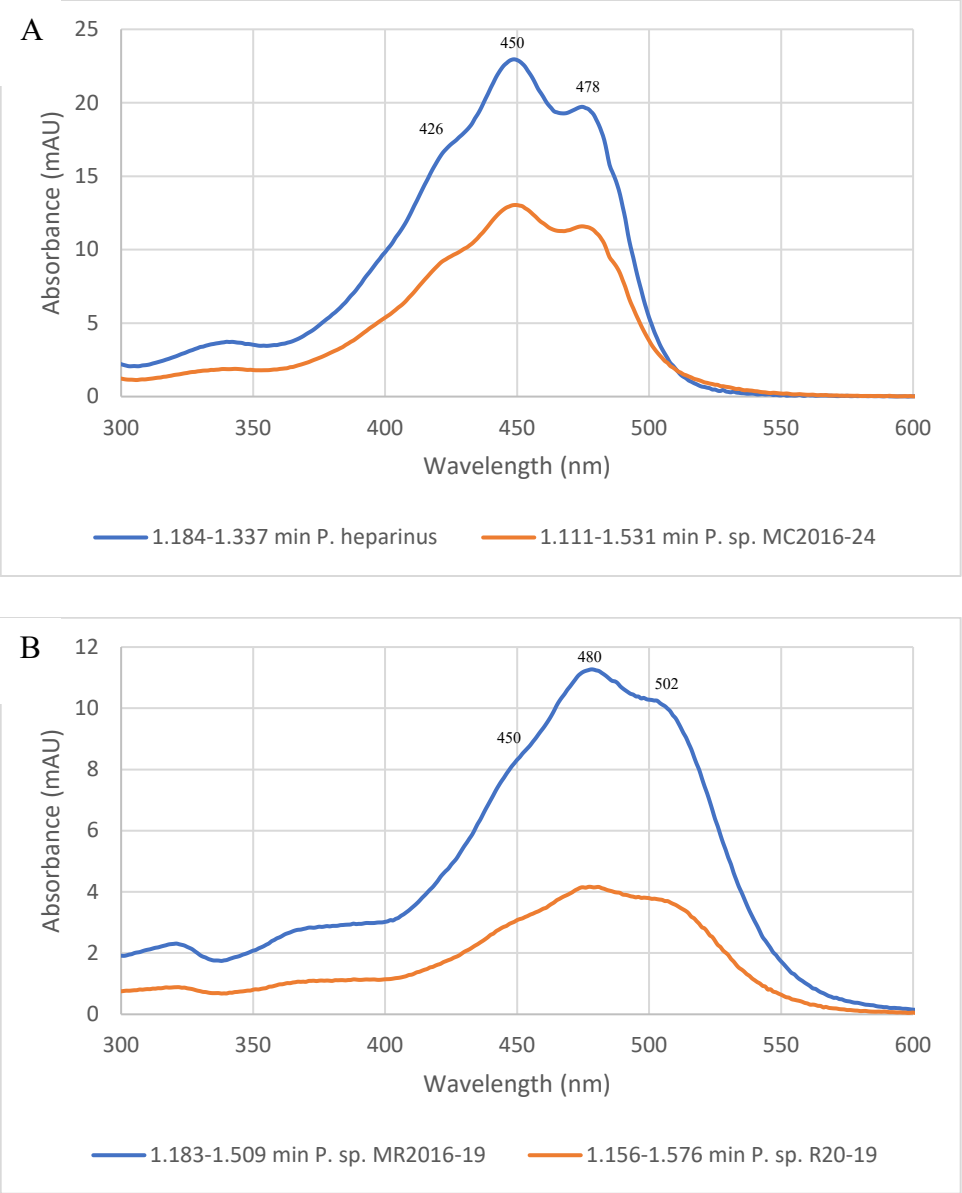

Figure 15. Absorption spectra of extracts from (A) the yellow *Pedobacter* at 450 nm and (B) the pink *Pedobacter* at 480 nm. Peak wavelengths correspond to those in Table 4. The range for the elution time at the peak wavelength is recorded in the key below the graph.

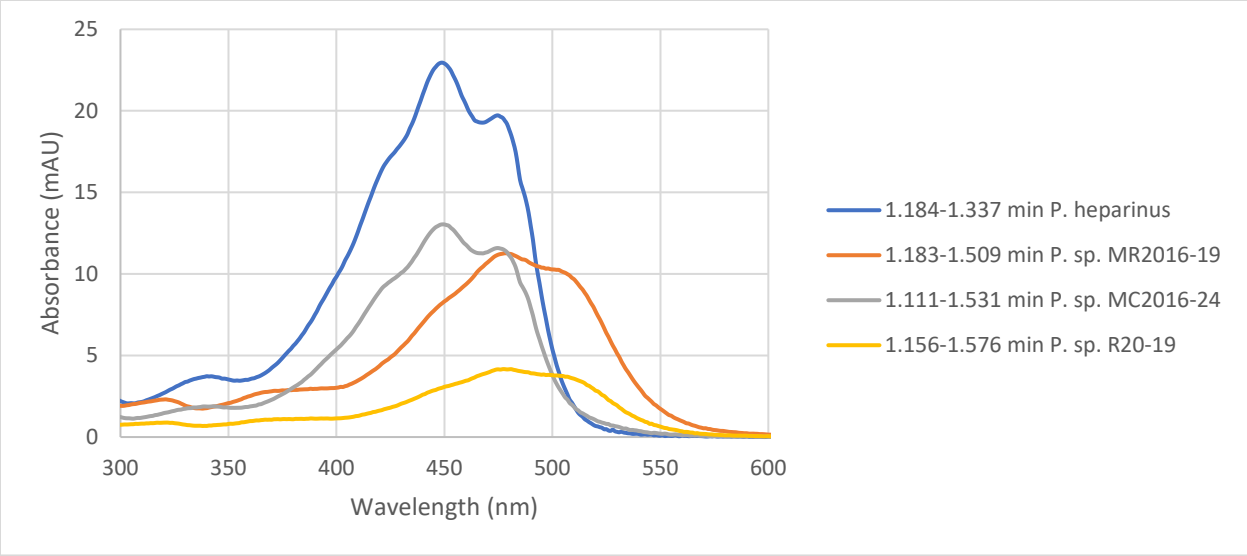

Figure 16. Absorption spectra of pink and yellow *Pedobacter*. The absorbance maxima are at 450 nm for the yellow organisms, and 480 nm for the pink organisms. The elution time and name of the organism is listed in the legend.

The yellow *Pedobacter* strains presented a molecular ion $[(C_{40}H_{56}O_2)+H]^+$ at m/z 569.3452 corresponding to the molecular weight of zeaxanthin or lutein, 568.88 g/mol.[21] The molecular formula of zeaxanthin and lutein is $C_{40}H_{56}0_2$. The difference in the two structures is the placement of the double bond in one ring, which examining the fragmentation pattern could decipher the exact carotenoid present. The characteristic APCI (+) MS pattern was 569.4 $[M+H]^+$, 551.4 $[M+H-18]^+$, and 477.4 $[M+H-92]^+$ aligning with the APCI (+) MS pattern of zeaxanthin.[7] The pink *Pedobacter* strains presented a molecular ion $[(M+H)+H]^+$ at m/z 592.4960 predicting this peak to be $C_{40}H_{48}O_4$ (Figure 18). This appears to be a precursor of astaxanthin, but instead with one double bond present in each of the two rings.

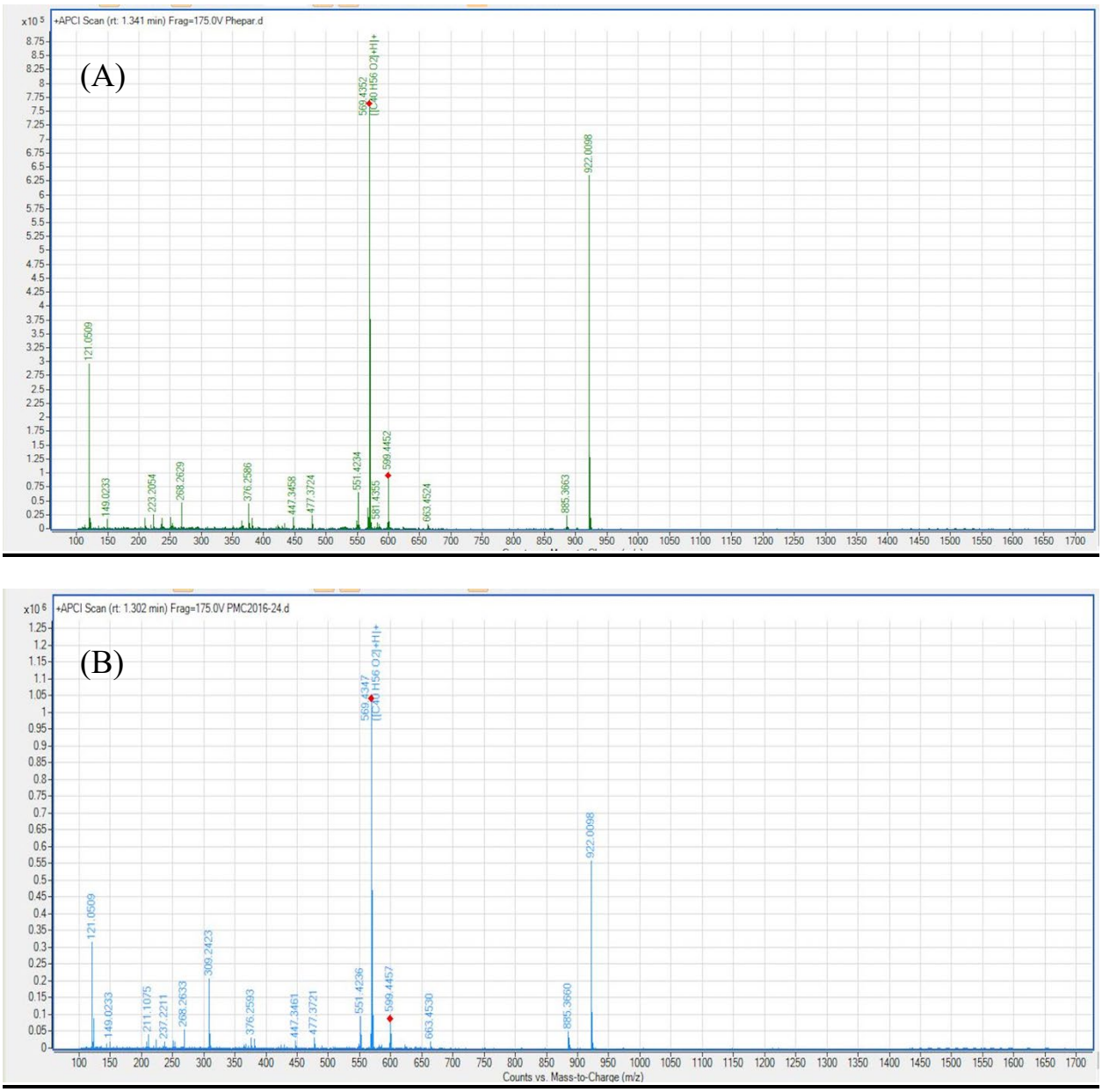

Figure 17. LC-MS spectra of (A) *P. heparinus* and (B) *P.* sp. MC2016-24 in positive APCI mode. The largest peak represents zeaxanthin.

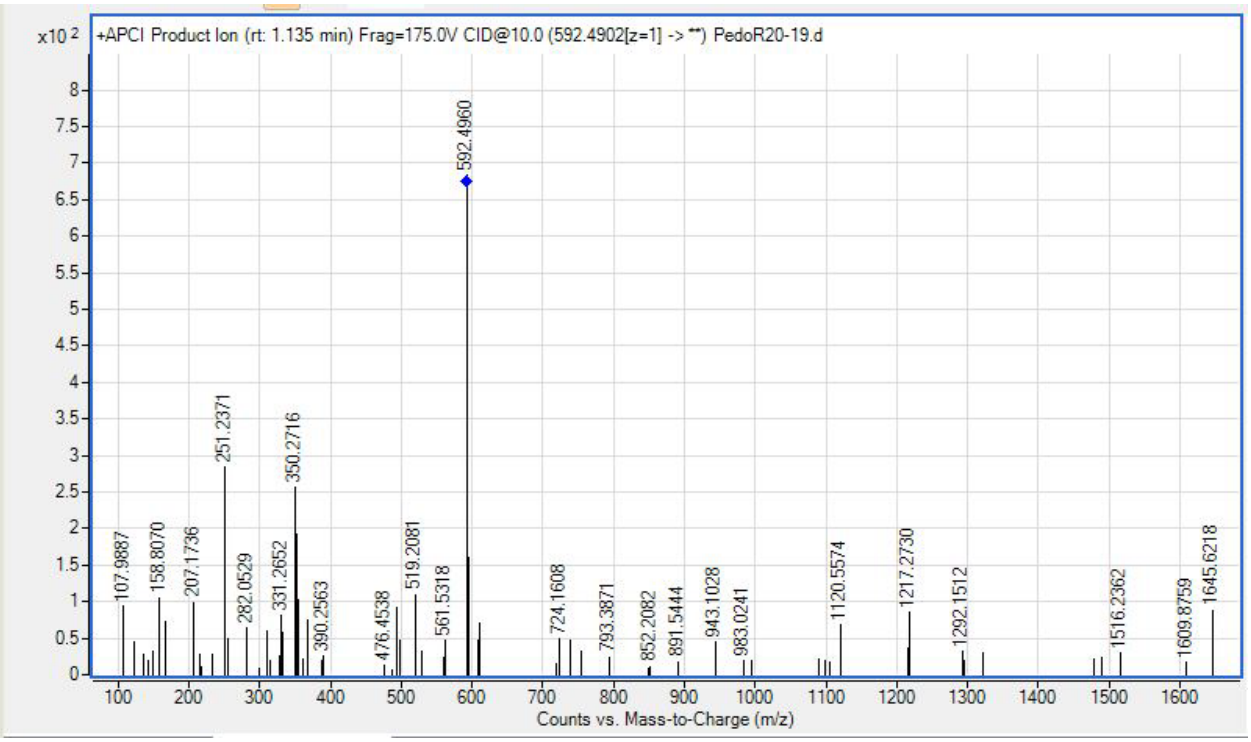

Figure 18. LC-MS spectra of the pink strains in positive APCI mode.

Conclusion

P. alluvionis, *P.* sp. MR2016-19, *P.* sp. R20-19, and the other pink *Pedobacter* strains should be reclassified into the new genus *Roseopedobacter* due to their genomic differences. This is supported by the GTDB and rpoB trees which both have similar branching patterns that show the divergence of the yellow strains from non-pink and non-orange strains. Also, both trees show that the pink strains share a common ancestor and cluster together as do the yellow strains. The AAI values shared between the pink strains were above 75, indicating that they should be classified into the same genus, *Roseopedobacter*. Moreover, the AAI values suggested that the pink and yellow strains belong to separate genera. The pink strains shared many genes collectively, as do the yellow strains as illustrated in the Venn Diagram. An obvious difference that will distinguish the new genus is pigmentation and carotenoid biosynthesis genes when

compared to the yellow pigmented *Pedobacter* strains. Thus, the β-carotene ketolase present within the pink strains should catalyze the reaction to generate pink-pigmented carotenoids, canthaxanthin and astaxanthin. As opposed to the yellow strains, which do not contain β-carotene ketolase, producing orange and yellow pigmented carotenoids, β-carotene and zeaxanthin. The UV-vis spectra revealed that the pink strains presented a peak wavelength at 480 nm rendering a pink color from absorbing light in the blue-green region. Contrastingly, the yellow strains absorb light in the violet-blue region and possess a peak wavelength at 450nm. Zeaxanthin was the yellow pigmented carotenoid present in the yellow *Pedobacter* species. The molecular ion seen in the MS of the pink strains appeared to be a precursor of astaxanthin with a molecular formula of $C_{40}H_{48}O_4$. Overall, this study discovered chemosystematic markers of carotenoids in *Pedobacter* because the study of these pigments is limited within this genus. Also, these pigments revealed its potential as a natural colorant, in hopes of substituting for synthetic dyes. Pedobacter species have the potential to produce carotenoids that can benefit industries as opposed to their utilization of harmful synthetic dyes.

Acknowledgements

The author thanks Joanne and Arthur Haberburger for financial support. and the Department of Biochemistry and Chemistry, especially Dr. Holly Bendorf, for maintaining and allowing use of the LC-MS instrument.

References

(1) Forgacs E.; Cserháti T.; Oros G. Removal of synthetic dyes from wastewaters: a review. *Environ Int*. **2004**, *30*, 953–971.

(2) Tuli H.S.; Chaudhary P.; Beniwal V.; Sharma A.K. 2014. Microbial pigments as natural color sources: current trends and future perspectives. *J Food Sci Technol*. **2014**, *52*, 4669–4678.

(3) Jiménez M.E.P.; Pinilla C.M.B.; Rodrigues E.; Brandelli A. 2018. Extraction and partial characterisation of antioxidant pigment produced by *Chryseobacterium* sp. kr6. *Nat Prod Res*. **2018**, *33*, 1541-1549.

(4) Venil C.K.; Khasim A.R.; Aruldass C.A.; Ahmad WA. Safety evaluation of flexirubin from Chryseobacterium artocarpi CECT 8497: Acute, sub-acute toxicity and mutagenicity studies. *Process Safety and Environmental Protection*. **2017**, *112*, 362-370.

(5) Joshi V.K.; Attri D.; Bala A.; Bhushan S. 2003. Microbial pigments. *Indian Journal of Biotechnology*. **2003**, *2*, 362-369.

(6) Henke N.A.; Peters-Wendisch P.; Wendisch V.F.; Heider S. C50 Carotenoids: Occurrence, Biosynthesis, Glycosylation, and Metabolic Engineering for their Overproduction. *John Wiley & Sons*. **2017**, 107-126.

(7) Vila E.; Hornero-Méndez D.; Azziz G.; Lareo Claudia.; Saravia V. Carotenoids from heterotrophic bacteria isolated from Fildes Peninsula, King George Island, Antarctica. *Biotechnology Reports*. **2019**, *21*.

(8) Rebelo B.A.; Farrona S.; Ventura M.R. Abranches R. Canthaxanthin, a Red-Hot Carotenoid: Applications, Synthesis, and Biosynthetic Evolution. Plants. *Basel*. **2020**, *9*, 1039.

(9) Chatzivasileiou, A.O.; Ward V.; Edgar S.M.; Stephanopoulos, G. Two-step pathway for isoprenoid synthesis. *PNAS*. **2019**, *116*, 506-511.

(10) Pasamontes, L.; Hug D.; Tessier, M.; Hohmann, H.; Schierle, J.; van Loon, A.P. Isolation and characterization of the carotenoid biosynthesis genes of Flavobacterium sp. strain R1534. *Gene*. **1997**, *185*, 35-41.

(11) Correa-Llantén, D.N.; Amenábar, M.J.; Blamey, J.M. Antioxidant capacity of novel pigments from an Antarctic bacterium. *J Microbiol.* **2012**, *50***,** 374–379.

(12) Altschul S.F.; Gish W.; Miller W.; Myers E.W.; Lipman D.J. Basic local alignment search tool. *J. Mol. Biol*. **1990**, *215*, 403-410.

(13) Aziz R.K.; Bartels D.; Best A.A.; DeJongh M.; Disz T.; Edwards R.A.; Formsma K.; Gerdes S.; Glass E.M.; Kubal M.; Meyer F.; Olsen G.J.; Olson R.; Osterman A.L.; Overbeek R.A.; McNeil L.K.; Paarmann D.; Paczian T.; Parrello B.; Pusch G.D.; Reich C.; Stevens R.; Vassieva O.; Vonstein V.; Wilke A.; Zagnitko O. The RAST Server: rapid annotations using subsystems technology. *BMC Genomics*. **2014**, *9*, 75.

(14) Kumar S.; Stecher G.; Tamura K. MEGA7: Molecular Evolutionary Genetics Analysis version 7.0 for bigger datasets. *Molecular Biology and Evolution*. **2016**, *33*, 1870-1874.

(15) Řezanka T.; & Olšovská J; Sobotka M.; Sigler K. The Use of APCI-MS with HPLC and Other Separation Techniques for Identification of Carotenoids and Related Compounds. *Current Analytical Chemistry*. **2009**, *5*, 1-25.

(16) Nishida Y; Adachi K; Kasai H; Shizuri Y; Shindo K; Sawabe A; Komemushi S; Miki W; Misawa N. Elucidation of a carotenoid biosynthesis gene cluster encoding a novel enzyme, 2,2'-beta-hydroxylase, from Brevundimonas sp. strain SD212 and combinatorial biosynthesis of new or rare xanthophylls. *Appl Environ Microbiol*. **2005**, *8*, 4286-96.

(17) Takaya T; Anan M; Iwata Koichi. Vibrational relaxation dynamics of β-carotene and its derivatives with substituents on terminal rings in electronically excited states as studied by femtosecond time-resolved stimulated Raman spectroscopy in near-IR. *Physical Chemistry Chemical Physics*. **2017**, 20.

(18) Castangia I; Manca M; Razavi S; Nácher A; Diez Sales O; Peris J; Allaw M; Terencio M; Usach I; Manconi M. Canthaxanthin Biofabrication, Loading in Green Phospholipid Vesicles and Evaluation of In Vitro Protection of Cells and Promotion of Their Monolayer Regeneration. *Biomedicines*. **2022**, *10*, 157.

(19) Juliadiningtyas A; Pringgenies D; Heriyanto H; Salim K; Radjasa OK; Shioi Y; Limantara L; Hardo T; Brotosudarmo P. Preliminary Investigation of the Carotenoid Composition of Erythrobacter sp. Strain KJ5 by High-Performance Liquid Chromatography and Mass Spectrometry. *Philippine Journal of Science*. **2018**, 147.

(20) Jiao G; Hui J; Burton I; Thibault MH; Pelletier C; Boudreau J; Tchoukanova N; Subramanian B; Djaoued Y; Ewart S; Gagnon J; Ewart K; Zhang J. Characterization of Shrimp Oil from Pandalus borealis by High Performance Liquid Chromatography and High Resolution Mass Spectrometry. *Marine drugs*. **2015**, 13.

(21) Kim S; Chen J; Cheng T; Gindulyte A; He J; He S; Li Q; Shoemaker BA; Thiessen PA; Yu B; Zaslavsky L; Zhang J; Bolton EE. PubChem in 2021: new data content and improved web interfaces. *Nucleic Acids Res*. **2019**, *47*, D1388-D1395.

(22) Kaiser P; Geyer R; Surmann P; Fuhrmann H. LC-MS method for screening unknown microbial carotenoids and isoprenoid quinones. *J Microbiol Methods*. **2012**, *88*, 28-34.

(23) Lee I; Ouk KY; Park SC; Chun J. OrthoANI: An improved algorithm and software for calculating average nucleotide identity. *Int J Syst Evol Microbiol*. **2016**, *66*, 1100-1103.

(24) Nicholson CA; Gulvik CA; Whitney AM; Humrighouse BW; Bell ME; Holmes B; Steigerwalt AG; Villarma A; Sheth M; Batra D; Rowe LA; Burroughs M; Pryor JC; Bernardet J; Hugo C; Kämpfer P; Newman JD; McQuiston JR. Division of the genus Chryseobacterium: Observation of discontinuities in amino acid identity values, a possible consequence of major

extinction events, guides transfer of nine species to the genus Epilithonimonas, eleven species to the genus Kaistella, and three species to the genus Halpernia gen. nov., with description of Kaistella daneshvariae sp. nov. and Epilithonimonas vandammei sp. nov. derived from clinical specimens. *International Journal of Systematic and Evolutionary Microbiology*. **2020**, *70*, 4432-4450.

(25) Andrew G; Honors Thesis. *Lycoming College*. **2015**.

(26) Saitou N; Nei M. The neighbor-joining method: A new method for reconstructing phylogenetic trees. *Molecular Biology and Evolution*. **1987**, *4*, 406-425.

(27) Felsenstein J. (1985). Confidence limits on phylogenies: An approach using the bootstrap. *Evolution*. **1985**, *39*, 783-791.

(28) Kimura M. A simple method for estimating evolutionary rate of base substitutions through comparative studies of nucleotide sequences. *Journal of Molecular Evolution*. **1980**, *16*, 111-120.