Faculty of Science

ANALYZING THE *UREC* GENES IN *STREPTOMYCES* SP. ICC1 AND ICC4 FOR BIOMINERALIZATION POTENTIAL

2024 | FERNANDO BOUTHILLIER

B.Sc. HONOURS THESIS – BIOLOGY

THOMPSON RIVERS UNIVERSITY

ANALYZING THE *UREC* GENES IN *STREPTOMYCES* SP. ICC1 AND ICC4 FOR BIOMINERALIZATION POTENTIAL

by

FERNANDO BOUTHILLIER

A THESIS SUBMITTED IN PARTIAL FULFILLMENT

OF THE REQUIREMENTS FOR THE DEGREE OF

BACHELOR OF SCIENCE (HONS.)

in the

DEPARTMENT OF BIOLOGICAL SCIENCES

(Cellular, Molecular, and Microbial Biology)

This thesis has been accepted as conforming to the required standards by:

Naowarat Cheeptham (Ph.D.), Thesis Supervisor, Dept. Biological Sciences

Joanna Urban (Ph.D.), Secondary Supervisor, Dept. Biological Sciences

Natasha Ramroop Singh (Ph.D.), External examiner, Dept. Biological Sciences

Dated this 18th day of April 2024, in Kamloops, British Columbia, Canada

ABSTRACT

Streptomyces sp. ICC1 and ICC4 were originally isolated from the Iron Curtain Cave which is an extreme environment featuring low levels of light and nutrients. In previous studies, other microbes isolated from the Iron Curtain Cave showed biomineralization potential through microbiologically induced carbonate precipitation (MICP) using urea hydrolysis which has the potential to be used for biocement production. The objective of this study is to determine the MICP potential of ICC1 and ICC4 through urease activity and calcium carbonate production as well as the presence of urease and carbonic anhydrase genes. *Streptomyces* sp. ICC1 and ICC4 that were grown on urea agar exhibited a pink colour after one-week incubation at 15°C indicative of active urease activity. ICC1 and ICC4 showed crystal-like formations when incubated on B4 agar at 15°C but could not be confirmed to be calcium carbonate. Polymerase chain reaction (PCR) primers were designed and tested to amplify all *ureC* gene copies in DNA extracted from ICC1 and ICC4. Gel electrophoresis and Sanger sequencing results confirmed the successful amplification of all *ureC* copies as expected from previous sequencing results from Gosse et al. in 2019. Bioinformatic analysis revealed that none of the *ureC* gene copies were associated with a complete seven-gene operon as seen in most ureolytic bacteria, which suggests the potential for limited urease activity than initially expected. Four carbonic anhydrase genes were detected in the ICC1 and ICC4 sequencing data suggesting the potential for a synergistic effect with urease enzymes.

Thesis Supervisor: Professor Naowarat Cheeptham (Ph.D.)

ACKNOWLEDGEMENTS

I would like to thank UREAP for funding my research leading up to this project. Thank you to Dr. Naowarat Cheeptham and Dr. Joanna Urban for supervising my project and Dr. Natasha Ramroop Singh for being a member of my examining committee. I would like to thank Dr. Don Nelson for his work as a secondary supervisor, and along with Dr. Eric Bottos, for assistance with molecular techniques that I used to conduct this research.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

BLAST	Basic Local Alignment Search Tool
CA	carbonic anhydrase
$CaCO_3$	calcium carbonate
CO_2	carbon dioxide
DNA	deoxyribonucleic acid
dNTP	deoxyribonucleotide triphosphate
GC	guanine/cytosine
ICC	Iron Curtain Cave
ICC1	*Streptomyces* sp. ICC1
ICC4	*Streptomyces* sp. ICC4
MICP	microbiologically induced carbonate precipitation
NCBI	National Centre for Biotechnology Information
PCR	Polymerase Chain Reaction
qRT-PCR	Qualitative Reverse Transcription-Polymerase Chain Reaction
RNA	ribonucleic acid

INTRODUCTION

Cave Bacteria

Caves are extreme habitats characterized by a low amount of nutrients, light, and temperature (Cheeptham et al. 2013). *Streptomyces* sp. ICC1 and ICC4 were isolated from the Iron Curtain Cave located near Chilliwack, BC by Ghosh et al. in 2017. Microorganisms living in these habitats are adapted to thrive in oligotrophic conditions. This is accomplished through the gain of genes that would help these organisms survive and the loss of non-essential genes (Iranzo et al. 2019). Iron Curtain Cave features a variety of speleothems that come in three types, soda straws, bacon, and popcorn-like structures (Ghosh et al. 2017). Both soda straws and bacon-type speleothems are made of calcium carbonate (Demeny et al. 2016). The temperature of the cave ranges from 4 to 12°C depending on the time of year. The cave also contains high iron content sediment and limestone structures (Ghosh et al. 2017). A previous study analyzed bacteria isolated from Iron Curtain Cave speleothems and found a variety of species that exhibited ureolytic ability. Two novel species, *Sphingobacterium* sp. PCS056 and *Pseudoarthrobacter* sp. SSS035 were also found to exhibit crystal formation ability consistent with microbiologically induced carbonate precipitation (Koning et al. 2022). Based on these previous results, it is hypothesized that the genes of the *urease* operon could be under selective pressures from the cave environment and evolve in a way that could have a beneficial industrial impact.

Ureolytic bacteria have been previously identified from other cave environments outside of the Iron Curtain Cave. From limestone caves in Sarawak, Malaysia, ureolytic species of *Sporosarcina*, *Bacillus*, and *Pseudogracilibacillus* were identified and able to produce $CaCO_3$ crystals (Omoregie et al. 2019). Strains of *Bacillus simplex*, *Rhodococcus degradans*, and *Strenostrophomonas maltophilia* were isolated from the Baralda Cave in Hungary and found to

induce amorphous calcium carbonate production (Enyedi et al. 2020). While less studied than other species, a variety of *Streptomyces* strains isolated from the Springtails' Cave in Belgium were specifically identified to exhibit urease activity and $CaCO_3$ formation (Maciejewska et al. 2017).

Streptomyces sp. ICC1 and ICC4

Streptomyces sp. ICC1 and ICC4 are Gram-positive, filamentous bacteria with high GC-content genomes. High GC content is typically associated with increased thermal stability and a broader temperature tolerance range in bacteria (Šmarda et al. 2014). *Streptomyces* have a unique life cycle consisting of vegetative growth, aerial hyphae, and sporulation. Due to the filamentous growth of *Streptomyces*, cells aggregate when grown in liquid culture which inhibits the ability to monitor growth by typical methods such as optical density. Alternative measurement forms such as packed cell volume, dry weight, and wet weight have been employed for the measurement of *Streptomyces* growth but are limited to larger quantities. An alternative method has been proposed using the adsorption and desorption of methylene blue to correlate with the amount of bacterial growth in aggregating bacteria (Fischer and Sawers 2013). Both ICC1 and ICC4 are novel and have only been isolated from the Iron Curtain Cave. ICC1 and ICC4 underwent whole genome sequencing in 2019 and were both found to show high similarity with *Streptomyces lavendulae* strain CCM 3239 (Gosse et al. 2019). However, in comparison to most *Streptomyces*, ICC1 and ICC4 showed increased amounts of certain genes such as biosynthetic gene clusters that are predicted to give them an advantage in the cave environment (Gosse et al. 2019). As another potential cave adaptation, both ICC1 and ICC4 were each found to contain five copies of the *ureC* gene in comparison to non-cave *Streptomyces* species such as *S. coelicolor* and *S. venezuelae*

which each contain two copies of the gene. The increased copy number implies the potential for ICC1 and ICC4 to have increased ureolytic activity making them worth investigating.

Biocement

The manufacturing of cement is a highly energy-intensive process that results in a large amount of CO_2 emissions which get trapped in the earth's atmosphere leading to global warming. It is estimated that cement production accounts for approximately 5-8% of total CO_2 emissions in the world (Kajeste and Hurme 2016) per year. Alternate methods of cement production are attractive options to lessen the impact of climate change. A variety of ureolytic bacteria have been studied for the production of biocement through the process of microbiologically induced carbonate precipitation (MICP). Urea hydrolysis starts with the breakdown of urea into carbamic acid and ammonia. Carbamic acid breaks down on its own to form carbonic acid and ammonia (Prajapati et al. 2023). Following this, in water, ammonia forms ammonium and hydroxide ions while carbonic acid forms bicarbonate and hydrogen ions. The release of hydroxide ions increases the pH which shifts the bicarbonate equilibrium towards the formation of carbonate ions. When calcium is found in the environment of these bacteria, the carbonate formed from the breakdown of urea can bind with the positively charged free calcium ions that are attracted to the negatively charged bacterial cell wall to form strong calcium carbonate crystals (Chuo et al. 2020).

$$CO(NH_2)_2 + H_2O_{(l)} \rightarrow NH_2COOH_{(aq)} + NH_{3\,(aq)}$$

$$NH_2COOH_{(aq)} + H_2O_{(l)} \rightarrow NH_{3\,(aq)} + H_2CO_{3\,(aq)}$$

$$H_2CO_{3\,(aq)} \leftrightarrow HCO_3^-{}_{(aq)} + H^+{}_{(aq)}$$

$$2NH_{3\,(aq)} + 2H_2O_{(l)} \rightarrow 2NH_4^+{}_{(aq)} + 2OH^-{}_{(aq)}$$

$$HCO_3^-{}_{(aq)} + H^+{}_{(aq)} + 2NH_4^+{}_{(aq)} + 2OH^-{}_{(aq)} \rightarrow (CO_3)^{2-} + 2NH_4^+{}_{(aq)} + 2H_2O_{(l)}$$

$$Ca^{2+}{}_{(aq)} + \text{Cell Surface} \rightarrow \text{Cell Surface} - Ca^{2+}{}_{(aq)}$$

$$\text{Cell Surface} - Ca^{2+}{}_{(aq)} + (CO_3)^{2-}{}_{(aq)} \rightarrow \text{Cell} - CaCO_{3\,(s)}$$

When calcium carbonate crystals are produced in the presence of soil or sand particles, bonds can form between the substances leading to the formation of biocement (reference). Previous studies have used *Sporosarcina pasteurii* for biocement production based on a variety of characteristics. It is an alkaliphile, so it thrives in the high pH conditions required for MICP. *S. pasteurii* is also non-pathogenic, making it non-harmful to other organisms (Wu et al. 2021).

urease **Operon**

The *urease* operon found in *S. pasteurii* consists of three genes that encode the trimeric structure of the enzyme. *ureA* encodes the gamma subunit, the beta subunit is encoded by *ureB*, and the alpha subunit is encoded by *ureC* gene. Four accessory genes are also typically found in the *urease* operon. *ureD, ureF*, and *ureG* typically form a secondary trimeric complex, while *ureE* is involved in the transportation of nickel ions to the active site of the enzyme (Wu et al. 2021).

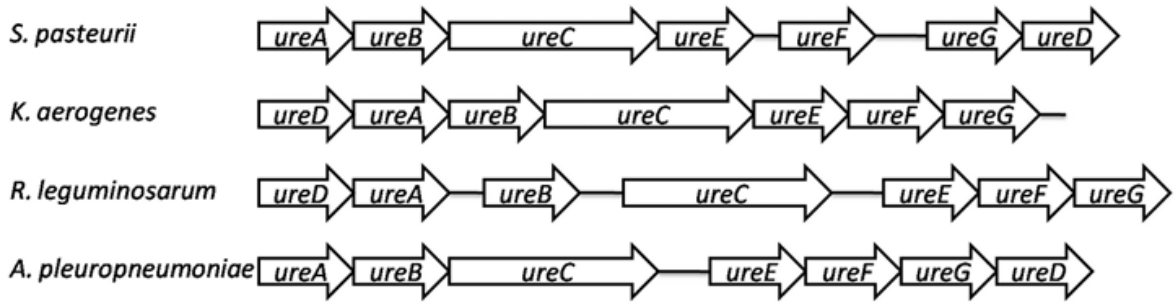


Figure 1. The common structure of the *urease* operon as seen across four different bacterial species featuring seven total genes (Pei and Hu 2023).

In a previous study, the *ureE* gene was knocked out in *Klebsiella aerogenes* resulting in a 50% decrease in the overall urease activity (Mulrooney et al. 2005). The importance of the subunit

encoded by the *ureDFG* genes is also observed in cloning studies where urease activity is improved by the co-expression of *ureABC* and *ureDFG* complexes (Liu et al. 2017). The regulation of *urease* operon is also typically dependent on the availability of nitrogen in the environment. Previous studies with ureolytic *Klebsiella pneumoniae* have observed that expression of the *urease* operon is independent of the concentration of urea present in the environment but is instead based on the need for nitrogen as a source for growth. Nitrogen assimilation control protein (NAC) has been studied to be responsible for the nitrogen-based control in *Klebsiella pneumoniae*, but a variety of regulatory factors have been found in other species such as *Bacillus subtilis* (Bender 2010) (Wray et al. 1997).

Carbonic anhydrase enzymes have also been predicted to play a role in calcium-carbonate precipitation synergistically with urease enzymes (Clarà Saracho and Marek 2024). Carbonic anhydrase enzymes work to hydrate CO_2 into carbonic acid which then dissociates to bicarbonate and hydrogen ions as consistent with the ureolytic pathway (Zheng and Qian 2020). A study analyzing *Bacillus megaterium*, a bacterium carrying both carbonic anhydrase and urease enzymes, found increased $CaCO_3$ precipitation ability in comparison to isolated solutions of either enzyme suggesting a synergistic effect (Dhami et al. 2014). Increased expression of the *ureC* gene was also found in *S. pasteurii* in the presence of bicarbonate which is the product that carbonic anhydrase produces from hydrating CO_2 (Clarà Saracho and Marek 2024).

Objective

The objective of this study is to understand the biomineralization potential of *Streptomyces* sp. ICC1 and ICC4 and to correlate this activity with the *ureC* genes to understand the impact of the cave environment on this activity. To accomplish this, three smaller objectives were formed for this study. The first objective was to determine the urease activity and calcium carbonate crystal

ability of ICC1 and ICC4 through growth on urea agar and B4 media. The second objective was to determine if the *ureC* genes are present as expected based on the previous sequencing results, PCR primers were developed and tested to successfully amplify the *ureC* genes. Additionally, because PCR primers could be used in future studies such as gene expression analysis with qRT-PCR, PCR amplification conditions were optimized. Lastly, given their importance to MICP, urease accessory genes and carbonic anhydrase genes were detected bioinformatically using translated BLAST to determine if and where these genes are present in the genomes of ICC1 and ICC4.

MATERIALS AND METHODS

Culturing of *Streptomyces* sp. ICC1 and ICC4

Bacteria were first streaked from a -80°C glycerol stock on Hickey-Tresner agar plates using a 4-way streaking technique and left in a refrigerated incubator to grow at 15°C. Isolated colonies were obtained approximately 4-6 days with ICC1 typically growing more quickly. Gram-staining was performed on samples to confirm the presence of Gram-positive filamentous bacteria.

Once isolated colonies were confirmed, they were individually cultured into 100 ml of liquid Hickey-Tresner media stored in 250 ml Erlenmeyer flasks. Flasks were then set to incubate at 15°C and shaken at 150 RPM for 4-6 days to obtain stationary growth and adequate biomass for DNA extraction.

Methylene Blue Growth Curve Assay

The methylene blue growth curve assay was carried out based on the method by Fischer and Sawers with exceptions as follows (Fischer and Sawers, 2013). Absorbance readings were performed by a NovaSpec II spectrophotometer. 200 ml of HT-media was prepared in a 500 ml Erlenmeyer flask. Each flask was inoculated with an isolated colony from a HT-media plate

previously incubated at 15°C. Once inoculated flasks, were placed in an incubator shaker at 15°C and shaken at 150 RPM. 1.8 ml samples were collected in 2 ml centrifuge tubes at each time interval for analysis by the growth assay. Flasks were swirled before collection to evenly distribute cells within the media. Initial centrifugations were performed at 14,000 RPM for 4 minutes to form a stable pellet. All media was removed by pipetting before adding the methylene blue adsorption dye. No glass beads were used in the protocol. Shaking incubations at 80°C were performed in 800 ml of water in a 1 L beaker with a 3 cm magnetic stirrer to create a whirlpool at 400 RPM. Samples were first mixed by vortex and then placed in a 4-tube foam rack that was added to the 80°C for 10 minutes. After incubation, samples were cooled to room temperature and centrifugation was performed for 2 minutes at 14,000 RPM and 80 µl of the resulting supernatant was used for testing adsorption. An additional centrifugation step was performed after testing for adsorption to remove the remaining supernatant. For ICC4 only, a wash with distilled water was also performed due to the unstable pellet preventing the complete removal of the adsorption dye. Following the addition of 1.5 ml of distilled water, washes were performed with distilled water instead of a low-concentration methylene blue solution. Lastly, 100 µl of the final supernatant was used to test for absorbance after desorption. The assay was performed in triplicate and absorbance readings were carried out twice per sample to minimize variation.

Culturing on Urea Agar Plates

Urea agar plates were made following the manufacturer's instructions [BBLTM Urea Agar Base Concentrate 10× (BD, Franklin Lakes, New Jersey, United States]. Isolated colonies were streaked on individual plates and incubated for seven days at 15°C. *Escherichia coli* was used as a negative control while *Klebsiella pneumoniae* was used as a positive control to confirm the proper preparation of the plates.

Culturing on B4 Media

B4 media agar was prepared following the protocol from Marvasi et al. in 2012. 50 ml of media was poured into deep Petri dishes to prevent drying over the long incubation. Liquid cultures were also made the same as plates, but without agar, and stored as 20 ml solutions in 50 ml glass tubes. Modified B4 media was also made that included urea in place of yeast extract at the same concentration. Plates were stored at 8°C and checked for calcium carbonate crystal formation under a dissecting microscope after four weeks of incubation as consistent with previous research on Iron Curtain Cave bacteria from Koning et al. in 2022. Liquid cultures were also checked after 4 weeks by pipetting 50 µl from the bottom of each tube onto a slide. Observations were done under light microscopy at 400x magnification. Two polarizing filters were used to observe calcium carbonate crystals. One was placed on top of the slide of interest while the other was placed on top of the light source. As light passes through the first filter, it becomes polarized. When the filter above the slide is rotated 90 degrees out of place relative to the filter above the light source, no light can pass through the top filter. Liquid crystals such as calcium carbonate have the ability to rotate the plane of polarized light in a way that allows light to pass through the top filter. As a result, light patches are present through a microscope in the presence of these crystals.

DNA Extraction

DNA Extractions were performed from 1.8 ml of confluent liquid culture using the QIAGEN DNeasy Microbial Kit following manufacturers instruction. The cell lysis step was increased in length by three minutes due to the toughness of ICC1 and ICC4 cells. DNA quality and quantity was determined using agarose gel electrophoresis and Nanodrop spectrophotometer.

Polymerase Chain Reaction Conditions for *ureC* Gene Copies

PCR was performed to obtain amplicons of all the *ureC* genes from *Streptomyces* sp. ICC1 and ICC4. Five sets of primers were created to amplify all five copies of the gene present in both species' genomes. Primers were designed and tested in silico using Primer-BLAST before ordering (Table 1). PCR reactions were set up in 50 µl reactions following the protocol from the QIAGEN Taq PCR Core Kit. PCR running conditions were as follows, an initial denaturation was run at 95°C for four minutes. This was followed by 35 cycles of denaturation for 30 seconds at 95°C, primer annealing for 45 seconds at a variable temperature, and extension for one minute at 72°C. After the 35 cycles were complete, a final extension was run for five minutes at 72°C. After completion of the PCR reaction, 5 µl of PCR product was mixed with 2 µl of 5x DNA Loading Dye and run on a 100 ml 1.0 % agarose gel. Gel electrophoresis was run at 100 volts until adequate separation of fragments was visible. Annealing temperatures were initially predicted using Primer-BLAST and adjusted based on experimental results.

Table 1. Nucleotide sequences of the PCR primers designed to target all copies of the *ureC* gene across ICC1 and ICC4.

Primer name:	**Sequence (5' -> 3')**
*ureC*1F	GGCCGCATCGGAGAATCCTG
*ureC*1R	GAGCGCCCAGTTGATCAGTC
*ureC*2F	ATCAAGGACGGGTTCATCGC
*ureC*2R	TGAAGTGGATGTGGGTGTCG
*ureC*3F	TATCGGTGGCGTTCGTCG
*ureC*3R	AGTTGCGGAGCATGTCCTTC
*ureC*4F	GAGTACGCCTCCGTCTTCG
*ureC*4R	ACGATCATGCCGTAGTGCTC
*ureC*5F	CCATCCACTCGTACCACACC
*ureC*5R	AACTCCTTGCCCAGTTCGAG

Note. F = Forward Primer, R = Reverse Primer

DNA Purification and Sequencing

PCR products were purified using the QIAquick PCR Purification Kit to remove DNA Polymerase, dNTPs, and contaminating salts. Elutions were performed in 30 μl of Tris-HCl at pH 8. Purified products were then added individually to 8-well PCR strip tubes for Sanger sequencing at UBC Sequencing + Bioinformatics Consortium. 20 tubes were set up following the User-Prepared, Pre-Mixed protocol. Two samples were sent for each sequence, one containing forward primer and one containing reverse primer. *ureC*1-3 fragments were diluted with distilled water while *ureC* 4 and 5 fragments were diluted in Tris-HCl. Sequencing results were returned in FASTA file format with a chromatogram for each sequence to verify quality.

Bioinformatic Analysis

Bioinformatic analysis was carried out using Nucleotide and Protein BLAST. Individual FASTA files from the sequenced products were compared to the NCBI database first using

MegaBLAST, followed by a search using BLASTn (Nucleotide BLAST) if the initial search was unsuccessful. Accessory genes were found through the annotations deposited in the NCBI database. Translated BLAST was performed on annotated urease accessory genes to examine similarity to other species and determine if any unannotated regions of interest were present in the ICC1 and ICC4 genomes.

RESULTS

Growth Curve for ICC1 and ICC4

The absorbance reading at 660 nm for the methylene blue dye was measured to be 2.280. After 24 hours of incubation at 15°C, ICC4 did not show a visible pellet after initial centrifugation, so adsorption and desorption measurements were not carried out. A small pellet was visible after 48 hours, but adsorption readings showed a negligible difference in comparison to blank, so desorption was not carried out (Figure 1). After 56.5 hours, the change in absorbance after adsorption was also negligible, so desorption steps were not carried out. From 72 to 97 hours of incubation a decrease in the absorbance after adsorption was observed and an opposite correlation was observed after desorption (Figure 1). While a decrease in the adsorption/desorption ratio was observed, the decrease was not drastic as would be expected with exponential growth.

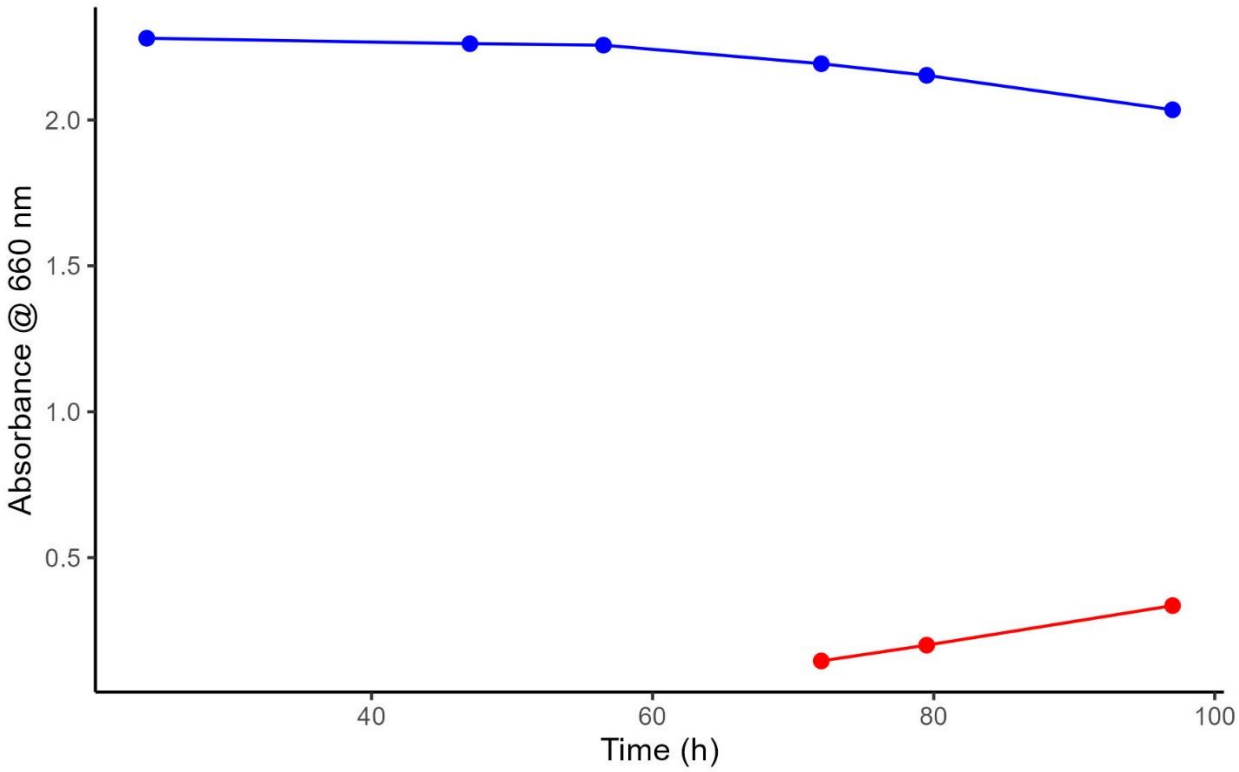


Figure 2. Growth curve results for *Streptomyces* sp. ICC4 after 97 hours of incubation in 200 ml Hickey-Tresner broth at 15°C. Absorbance after adsorption = blue, Absorbance after desorption = red

After 24 hours of incubation at 15°C, ICC1 did not show a visible pellet after initial centrifugation so adsorption and desorption absorbance measurements were not carried out. After 45 hours a significant decrease was observed in the absorbance measurement after adsorption, so the desorption steps were carried out and the supernatant was measured (Figure 2). A significant decrease was again observed after 52 hours in the absorbance value after adsorption with a parallel increase in absorbance after desorption (Figure 2). However, after 68 hours, oddly an increase in absorbance was observed after adsorption, and a decrease was observed after desorption in comparison to the measurements at 52 hours (Figure 2).

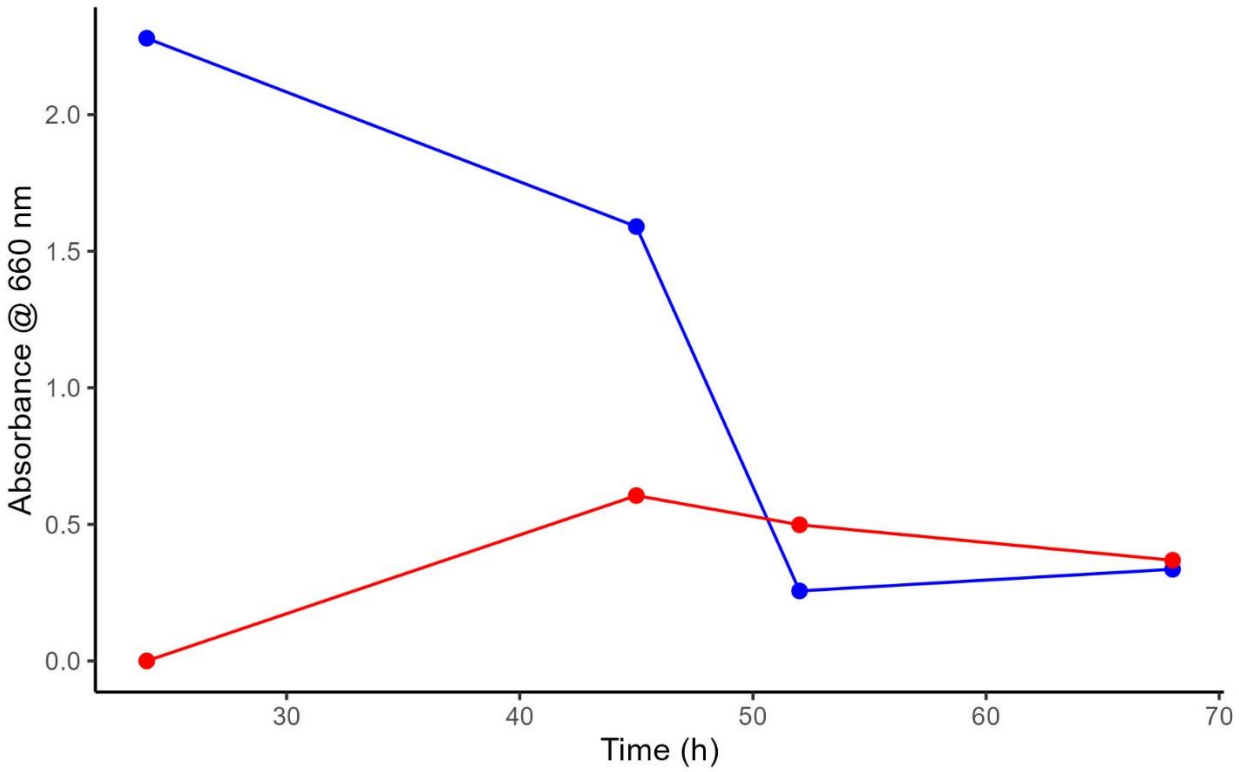


Figure 3. Growth curve results for *Streptomyces* sp. ICC1 after 68 hours of incubation in 200 ml Hickey-Tresner broth at 15°C. Absorbance after adsorption = blue, Absorbance after desorption = red

Urease Activity Test on Phenol Red Urea Agar

Both *Streptomyces* sp. ICC1 and ICC4 showed urease activity when grown on phenol red urea agar. A pink colour was observed after a one-week incubation at 15°C. Gram-staining performed after the incubation showed a Gram-positive and filamentous phenotype confirming the sole presence of *Streptomyces* on the plates. The negative control plate showed no colour change while the positive control turned pink after a one-week incubation (Figure 3).

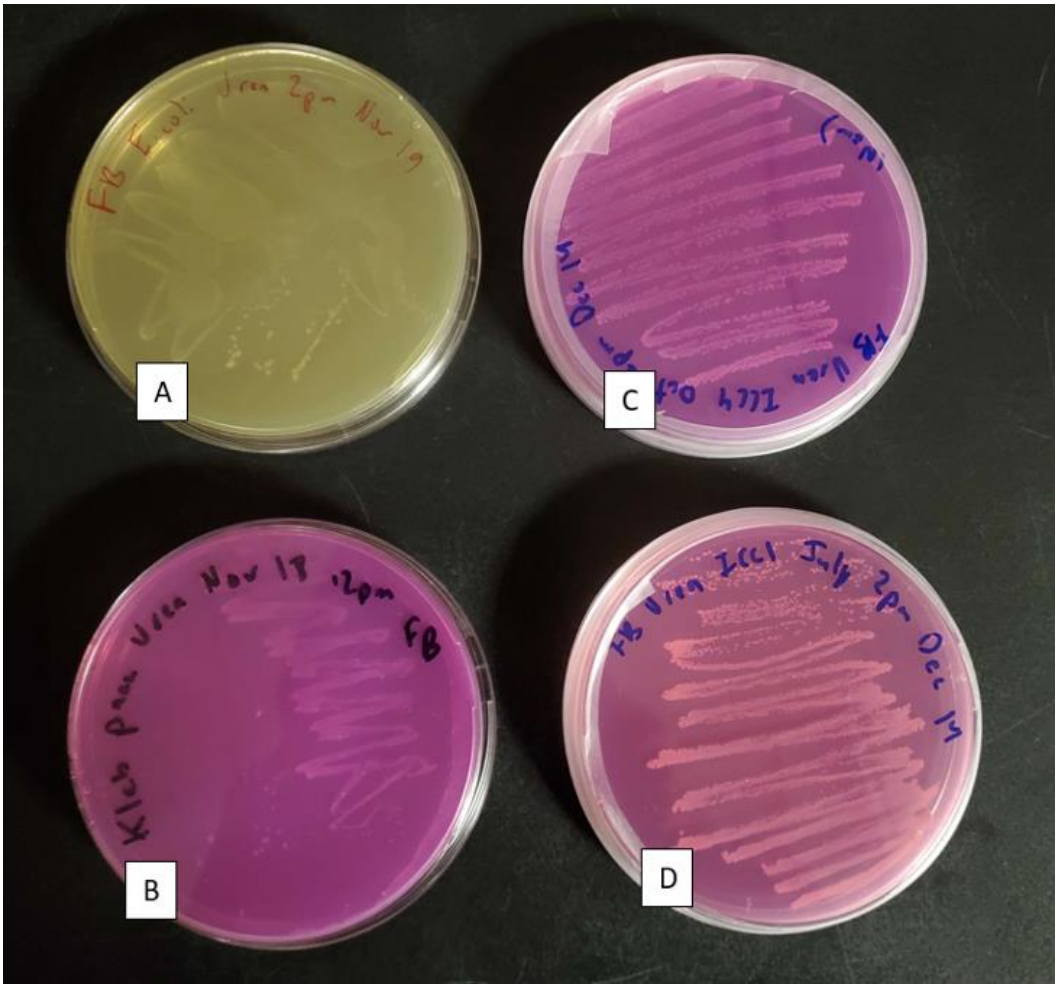


Figure 4. Urea agar plates after a one-week incubation at 15°C. A = *Escherichia coli* (negative control), B = *Klebsiella pneumoniae* (positive control), C = *Streptomyces* sp. ICC4, D = *Streptomyces* sp. ICC1

Calcium Carbonate Precipitate Ability on B4 Media

Both ICC1 and ICC4 showed significant growth on the default B4 media plates and liquid broth after four weeks of incubation at 8°C. Small crystal-like structures were visible in the media of both species in the agar plates (Figure 4).

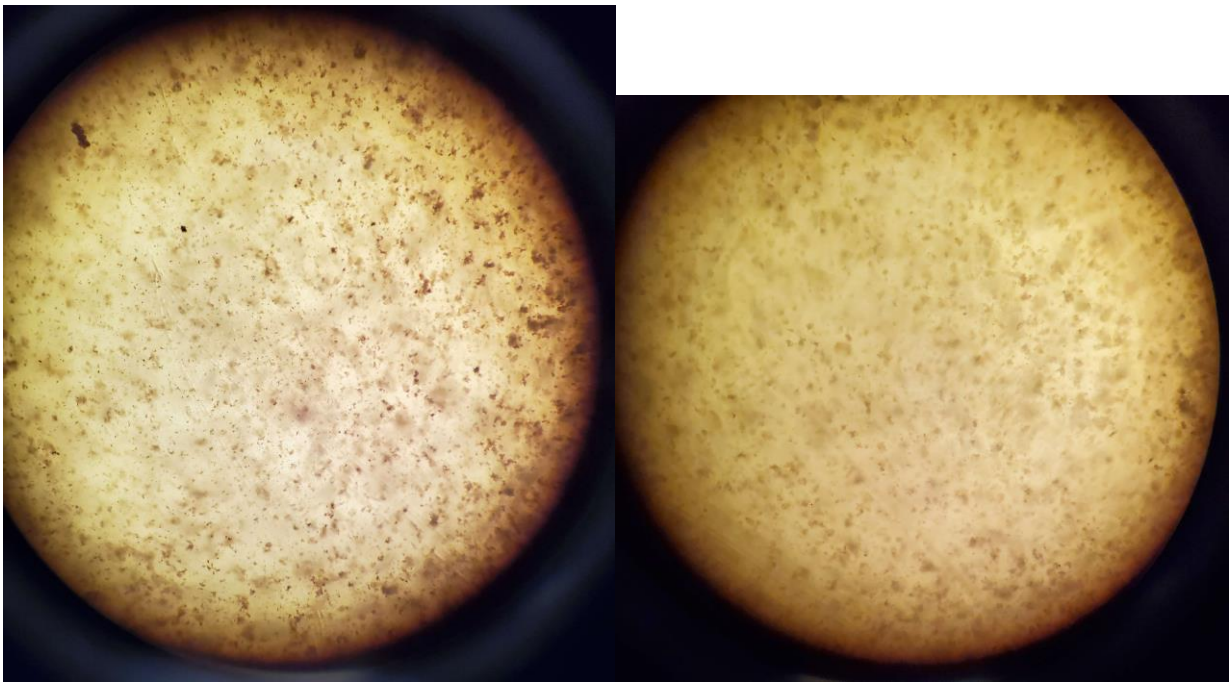


Figure 5. Similar crystal-like structures were observed in the B4 calcium-infused media for both ICC1 and ICC4 after a four-week incubation at 8°C. Left = ICC4, Right = ICC1

For observation of crystals in liquid media, polarized films were used to specifically filter out light sources outside of those emitted by liquid crystals. Only one ICC4 sample showed a significant light pattern that would be expected for calcium carbonate crystal formation (Figure 5).

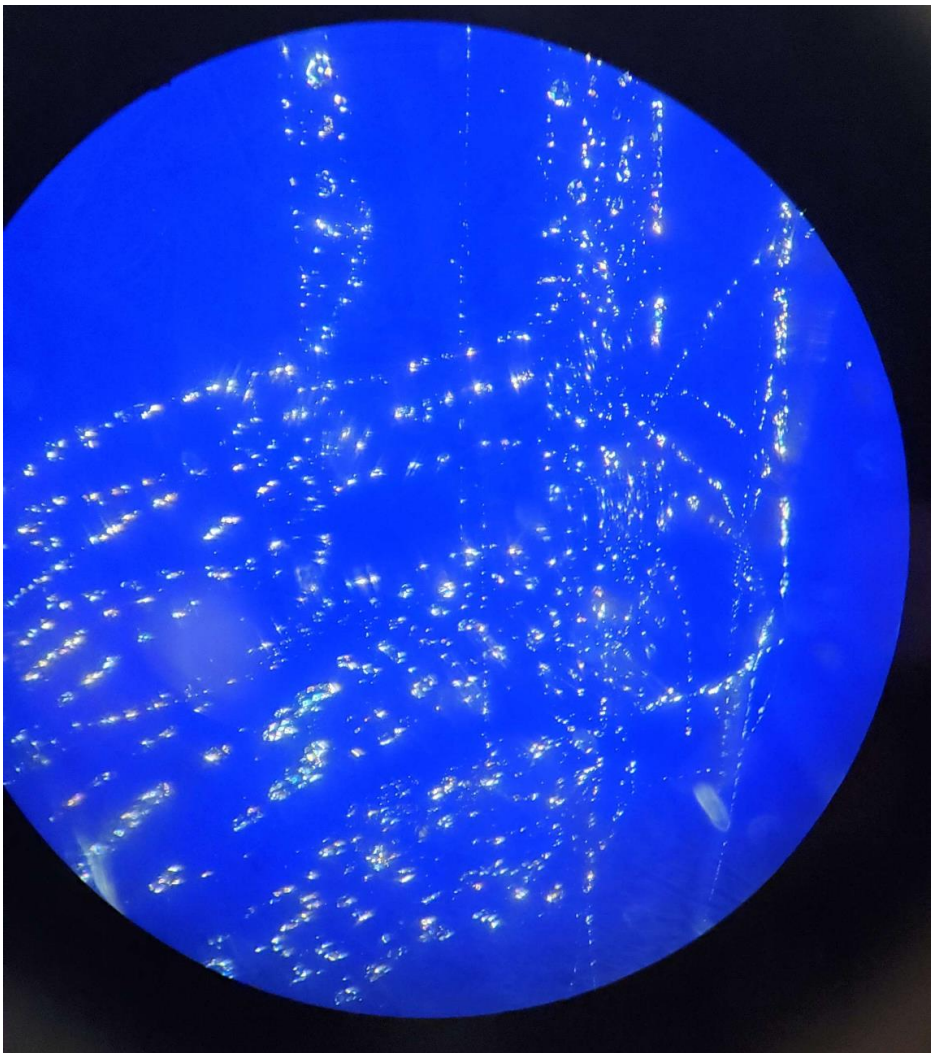


Figure 6. Scattered zones of light are visible under 400x magnification using light microscopy after polarizing filters were applied to observe potential calcium carbonate crystals by blocking out other light sources on ICC4 B4 calcium-infused liquid media after a four-week incubation at 8°C.

B4 Media agar plates created with urea as the alternative nitrogen source in place of yeast extract proved to be unsuccessful at growing ICC1 and ICC4 colonies after four weeks of incubation at 15°C.

Optimal PCR Conditions for Amplifying the *ureC* Gene Copies

All PCR reactions were originally performed with an annealing temperature of 54°C. *ureC*5 showed high amplification at this temperature, so adjustments were not made for this set. *ureC*1 was amplified with an annealing temperature of 56°C and high amplification of expected size was also observed for this set, so further adjustments were not made (Figure 6).

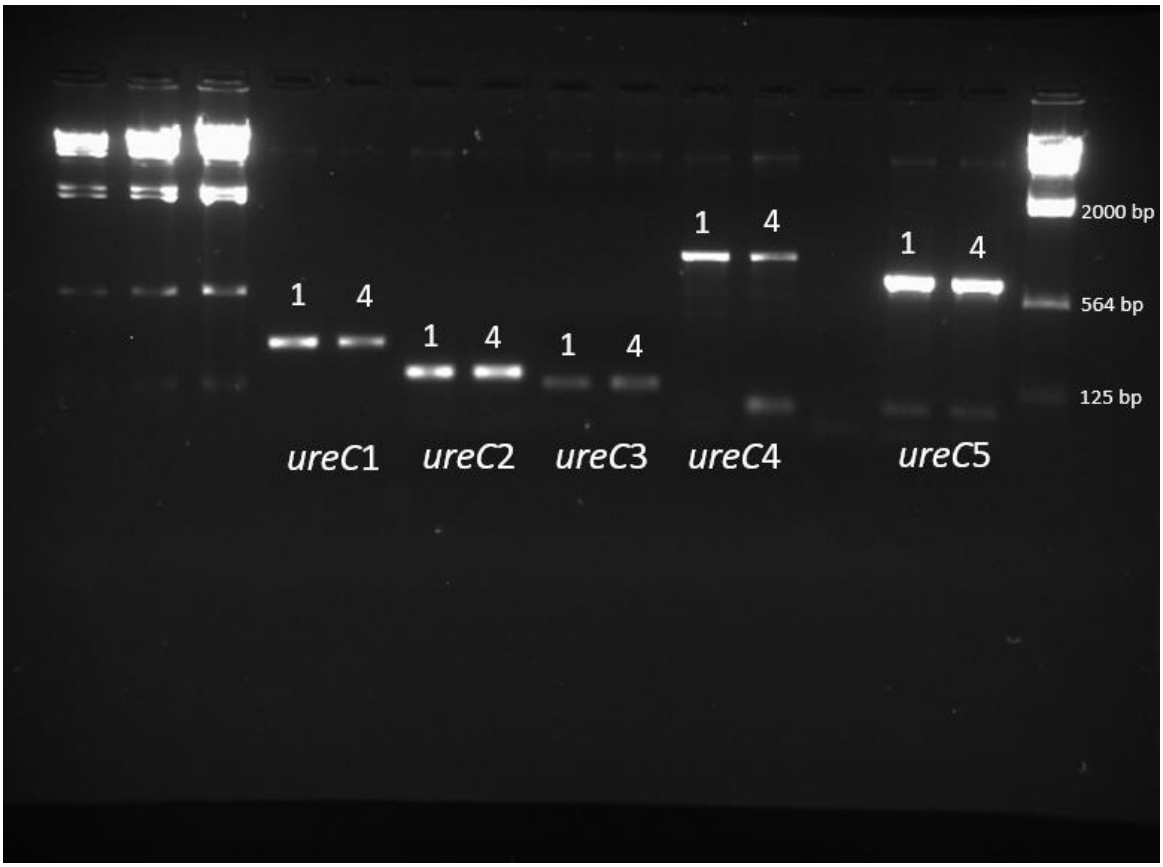


Figure 7. Gel electrophoresis analysis of all *ureC* PCR products after amplification with an annealing temperature of 54°C. An annealing temperature of 56°C was used for *ureC*1 due to the increased predicted melting temperature.

The predicted melting temperature of *ureC*2 primers was 60°C. Based on this, a gradient PCR was set up for the range of 55.7 to 62.0°C. The best amplification was observed at 55.7°C with an increase in primer-dimer as temperature increased (Figure 7). Amplification was lower with

*ureC*3, so a lower temperature gradient from 52.0 to 58.1°C was used as non-specific amplification was not a concern. An opposite effect from *ureC*2 was observed where amplification increased, and primer-dimer decreased as the temperature increased (Figure 7). A *ureC*4 gradient was set up for 55.7 to 62°C based on the predicted melting temperature of 60°C and the desire to remove non-specific amplification. Only the PCR reaction at 60°C showed good amplification with no products of interest being observed at all at 61.2°C and 62°C. High amounts of primer-dimer were observed for all *ureC*4 amplifications (Figure 7).

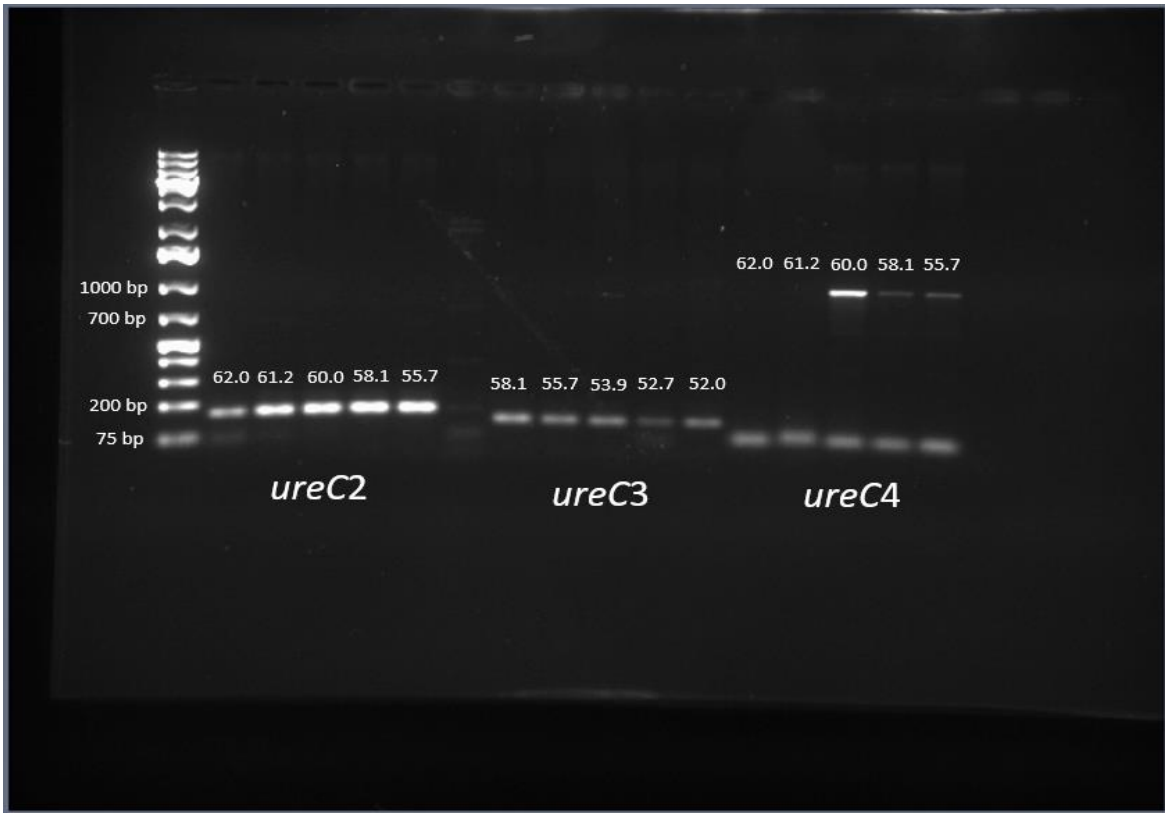


Figure 8. Gel electrophoresis analysis of PCR products from gradient PCR using primers *ureC*2-4 on ICC1 DNA. Annealing temperatures used appear above each lane in degrees Celsius.

Further optimization of PCR amplification was performed on ICC4 DNA to determine if the same PCR conditions were optimal. Given the similarity in DNA sequences, more narrow ranges were selected for *ureC*2 and *ureC*4 as they showed good amplification. A range of 54.5 to 58.2°C was used for *ureC*2 and a higher range of 58.2 to 61°C was used for *ureC*3 and *ureC*4 as they displayed the poorest amplification previously. Amplification appeared to improve slightly as temperature increased for *ureC*3 while the opposite trend was observed with *ureC*2 (Figure 8). As consistent with Figure 7, only one successful amplicon was visible for *ureC*4, however it appeared at 61°C as opposed to 60°C (Figure 8).

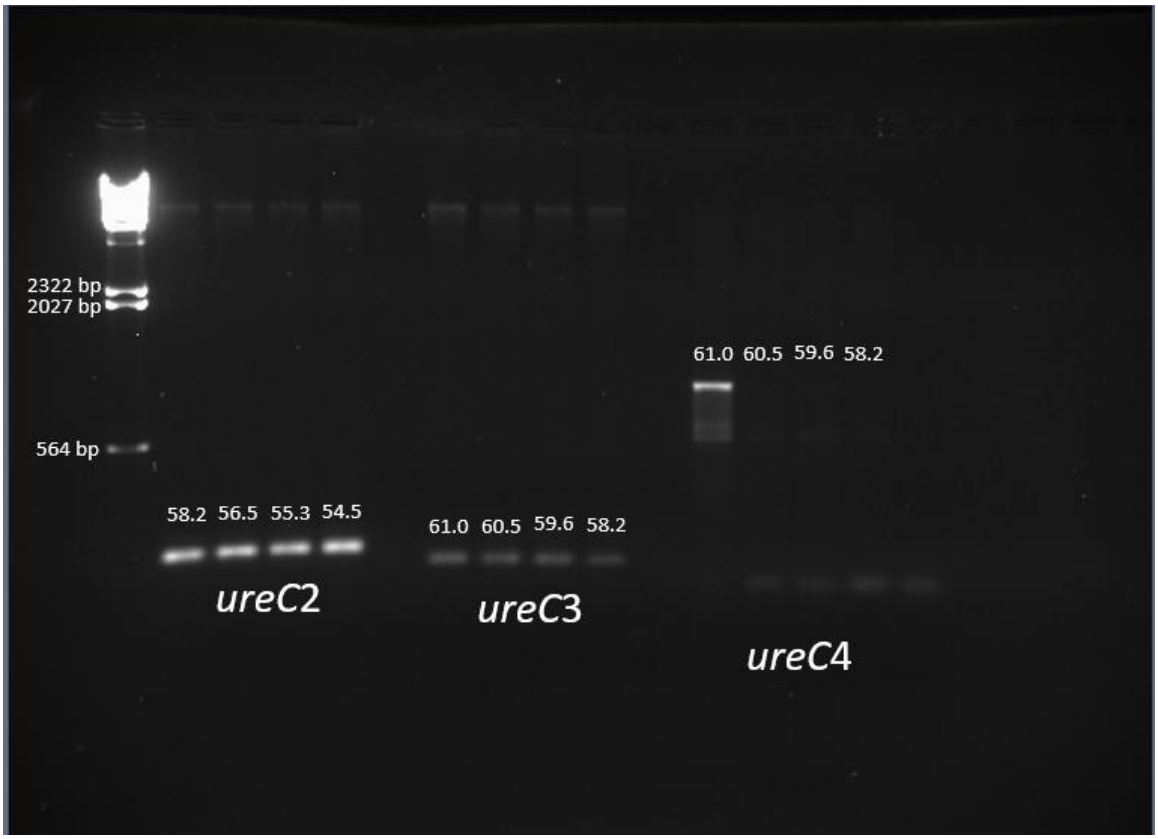


Figure 9. Gel electrophoresis analysis of PCR products from gradient PCR using primers ureC2-4 on ICC4 DNA. 30 cycles of PCR amplification were performed here in comparison to the 35 of other PCR reactions. Annealing temperatures used appear above each lane in degrees Celsius.

Sanger Sequencing of the *ureC* PCR Products

After PCR purification, successfully recovered products were shipped for Sanger Sequencing. FASTA sequences were compared to the NCBI database using BLAST to detect similarity. ICC1 and ICC4 showed the highest similarity hits for all sequenced products (Table 2). Chromatogram results indicated high quality sequences outside of the initial base pairs for most sequences. ICC1*ureC*2F showed poor chromatogram quality, indicating that sequence results for this sample are unreliable.

Table 2. Nucleotide BLAST analysis of the sequencing results. The highest hit for each FASTA sequence was summarized.

PCR Primers:	**Query Coverage:**	**Percent Identification:**	**Query length (bp):**	**Genome region (bp):**	**Expected Size (bp):**
ICC1*ureC*1F	100%	100%	216	6507302 to 6507517	259
ICC4*ureC*1F	91%	93%	247	6377377 to 6377596	259
ICC1*ureC*2F	44%	90%	238	6507304 to 6507410	137
ICC4*ureC*2F	94%	95%	135	898971 to 899094	137
ICC1*ureC*3F	84%	96%	99	4697689 to 4697771	102
ICC4*ureC*3F	76%	93%	97	8032472 to 8032543	102
ICC1*ureC*4F	96%	99%	903	992054 to 992919	904
ICC4*ureC*4F	96%	99%	904	1176926 to 1177791	904

ICC1*ureC*5F	91%	99%	705	5821626 to 5822271	710
ICC4*ureC*5F	94%	99%	702	2468329 to 2468990	710
ICC1*ureC*1R	91%	99%	262	6507328 to 6507569	259
ICC4*ureC*1R	95%	97%	257	6377319 to 6377557	259
ICC1*ureC*2R	88%	100%	129	670103 to 670216	137
ICC4*ureC*2R	85%	97%	139	898938 to 899052	137
ICC1*ureC*3R	85%	92%	93	4697731 to 4697808	100
ICC4*ureC*3R	79%	91%	104	8032514 to 8032591	100
ICC1*ureC*4R	91%	99%	911	992107 to 992939	903
ICC4*ureC*4R	94%	99%	909	1176956 to 1177810	903
ICC1*ureC*5R	86%	99%	678	5821738 to 5822321	710
ICC4*ureC*5R	93%	99%	715	2468365 to 2469026	710

Detection of Carbonic Anhydrase and Urease Accessory Genes

Through bioinformatic analysis, it was found that none of the *ureC* gene copies were associated with a classic completed *urease* operon consisting of seven genes. All five copies contained an associated gamma and beta subunit in a nearby region to form a proper urease core

enzyme, but all sets lacked accessory genes (Table 3). *ureC*1 and *ureC*2 both contained operons containing five of the seven typical urease genes. *ureC*5 contained the only *ureF* gene but was lacking all the other accessory genes. *ureC*3 and 4 were lacking all accessory genes. Results were consistent across ICC1 and ICC4 (Table 3). Two sequences were also detected to be *ureD* and *ureG* genes, but no core urease enzyme genes were found near this region (Table 3).

Table 3. Bioinformatic results detecting the presence of urease accessory genes in the NCBI database sequence of ICC1 and ICC4.

Sample:	***ureA*** **Sequence Start**	***ureB*** **Sequence Start**	***ureC*** **Sequence Start**	***ureD*** **Sequence Start**	***ureE*** **Sequence Start**	***ureF*** **Sequence Start**	***ureG*** **Sequence Start**	**Urea Transporter**
ICC1*ureC*1	6510083	6508764	6506961	6504544	-	-	6505581	6510552
ICC1*ureC*2	668140	669339	669827	673025	-	-	672360	667126
ICC1*ureC*3	4699753	4699372	4697576	-	-	-	-	-
ICC1*ureC*4	993473	993095	991293	-	-	-	-	-
ICC1*ureC*5	5823485	5823162	5821442	-	-	5820660	-	-
ICC4*ureC*1	6374503	6375702	6376206	6379382	-	-	6378687	6373374
ICC4*ureC*2	896975	898174	898662	901863	-	-	901195	895961
ICC4*ureC*3	8034143	8034155	8032359	-	-	-	-	-
ICC4*ureC*4	1178344	1177966	1176164	-	-	-	-	-
ICC4*ureC*5	2470184	2469858	2468144	-	-	2467362	-	-
Unknown ICC1	-	-	-	3766329	-	-	3765610	-
Unknown ICC4	-	-	-	5117064	-	-	5116348	-

Four carbonic anhydrase genes were found in ICC1 and ICC4. The sizes of each enzyme were inconsistent throughout the four copies. ICC1 and ICC4 did however show the same respective enzyme sizes, but they were encoded in separate areas of the genome (Table 4).

Table 4. Bioinformatic results detecting the presence of carbonic anhydrase genes in the NCBI database sequence of ICC1 and ICC4.

Sample:	**Sequence Length (AAs):**	**Genome Location: (Start position in bp)**
ICC1 CA1	245	3839356
ICC1 CA2	210	8917880
ICC1 CA3	195	8332101
ICC1 CA4	243	390762
ICC4 CA1	245	4460043
ICC4 CA2	210	8928794
ICC4 CA3	195	550684
ICC4 CA4	243	214389

Note. AAs = amino acids, bp = base pairs

DISCUSSION

The results of the growth curve assay were inconclusive in determining an accurate exponential phase for *Streptomyces* sp. ICC4. After 97 hours of incubation, a significant decrease in supernatant absorbance after adsorption was not yet observed, however, absorbance after desorption increased more sharply. The ICC4 cells formed large spherical pellets that did not appear to take up the dye very well. Similar pellet-like structures have been observed in *Streptomyces brollosae* NEAE-115 during the production of L-asparaginase (El-Naggar et al. 2019). Due to limited pellet size and negligible change in absorbance, we can confidently assume that after 48 hours, ICC4 has not exited the lag phase, but the results do not support an accurate exponential phase. In comparison to ICC4, the ICC1 supernatant showed a significant decrease in absorbance after adsorption after 45 hours. A further significant decline was observed after 52 hours. These results would indicate that the exponential phase was reached by 52 hours of incubation. However, while the absorbance after desorption readings were highest at 45 hours and

decreased by 52 hours, after 68 hours, absorbance readings after desorption further decreased, but the expected parallel increase in absorbance after adsorption was also observed. This result is unexpected even if the stationary phase was achieved as a property of methylene blue is to stain both dead and living cells (Borzani and Vairo 1959). My results would indicate that ICC1 has an earlier exponential phase in comparison to ICC4 but given the unexpected data and limited trials. Further optimization and other forms of growth measurement should be employed to validate the predicted growth patterns.

The results of the urea agar test would indicate that both ICC1 and ICC4 have active urease enzymes as a pink color was observed after one week of incubation which is consistent with the formation of ammonium causing an increase in the pH of the agar. High percentage identifications to ICC1 and ICC4 were found in all copies of the *ureC* genes amplified with the PCR primers in this study. This result confirms that isolates of *Streptomyces* sp. ICC1 and ICC4 grown in the lab have not shown significant change in the *ureC* genes in comparison to when previously sequenced (Gosse et al. 2019). This also supports the design of PCR primers to successfully amplify their expected products. NCBI Sequencing data can reliably support further studies on these bacteria. The *ureC*4 primer set still showed some non-specific products and primer-dimer when fully optimized. *ureC*4 and *ureC*5 also encode products significantly above 150 bp which can decrease amplification efficiency skewing results when performing qRT-PCR (Ruiz-Villalba 2017). Therefore, new primer sets for *ureC*4 and *ureC5* should be made if qRT-PCR is to be performed. While five copies of the *ureC* gene were detected, no full *urease* operon as expected based on previous MICP bacteria was found in either ICC1 or ICC4. Notably, the *ureE* gene was missing from all operons which is typically the nickel carrier for the active urease enzyme (Wu et al. 2021). It is possible that ICC1 and ICC4 may contain a separate form of operon from other species where

another co-factor has taken on the role of nickel. However, in previous studies from the Iron Curtain Cave, nickel was present in both the popcorn and soda straw type speleothems which indicates that a lack of a nickel would not be a factor for a potential change in co-factor (Koning et al. 2022). The complete lack of accessory genes associated with the *ureC3* and *ureC4* genes is concerning that these urease enzymes may exhibit low levels of activity as previous research in other bacteria has indicated their importance to urease expression (Liu et al. 2017). Four carbonic anhydrase genes were detected in each species through bioinformatic analysis supporting the theory that they could act synergistically with the *urease* operon. All carbonic anhydrase copies showed high similarity with those from other *Streptomyces* species. However, carbonic anhydrase is not well studied in *Streptomyces* outside of *S. kunmingensis* which showed poor similarity with the carbonic anhydrase genes in ICC1 and ICC4 (Sangeetha et al. 2022). Further analysis of the expression patterns of these genes would be necessary to confirm these predictions.

The formation of calcium carbonate crystals is critical to understanding if ICC1 and ICC4 are good candidates for the formation of biocement. After a 4-week incubation at 8°C, crystal-like structures were observed in the agar of both ICC1 and ICC4. However, the structures visible did not show a correlation with those observed from previous MICP studies of the Iron Curtain Cave and other previous work with *Sporosarcina pasteurii* (Ghosh et al. 2019 and Koning et al. 2022). While ICC4 showed a light scattering pattern consistent with the presence of calcium carbonate crystals in liquid media that was not visible with ICC1, both cultures showed almost identical structures on agar plates. Given that other crystals appear similar in appearance at low resolution and also reflect light, these results are not definitive. To confirm if the crystals observed were of calcium carbonate, a higher level of microscopy could be employed such as a scanning electron microscope (Ghosh et al. 2019). A temperature of 8°C was used as an incubation due to previous

success with this temperature for other bacteria from the Iron Curtain Cave to induce calcium carbonate crystal formation. However, given that 15°C appeared optimal for growth for both ICC1 and ICC4, this incubation temperature may have been a better choice for calcium carbonate formation. Higher temperatures have been used for this incubation in other species of bacteria that do not require colder temperatures for growth (Hoffmann et al. 2021). Given the potentially suboptimal temperature, it is possible it may take longer for larger calcium carbonate aggregates to form. The lack of successful growth with the B4 media substituting yeast extract for urea can likely be attributed to the loss of essential elements for bacterial growth such as sulfur and phosphate present in the yeast extract, but not in the urea.

In conclusion, the urea agar test indicated that *Streptomyces* sp. ICC1 and ICC4 have urease activity and sequencing results after PCR of the five expected *ureC* gene copies confirmed the presence of these genes at high similarity with previous sequencing results. PCR primers developed were successful and have the potential to be used in future studies to analyze *ureC* expression. However, bioinformatic analysis for the presence of the urease accessory genes indicated that the complete operon typical of other bacterial species was not present for any *ureC* gene copy meaning a different form of operon may be used for ICC1 and ICC4 or activity levels may be lower than previously anticipated. Results from the B4 media test were also inconclusive as definitive calcium carbonate crystals were not observed after a four-week incubation. Higher resolution tests are needed to determine the identity of the structures observed. Experimenting with a longer incubation time or modifying the incubation temperature are possible alternatives to induce calcium carbonate crystal formation.

In future studies, different media types and conditions such as pH and incubation temperature should be tested with ICC1 and ICC4 to determine what is optimal for calcium

carbonate crystal formation given the lack of conclusive results. To analyze which *ureC* copies are being expressed and at what levels, RNA extractions should be performed on ICC1 and ICC4 and qRT-PCR should be performed on all the *ureC* gene copies. Successful PCR primers from this study could be used to perform this objective. To analyze the efficiency of the urease enzymes from ICC1 and ICC4 themselves, each copy of the *urease* operon could be cloned and overexpressed. Enzymes could then be purified and studied using a variety of biochemical assays such as Michaelis-Menten kinetics. Structural biology techniques such as X-ray crystallography could also be performed on purified enzymes so the structures of ICC1 and ICC4 enzymes could be compared to other known urease enzymes to better understand their activity.

LITERATURE CITED

Bender, R.A. 2010, October. A NAC for regulating metabolism: The nitrogen assimilation control protein (NAC) from Klebsiella pneumoniae. doi:10.1128/JB.00266-10.

Borzani, W., and Vairo, M.L.R. 1959. ADSORPTION OF METHYLENE BLUE AS A MEANS OF DETERMINING CELL CONCENTRATION OF DEAD BACTERIA IN SUSPENSIONS. https://doi.org/10.3109/10520296009114720

Cheeptham, N. 2013. Advances and Challenges in Studying Cave Microbial Diversity. pp. 1–34. doi:10.1007/978-1-4614-5206-5_1.

Chuo, S.C., Mohamed, S.F., Setapar, S.H.M., Ahmad, A., Jawaid, M., Wani, W.A., Yaqoob, A.A., and Ibrahim, M.N.M. 2020, November 1. Insights into the current trends in the utilization of bacteria for microbially induced calcium carbonate precipitation. MDPI AG. doi:10.3390/ma13214993.

Clarà Saracho, A., and Marek, E.J. 2024. Uncovering the Dynamics of Urease and Carbonic Anhydrase Genes in Ureolysis, Carbon Dioxide Hydration, and Calcium Carbonate Precipitation. Environ Sci Technol **58**(2): 1199–1210. American Chemical Society. doi:10.1021/acs.est.3c06617.

Demény, A., Németh, P., Czuppon, G., Leél-Ossy, S., Szabó, M., Judik, K., Németh, T., and Stieber, J. 2016. Formation of amorphous calcium carbonate in caves and its implications for speleothem research. Sci Rep **6**. Nature Publishing Group. doi:10.1038/srep39602.

Dhami, N.K., Reddy, M.S., and Mukherjee, A. 2014. Synergistic role of bacterial urease and carbonic anhydrase in carbonate mineralization. Appl Biochem Biotechnol **172**(5): 2552–2561. Humana Press Inc. doi:10.1007/s12010-013-0694-0.

El-Naggar, N.E.A., Moawad, H., El-Shweihy, N.M., El-Ewasy, S.M., Elsehemy, I.A., and Abdelwahed, N.A.M. 2019. Process development for scale-up production of a therapeutic L-asparaginase by Streptomyces brollosae NEAE-115 from shake flasks to bioreactor. Sci Rep **9**(1). Nature Publishing Group. doi:10.1038/s41598-019-49709-6.

Enyedi, N.T., Makk, J., Kótai, L., Berényi, B., Klébert, S., Sebestyén, Z., Molnár, Z., Borsodi, A.K., Leél-Őssy, S., Demény, A., and Németh, P. 2020. Cave bacteria-induced amorphous calcium carbonate formation. Sci Rep **10**(1). Nature Research. doi:10.1038/s41598-020-65667-w.

Fischer, M., and Sawers, R.G. 2013. A universally applicable and rapid method for measuring the growth of Streptomyces and other filamentous microorganisms by methylene blue adsorption-desorption. Appl Environ Microbiol **79**(14): 4499–4502. doi:10.1128/AEM.00778-13.

Ghosh, S., Paine, E., Wall, R., Kam, G., Lauriente, T., Sa-Ngarmangkang, P.C., Horne, D., and Cheeptham, N. 2017. In situ cultured bacterial diversity from Iron Curtain Cave, Chilliwack, British Columbia, Canada. Diversity (Basel) **9**(3). MDPI AG. doi:10.3390/d9030036.

Ghosh, T., Bhaduri, S., Montemagno, C., and Kumar, A. 2019. Sporosarcina pasteurii can form nanoscale calcium carbonate crystals on cell surface. PLoS One **14**(1). Public Library of Science. doi:10.1371/journal.pone.0210339.

Gosse, J.T., Ghosh, S., Sproule, A., Overy, D., Cheeptham, N., and Boddy, C.N. 2019. Whole genome sequencing and metabolomic study of cave Streptomyces isolates ICC1 and ICC4. Front Microbiol 10(MAY). Frontiers Media S.A. doi:10.3389/fmicb.2019.01020.

Hoffmann, T.D., Paine, K., and Gebhard, S. 2021. Genetic optimisation of bacteria-induced calcite precipitation in Bacillus subtilis. Microb Cell Fact **20**(1). BioMed Central Ltd. doi:10.1186/s12934-021-01704-1.

Iranzo, J., Wolf, Y.I., Koonin, E. V., and Sela, I. 2019. Gene gain and loss push prokaryotes beyond the homologous recombination barrier and accelerate genome sequence divergence. Nat Commun **10**(1). Nature Research. doi:10.1038/s41467-019-13429-2.

Kajaste, R., and Hurme, M. 2016. Cement industry greenhouse gas emissions - Management options and abatement cost. J Clean Prod **112**: 4041–4052. Elsevier Ltd. doi:10.1016/j.jclepro.2015.07.055.

Koning, K., McFarlane, R., Gosse, J.T., Lawrence, S., Carr, L., Horne, D., Van Wagoner, N., Boddy, C.N., and Cheeptham, N. 2022. Biomineralization in Cave Bacteria—Popcorn and Soda Straw Crystal Formations, Morphologies, and Potential Metabolic Pathways. Front Microbiol **13**. Frontiers Media S.A. doi:10.3389/fmicb.2022.933388.

Liu, X., Zhang, Q., Zhou, N., and Tian, Y. 2017, March 1. Expression of an Acid Urease with Urethanase Activity in E. coli and Analysis of Urease Gene. Humana Press Inc. doi:10.1007/s12033-017-9994-x.

Maciejewska, M., Adam, D., Naômé, A., Martinet, L., Tenconi, E., Calusinska, M., Delfosse, P., Hanikenne, M., Baurain, D., Compère, P., Carno l, M., Barton, H.A., and Rigali, S. 2017. Assessment of the potential role of Streptomyces in cave moonmilk formation. Front Microbiol **8**(JUN). Frontiers Media S.A. doi:10.3389/fmicb.2017.01181.

Marvasi, M., Gallagher, K.L., Martinez, L.C., Molina Pagan, W.C., Rodríguez Santiago, R.E., Castilloveitía Vega, G., and Visscher, P.T. 2012. Importance of B4 Medium in Determining Organomineralization Potential of Bacterial Environmental Isolates. Geomicrobiol J **29**(10): 916–924. doi:10.1080/01490451.2011.636145.

Mulrooney, S.B., Ward, S.K., and Hausinger, R.P. 2005. Purification and properties of the Klebsiella aerogenes UreE metal-binding domain, a functional metallochaperone of urease. J Bacteriol **187**(10): 3581–3585. doi:10.1128/JB.187.10.3581-3585.2005.

Omoregie, A.I., Ong, D.E.L., and Nissom, P.M. 2019. Assessing ureolytic bacteria with calcifying abilities isolated from limestone caves for biocalcification. Lett Appl Microbiol **68**(2): 173–181. John Wiley and Sons Inc. doi:10.1111/lam.13103.

Pei, D., and Hu, B. 2023. A novel urease gene structure of Sporosarcina pasteurii with double operons. doi:10.21203/rs.3.rs-3204608/v1.

Prajapati, N.K., Agnihotri, A.K., and Basak, N. 2023. Microbial induced calcite precipitation (MICP) a sustainable technique for stabilization of soil: A review. *In* Materials Today: Proceedings. Elsevier Ltd. pp. 357–361. doi:10.1016/j.matpr.2023.07.303.

Ruiz-Villalba, A., van Pelt-Verkuil, E., Gunst, Q.D., Ruijter, J.M., and van den Hoff, M.J. 2017. Amplification of nonspecific products in quantitative polymerase chain reactions (qPCR). Biomol Detect Quantif **14**: 7–18. Elsevier GmbH. doi:10.1016/j.bdq.2017.10.001.

Sangeetha, M., Sivarajan, A., Radhakrishnan, M., Siddharthan, N., and Balagurunathan, R. 2022. Biosequestration of carbon dioxide using carbonic anhydrase from novel Streptomyces kunmingensis. Arch Microbiol **204**(5). Springer Science and Business Media Deutschland GmbH. doi:10.1007/s00203-022-02887-w.

Šmarda, P., Bureš, P., Horová, L., Leitch, I.J., Mucina, L., Pacini, E., Tichý, L., Grulich, V., and Rotreklová, O. 2014. Ecological and evolutionary significance of genomic GC content diversity in monocots. Proc Natl Acad Sci U S A **111**(39): E4096–E4102. National Academy of Sciences. doi:10.1073/pnas.1321152111.

Wray, L. V, Ferson, A.E., and Fisher, S.H. 1997. Expression of the Bacillus subtilis ureABC Operon Is Controlled by Multiple Regulatory Factors Including CodY, GlnR, TnrA, and Spo0H. *In* JOURNAL OF BACTERIOLOGY. doi: 10.1128/jb.179.17.5494-5501.1997

Wu, Y., Li, H., and Li, Y. 2021, November 1. Biomineralization induced by cells of Sporosarcina pasteurii: Mechanisms, applications and challenges. MDPI. doi:10.3390/microorganisms9112396.
Zheng, T., and Qian, C. 2020. Influencing factors and formation mechanism of CaCO3 precipitation induced by microbial carbonic anhydrase. Process Biochemistry **91**: 271–281. Elsevier Ltd. doi:10.1016/j.procbio.2019.12.018.

APPENDIX

Media Composition per 1000 ml of dH_20

B4 Media:

Yeast Extract: 4 g

Dextrose: 5 g

Calcium Acetate Hydrate: 2.42 g

Agar: 14 g

pH: 8.2

Hickey-Tresner Media:

Yeast Extract: 1 g

Beef Extract: 1 g

N-Z Amine A: 2 g

Dextrin: 10 g

pH: 7.3

*Tap water used in place of dH_20

Urea Agar:

Pancreatic Digest of Gelatin: 1 g

Dextrose: 1 g

Sodium Chloride: 5 g

Potassium Phosphate: 2 g

Urea: 20 g

Phenol Red: 12 mg

Raw Sequence Data of *ureC* Genes

>Icc1urec1F-ud 252 32 217 0.05

TACAACGATGACCGCTCACACTGCAGGACCAGCGCGGCAAGCTCGACGGCGACACCGAGCGCAACGACAA

CTTCCGCGTCCTGCGCTACCTCGCCAAGATCACGATCAACCCGGCGATCGCCACCGGCACCGCCGAGCAC

ATCGGCTCCCTGGAGACCGGCAAGCTCGCCGACATCGTGCTCTGGCCGATCGGCTCCTTCGCCGCCAAGC

CGAAGATGGTCATCAAGGGCGGACTGATCACCTGGGCGCTCA

>Icc4urec1F-ud 247 247 0 0.05

AGTCAGTGACAGTCAACACGGATGGGAATAGCGCGGGCAGCTCGACGGCGACCCGAGCGCAACGACAACT

TTCTCGTCCTGCGCTACCTCGCCAGATCACGATCAACCCACGATCGCCACCGGCACCGCCGAGCACATCG

GCTTCCGGAGACCGGCAAGCTCGCCGACATCGTGCTCTGGCCGATCGGCTCCTTCGCCGCCAAGCCGAAG

ACGGTCATCAAGGGCGGACCGTCCCTTGGGCGCTTAA

>Icc1urec2F-ud 238 238 0 0.05

TTCCTGCTATACGTATTCTGCTGCGACAGCGCCTTCATCCTGGGTTCAGGAAGGTGCGGAGGTTTAGAAA

AGAGCCTTGGGATAAGATGAGAAAATCCGGTCATCGCCCCGTTTGAGCCGAGGCATCGGATCCCTGGAGA

CCGGCTGGCTCCCCGACATCGTGCTCTGGCCGATCGGCTCCTCTTCCGCCGAGCCGAAGATGGTCATCAA

GGGGGGACTGATACTTGGTTTTTTTCAA

>Icc4urec2F-ud 135 29 15 0.05

GAGGCCCAGATCCGGCACCGCACACCCAGTCCGGGGTCCACCCCAAGCTTGGGTCATCCGGGCCCGGCAC

CGAGGTGATCTCCGGCGAGCACCTGATCGCCACGGCCGGCGGCATCGACACCCACATCCACTTCA

>Icc1urec3F-ud 99 99 0 0.05

TTGGTCCCGGCTCGGCGAACTGGATGGCCACCCGGCGCCGCCGCGTCGCGGTCCGCGGGACCCGGGGCAT

CGGGCCGAAGGACATGCTCCGCAACTAAA

>Icc4urec3F-ud 97 97 0 0.05

TAGGCCCGATTGAGCAATTGTGGCCACCCGGGCGCCCGCCGCGTCGCGGTCCGCGGGACCGGGGCATCGG

GCCGAAGGACATGCTTCCGCAACTAAA

>Icc1urec4F-ud 903 61 810 0.05

TAGCCCCAAGGACTGCCGGCTTGGCGACTCCGGGCTGACCGTCTGCGTCGAGTCCGATTCCCAGAAGCCC

GGCGACGAGTTCCTCGCCGGCTTCGGCAAGACCGCCCGCGACGGACTGCACCTGAAGGCCGCCGCCGTCC

GCGACACCTGCGACGTGGTGATCAGCAACGTGCTGGTCATCGACGCCGTCCTCGGCATCCGCAAGGTCTC

CATCGGCATCCGCGAGGGCCGCATCCACGCGATCGGCCGCGCCGGAAACCCGGACACCCTGGACGGGGTG

GACGTCGTGGTCGGCACGGGCACGAGCATCGTCTCCGGCGAGGGCATGATCGCCACCGCCGGAGCCGTCG

ACACCCACGTGCACCTGCTGTCCCCGCGCATCATGGAGGCCTCGCTCGCGGCGGGCGTCACCACGATCAT

CGGCCAGGAGTTCGGCCCGGTCTGGGGCGTCGGAGTCAACTCCCCGTGGGCGCTCAAGCACGCCTTCAAC

GCCTTCGACGCCTGGCCCGTCAACATCGGCTTCCTCGCCCGCGGCTCCTCCTCCGACCCGGCCCCGCTGA

TCGAGGCGCTGGCCGAGGGCGGTGCGAGCGGCTTCAAGGTCCACGAGGACATGGGCGCCCACACCCGCGC

CCTGGACACTGCCCTGCGGGTCGCGGAGGAGTACGACGTACAGGTCGCCCTGCACAGCGACGGCCTCAAC

GAGTGCCTGTCCGTCGAGGACACCCTGCGGGTGCTGGACGGCCGCACCATCCACGCCTTCCACATCGAGG

GCTGCGGCGGCGGACACGTCCCCAACGTCCTGAAGATGGCCGGCGTGCCGAACGTGATCGGCTCCTCCAC

CAACCCCACCCTGCCCTTCGGCCGGGACGCGGTCGCCGAGCGCGGTGGGGTAAAAAAAAAAAC

>Icc4urec4F-ud 904 32 44 0.05

GAGGTTCTTGACGTCGATCCGGCTCGGCGACTCCGGGCTGACCGTCTGCGTCGAAGTCCGATTCCCAGAA

GCCCGGCGACGTAGTTCCTCGCCGGCTTCGGCAAGACCGCCCGCGACGGACTGCACCTGAAGGCCGCCGC

CGTCCGCGACACCTGCGACGTGGTGATCAGCAACGTGCTGGTCATCGACGCCGTCCTCGGCATCCGCAAG

GTCTCCATCGGCATCCGCGAGGGCCGCATCCACGCGATCGGCCGCGCCGGAAACCCGGACACCCTGGACG

GGGTGGACGTCGTGGTCGGCACGGGCACGAGCATCGTCTCCGGCGAGGGCATGATCGCCACCGCCGGAGC

CGTCGACACCCACGTGCACCTGCTGTCCCCGCGCATCATGGAGGCCTCGCTCGCGGCGGGCGTCACCACG

ATCATCGGCCAGGAGTTCGGCCCGGTCTGGGGCGTCGGAGTCAACTCCCCGTGGGCGCTCAAGCACGCCT

TCAACGCCTTCGACGCCTGGCCCGTCAACATCGGCTTCCTCGCCCGCGGCTCCTCCTCCGACCCGGCCCC

GCTGATCGAGGCGCTGGCCGAGGGCGGTGCGAGCGGCTTCAAGGTCCACGAGGACATGGGCGCCCACACC

CGCGCCCTGGACACTGCCCTGCGGGTCGCGGAGGAGTACGACGTACAGGTCGCCCTGCACAGCGACGGCC

TCAACGAGTGCCTGTCCGTCGAGGACACCCTGCGGGTGCTGGACGGCCGCACCATCCACGCCTTCCACAT

CGAGGGCTGCGGCGGCGGACACGTCCCCAACGTCCTGAAGATGGCCGGCGTGCCGAACGTGATCGGCTCC

TCCACCAACCCCACCCTGCCCTTCGGCCGGGACGCGGTCGCCGAGAAGGGGGGATTAAAAAAAA

>Icc1urec5F-ud 705 93 22 0.05

TAGCCCTTCACAAGAAGTTTCGAATTATCACCGTGGGTAACCGAAGCCGAACGTCCTGCCAGCTTCACCA

ACCCGACCCGGCCGCACACCGTCAACACCGTCGAGGAGCACCTCGACATGCTGATGGTCTGCCACCACCT

CAACCCCTCCGTGCCCGAGGACCTGGCCTTCGCCGAGTCCCGGATCCGGCCCTCCACCATCGCCGCCGAG

GACGTCCTGCACGACCTCGGGGCCATCTCGATCATCTCTTCCGACGCCCAGGCCATGGGCCGCGTCGGCG

AGGTGGTCCTGCGCACCTGGCAGACCGCCCACGTGATGAAGAAGCGGCGCGGGTTCCTGCCCGGGGACGG

CCCCGCCGACAACCACCGGGTCCGCCGCTACGTCGCCAAGTACACGATCAACCCCGCCGTGGCCCAGGGC

CTGGCCCGCGAGATCGGCTCGGTGGAGACCGGCAAACTCGCCGACCTGGTGCTGTGGAACCCCGCCTTCT

TCGGCGTGAAGCCCGAACTCGTCATCAAGGGCGGCCAGATCGCCTACGCGCAGATGGGCGATGCCAACGC

CTCGATCCCCACCCCGCAGCCGGTGCTGCCCCGCCCGATGTTCGGGGCGTACGGGCGCGCCCCCGGCCTG

AACTCGGTCAACTTCACCGCCCAGGCGGCCCTCGACGACGGACTGCCCGAACGGCTCGATGGAAAGGGGA

AATTA

>Icc4urec5F-ud 702 88 20 0.05

TCGCACTCCCCAGTGTTAAGAAACATCACCGTGGGTATCGAGCCGAACGTCCTGCCAGCTCCCCAACCCG

ACCCGGCCGCACACCGTCAACACCGTCGAGGAGCACCTCGACATGCTGATGGTCTGCCACCACCTCAACC

CCTCCGTGCCCGAGGACCTGGCCTTCGCCGAGTCCCGGATCCGGCCCTCCACCATCGCCGCCGAGGACGT

CCTGCACGACCTCGGGGCCATCTCGATCATCTCTTCCGACGCCCAGGCCATGGGCCGCGTCGGCGAGGTG

GTCCTGCGCACCTGGCAGACCGCCCACGTGATGAAGAAGCGGCGCGGGTTCCTGCCCGGGGACGGCCCCG

CCGACAACCACCGGGTCCGCCGCTACGTCGCCAAGTACACGATCAACCCCGCCGTGGCCCAGGGCCTGGC

CCGCGAGATCGGCTCGGTGGAGACCGGCAAACTCGCCGACCTGGTGCTGTGGAACCCCGCCTTCTTCGGC

GTGAAGCCCGAACTCGTCATCAAGGGCGGCCAGATCGCCTACGCGCAGATGGGCGATGCCAACGCCTCGA

TCCCCACCCCGCAGCCGGTGCTGCCCCGCCCGATGTTCGGGGCGTACGGGCGCGCCCCCGGCCTGAACTC

GGTCAACTTCACCGCCCAGGCGGCCCTCGACGACGGACTGCCCGAACGGCTCGCGGCAAGCGAAAATAAT

AA

>Icc1urec1R-ud 262 41 215 0.05

GGGCTGGGCTGGCTTAGCGATTTTGGCTGGCGGCGAGGAGCCGATCGGCCAGAGCACGATGTCGGCGAGC

TTGCCGGTCTCCAGGGAGCCGATGTGCTCGGCGGTGCCGGTGGCGATCGCCGGGTTGATCGTGATCTTGG

CGAGGTAGCGCAGGACGCGGAAGTTGTCGTTGCGCTCGGTGTCGCCGTCGAGCTTGCCGCGCTGGTCCTT

GCAGTGGTGAGCGGTCTGGAAGGCGCGGGTCCAGGATTCTCCGATGCGGCCA

>Icc4urec1R-ud 257 91 149 0.05

TCGCAAGTATCCTGGGCTGGGCGGCGAAGGAGCCGAATGGCCAGAAGCACGATGTCGGCGAAGCTTGCCG

GTCTCCCAGGGAGCCGTATGTGCTCGGCGGTGCCGGTGGCGATCGCCGGGTTGATCGTGATCTTGGCGAG

GTAGCGCAGGACGCGGAAGTTGTCGTTGCGCTCGGTGTCGCCGTCGAGCTTGCCGCGCTGGTCCTTGCAG

TGGTGAGCGGTCTGGAAGGCGCGGGTCCAGGATTCCCGATGCGGCCA

>Icc1urec2R-ud 129 31 93 0.05

TTCGTGCGTACGATAGTGCTCGCCGGAGATGACCTCGGTGCCGGGCCCGATGACCAGCTTGGGGTGGACC

CCGGACTGGGTGTGCGGGTTGCCGGACTTGCCGAGACCCGCGATGAACCCGTCCTTGAT

>Icc4urec2R-ud 139 50 81 0.05

TAGCCCATCCTTGGGACCGGTGGCTCGCCGGAGATGACCTCGGTGCCGGGCCCGATGACCAGCTTGGGGT

GGACCCCGGACTGCGGTGTGCGGGTTGCCGGCACTTGCCGAGGCCCGCGATGAACCCGTCCTTGATCGA

>Icc1urec3R-ud 93 93 0 0.05

TAGGCCGGGAAGGCTGGGGGACCGCGACGCCGGCGGCGCCGGGGAGCATAGGTCGCCGTTCTGCACGGCG

GCCTCGCGACAACGCCACCGATA

>Icc4urec3R-ud 104 62 40 0.05

GAGGCCAGATGCAGGCTCGGCGGGACCGCGAAGCGGCGGCGCCGGGGTGGGGCATCCAGGGTCGCCGTTC

TGCACGGCGGCCTCCGCGACGAACGCCACCGATA

>Icc1urec4R-ud 911 74 803 0.05

TGGTCACACGCGGTCGGGAAGGTGGGGGTTTGGTGGGAGGAAGCCGAATCACGTTCGGCACGCCGGCCAT

CTTCAGGACGTTGGGGACGTGTCCGCCGCCGCAGCCCTCGATGTGGAAGGCGTGGATGGTGCGGCCGTCC

AGCACCCGCAGGGTGTCCTCGACGGACAGGCACTCGTTGAGGCCGTCGCTGTGCAGGGCGACCTGTACGT

CGTACTCCTCCGCGACCCGCAGGGCAGTGTCCAGGGCGCGGGTGTGGGCGCCCATGTCCTCGTGGACCTT

GAAGCCGCTCGCACCGCCCTCGGCCAGCGCCTCGATCAGCGGGGCCGGGTCGGAGGAGGAGCCGCGGGCG

AGGAAGCCGATGTTGACGGGCCAGGCGTCGAAGGCGTTGAAGGCGTGCTTGAGCGCCCACGGGGAGTTGA

CTCCGACGCCCCAGACCGGGCCGAACTCCTGGCCGATGATCGTGGTGACGCCCGCCGCGAGCGAGGCCTC

CATGATGCGCGGGGACAGCAGGTGCACGTGGGTGTCGACGGCTCCGGCGGTGGCGATCATGCCCTCGCCG

GAGACGATGCTCGTGCCCGTGCCGACCACGACGTCCACCCCGTCCAGGGTGTCCGGGTTTCCGGCGCGGC

CGATCGCGTGGATGCGGCCCTCGCGGATGCCGATGGAGACCTTGCGGATGCCGAGGACGGCGTCGATGAC

CAGCACGTTGCTGATCACCACGTCGCAGGTGTCGCGGACGGCGGCGGCCTTCAGGTGCAGTCCGTCGCGG

GCGGTCTTGCCGAAGCCGGCGAGGAACTCGTCGCCGGGCTTCTGGGAATCGGACTCGACGCAGACGGTCA

GCCCGGAGTCGCCGAGCCGGACCCGGTCCTCGGCGCGGGGCGTGAGGGGAGGGGGCGCCTACATACATAC

A

>Icc4urec4R-ud 909 29 848 0.05

TTAGTCCGATATAGATAATCGGCAGGTGGGGTTGGTGGAGGAAGCCGATCACGTTCGGCACGCCGGCCAT

CTTCAGGACGTTGGGGACGTGTCCGCCGCCGCAGCCCTCGATGTGGAAGGCGTGGATGGTGCGGCCGTCC

AGCACCCGCAGGGTGTCCTCGACGGACAGGCACTCGTTGAGGCCGTCGCTGTGCAGGGCGACCTGTACGT

CGTACTCCTCCGCGACCCGCAGGGCAGTGTCCAGGGCGCGGGTGTGGGCGCCCATGTCCTCGTGGACCTT

GAAGCCGCTCGCACCGCCCTCGGCCAGCGCCTCGATCAGCGGGGCCGGGTCGGAGGAGGAGCCGCGGGCG

AGGAAGCCGATGTTGACGGGCCAGGCGTCGAAGGCGTTGAAGGCGTGCTTGAGCGCCCACGGGGAGTTGA

CTCCGACGCCCCAGACCGGGCCGAACTCCTGGCCGATGATCGTGGTGACGCCCGCCGCGAGCGAGGCCTC

CATGATGCGCGGGGACAGCAGGTGCACGTGGGTGTCGACGGCTCCGGCGGTGGCGATCATGCCCTCGCCG

GAGACGATGCTCGTGCCCGTGCCGACCACGACGTCCACCCCGTCCAGGGTGTCCGGGTTTCCGGCGCGGC

CGATCGCGTGGATGCGGCCCTCGCGGATGCCGATGGAGACCTTGCGGATGCCGAGGACGGCGTCGATGAC

CAGCACGTTGCTGATCACCACGTCGCAGGTGTCGCGGACGGCGGCGGCCTTCAGGTGCAGTCCGTCGCGG

GCGGTCTTGCCGAAGCCGGCGAGGAACTCGTCGCCGGGCTTCTGGGAATCGGACTCGACGCAGACGGTCA

GCCCGGAGTCGCCGAGCCGGACCCGGTCCCCGGCGCGGGGCGTGAGGGGGGGGGGTTTTACTACACATA

>Icc1urec5R-ud 678 226 429 0.05

TGCTCCAGGGGCCCCGCGCGGCGGGAATCGACGATTAGGCGGAGGCGCTGTACGTTGAAATCGGGCGGGG

AGCCCGGCGCGGGGTGGGGTCGAGGCGTTGGCATCGCTCTCTGCGCGTAGGCGATCCGGCCGCTCTTGAT

GACGAGTTCGGGCTTCACGCCGAAGAAGGCGGGGTTCCAAGCACCAGGTCGGCGAGTTTGCCGGTCTCCA

CCGAGCCGATCTCGCGGGCCAGGCCCTGGGCCACGGCGGGGTTGATCGTGTACTTGGCGACGTAGCGGCG

GACCCGGTGGTTGTCGGCGGGGCCGTCCCCGGGCAGGAACCCGCGCCGCTTCTTCATCACGTGGGCGGTC

TGCCAGGTGCGCAGGACCACCTCGCCGACGCGGCCCATGGCCTGGGCGTCGGAAGAGATGATCGAGATGG

CCCCGAGGTCGTGCAGGACGTCCTCGGCGGCGATGGTGGAGGGCCGGATCCGGGACTCGGCGAAGGCCAG

GTCCTCGGGCACGGAGGGGTTGAGGTGGTGGCAGACCATCAGCATGTCGAGGTGCTCCTCGACGGTGTTG

ACGGTGTGCGGCCGGGTCGGGTTGGTGGAGCTGGGCAGGACGTTCGGCTCGGAGACCACGGTGATGATGT

CGGGGGCGTGCCCGCCGCCCGCGCCCTCGTGGCAAGAGGGGGGGGGGA

>Icc4urec5R-ud 715 74 616 0.05

GAGTCTACTCGTAGTCGTGTTGGCGCCTGGGGCGGTGAAGTTGAACCGAAGTTCAGGCCGGGGGCGCGCC

CGTACGCCCCGAACATTCGGGCGGGGCAGCACCGGCTGCGGGGTTGGGGATCGAGGCGTTGGCATCGCCC

ATCTGCGCGTAGGCGATCTGGCCGCCCTTGATGACGAGTTCGGGCTTCACGCCGAAGAAGGCGGGGTTCC

ACAGCACCAGGTCGGCGAGTTTGCCGGTCTCCACCGAGCCGATCTCGCGGGCCAGGCCCTGGGCCACGGC

GGGGTTGATCGTGTACTTGGCGACGTAGCGGCGGACCCGGTGGTTGTCGGCGGGGCCGTCCCCGGGCAGG

AACCCGCGCCGCTTCTTCATCACGTGGGCGGTCTGCCAGGTGCGCAGGACCACCTCGCCGACGCGGCCCA

TGGCCTGGGCGTCGGAAGAGATGATCGAGATGGCCCCGAGGTCGTGCAGGACGTCCTCGGCGGCGATGGT

GGAGGGCCGGATCCGGGACTCGGCGAAGGCCAGGTCCTCGGGCACGGAGGGGTTGAGGTGGTGGCAGACC

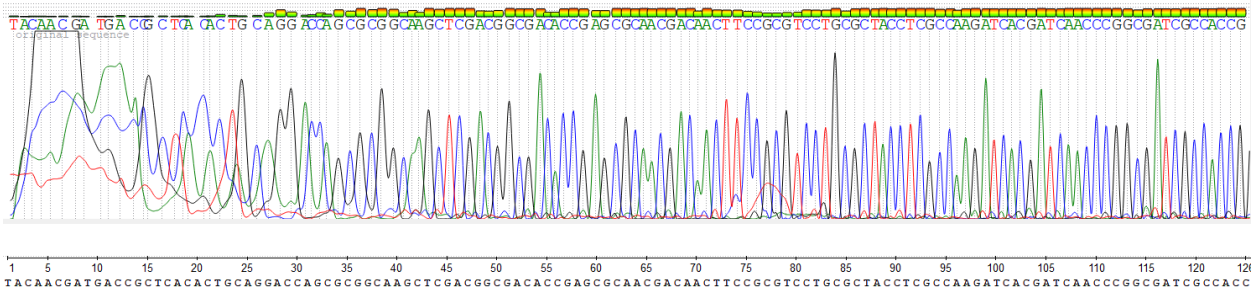
ATCAGCATGTCGAGGTGCTCCTCGACGGTGTTGACGGTGTGCGGCCGGGTCGGGTTGGTGGAGCTGGGCA

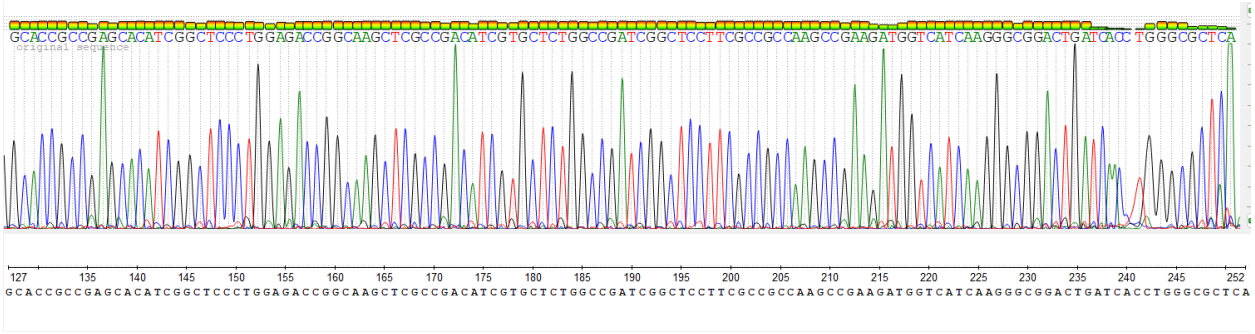
GGACGTTCGGCTCGGAGACCACGGTGATGATGTCGGGGGCGTGCCCGCCGCCCGCGCCCTCGGTGGGCAA

AGGGGGAGTGGGGGA

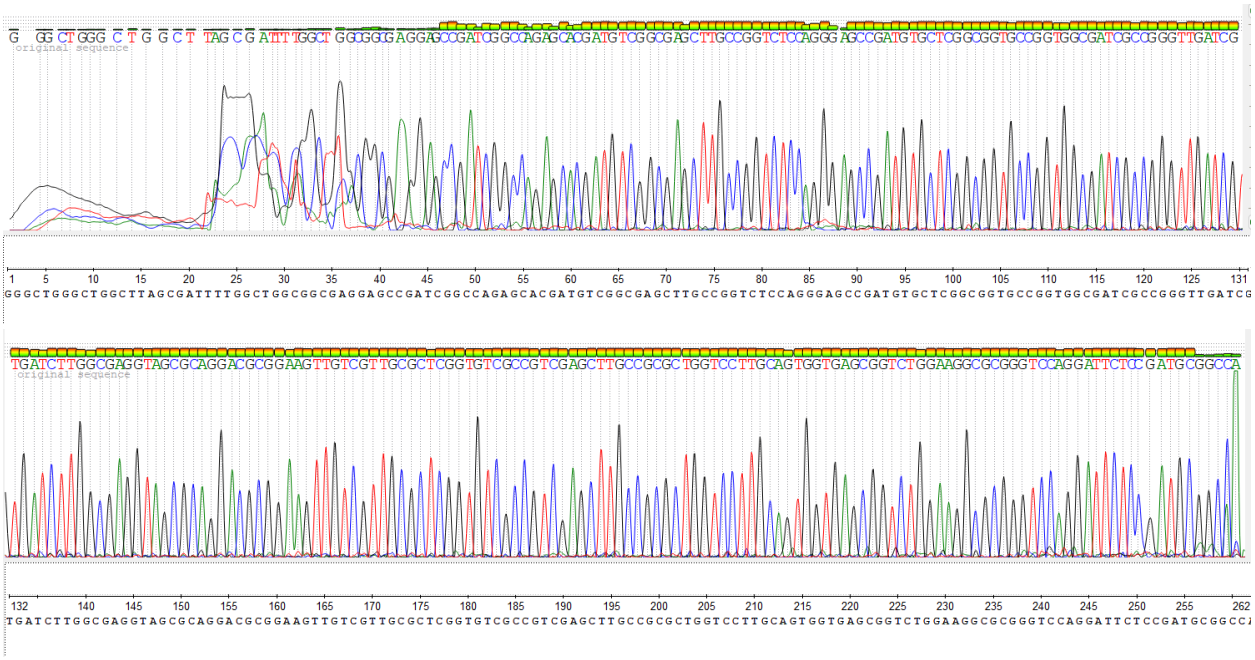
Chromatograms of *ureC* Genes Sanger Sequencing Data

ICC1*ureC*1F

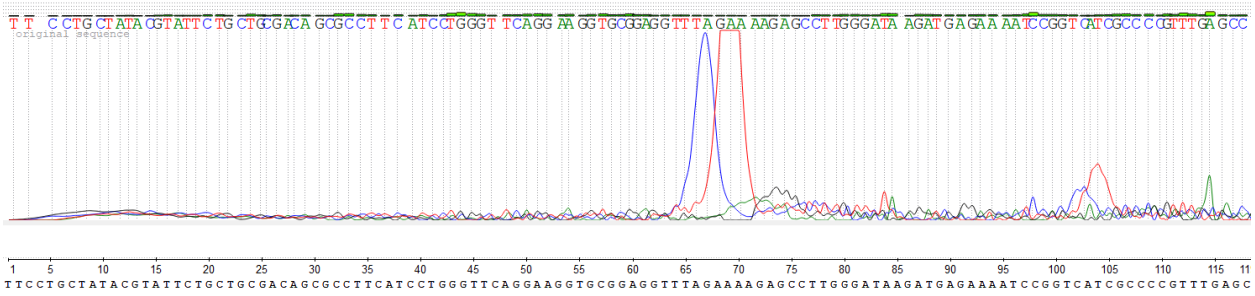


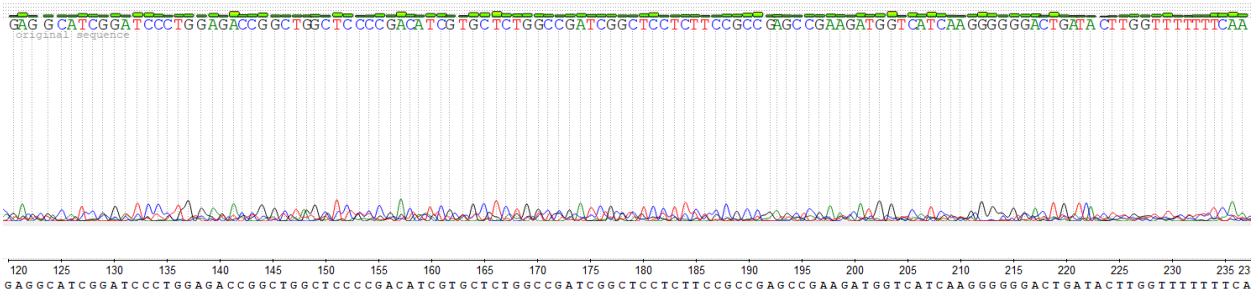


ICC1*ureC*1R

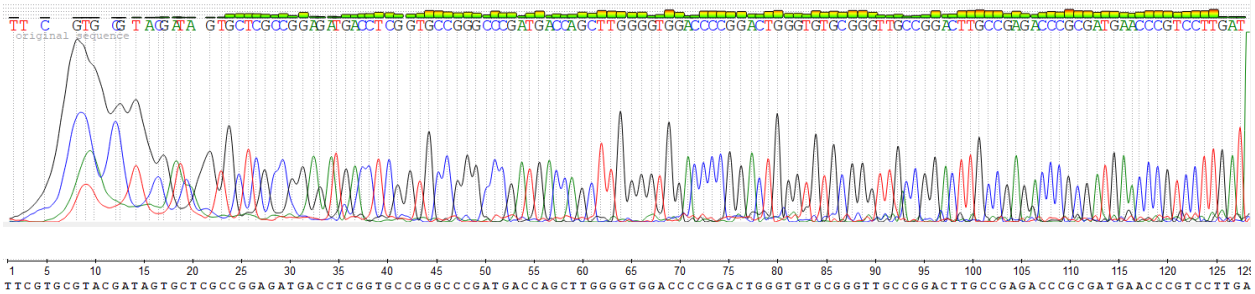


ICC1*ureC*2F

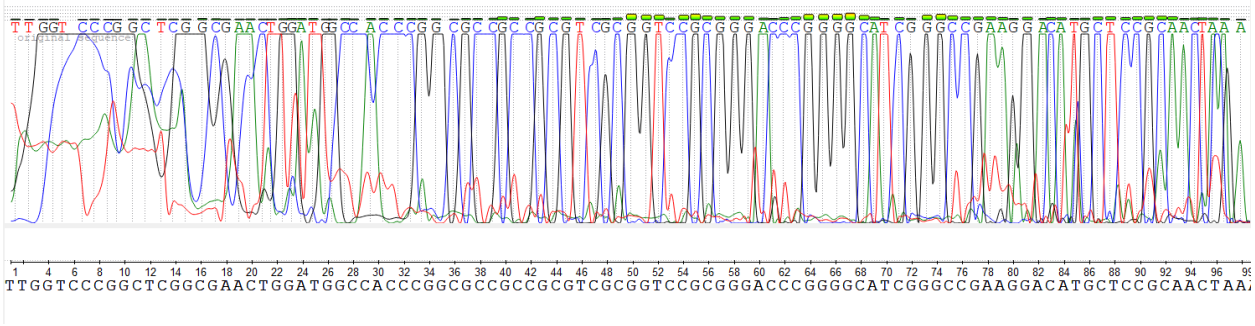




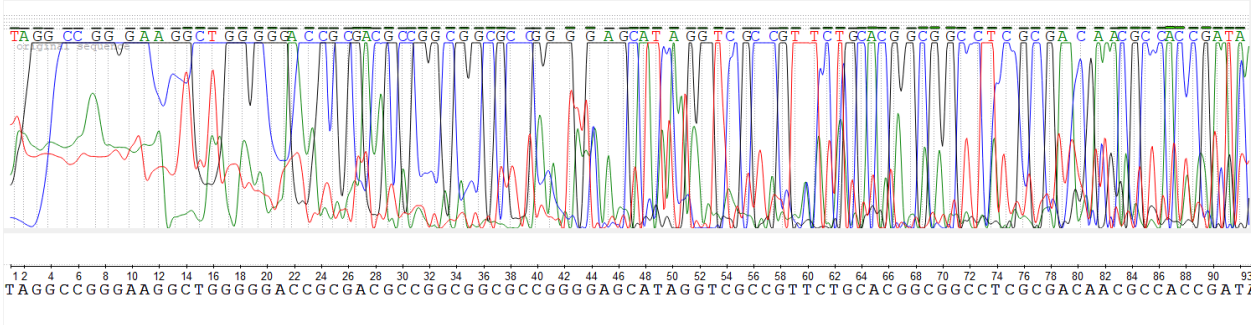
ICC1*ureC*2R



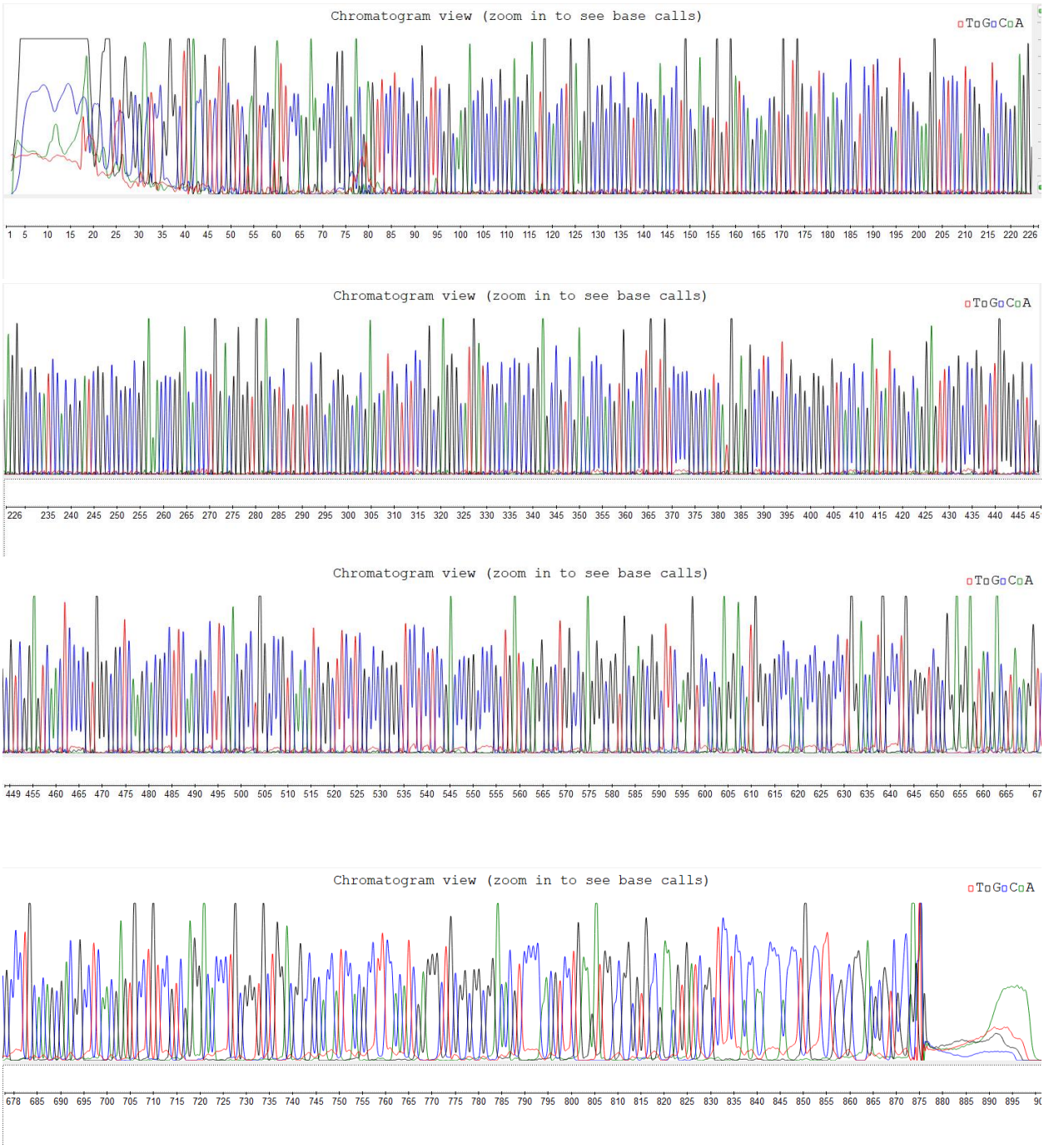
ICC1*ureC*3F



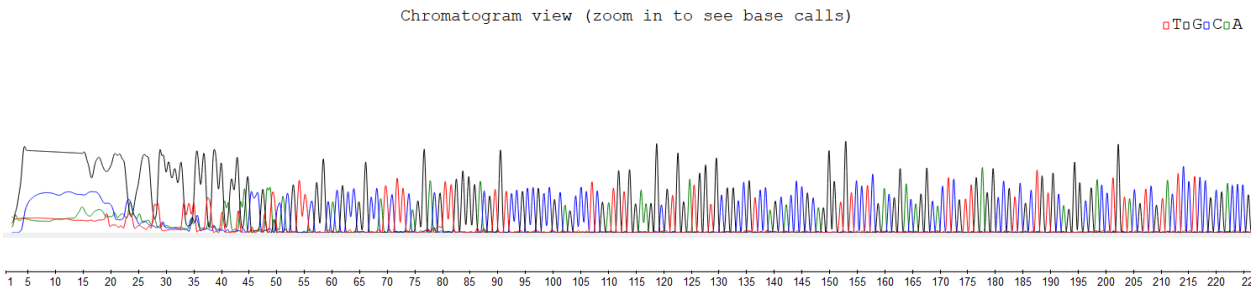
ICC1*ureC*3R

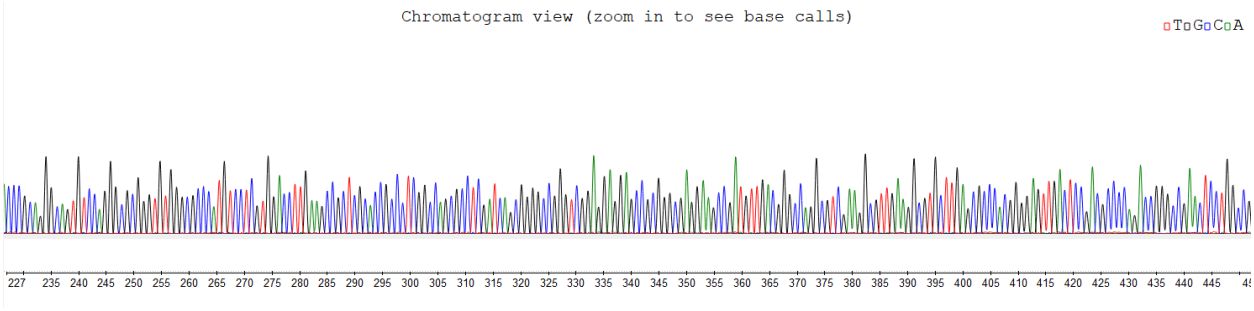


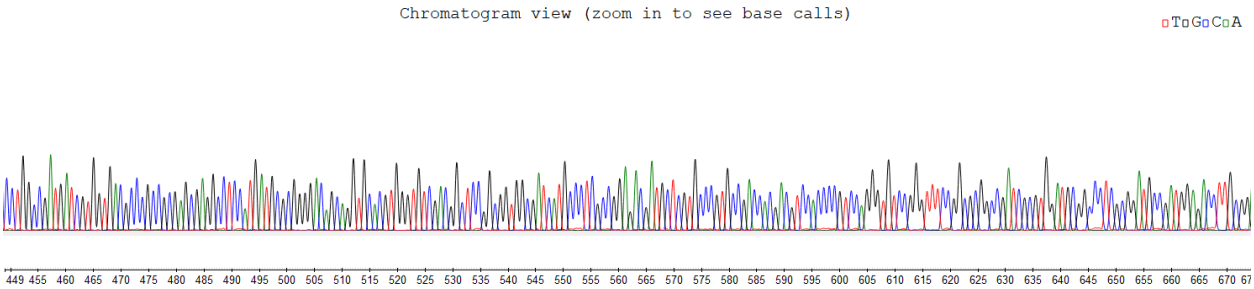
ICC1*ureC*4F

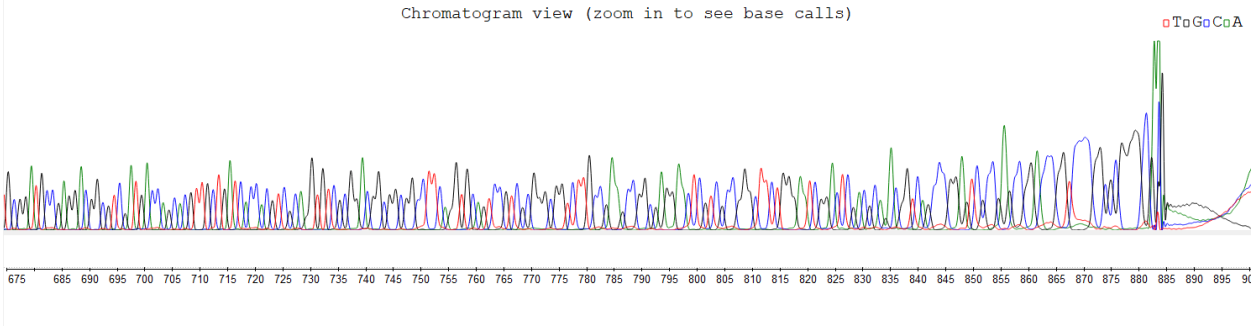

Chromatogram view (zoom in to see base calls)
T G C A
1 5 10 15 20 25 30 35 40 45 50 55 60 65 70 75 80 85 90 95 100 105 110 115 120 125 130 135 140 145 150 155 160 165 170 175 180 185 190 195 200 205 210 215 220 226
Chromatogram view (zoom in to see base calls)
T G C A
226 235 240 245 250 255 260 265 270 275 280 285 290 295 300 305 310 315 320 325 330 335 340 345 350 355 360 365 370 375 380 385 390 395 400 405 410 415 420 425 430 435 440 445 45
Chromatogram view (zoom in to see base calls)
T G C A
449 455 460 465 470 475 480 485 490 495 500 505 510 515 520 525 530 535 540 545 550 555 560 565 570 575 580 585 590 595 600 605 610 615 620 625 630 635 640 645 650 655 660 665 67
Chromatogram view (zoom in to see base calls)
T G C A
678 685 690 695 700 705 710 715 720 725 730 735 740 745 750 755 760 765 770 775 780 785 790 795 800 805 810 815 820 825 830 835 840 845 850 855 860 865 870 875 880 885 890 895 90

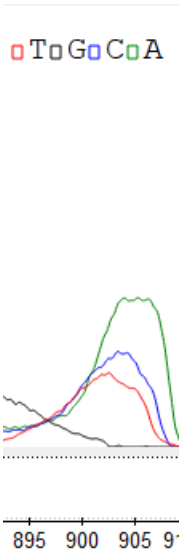
ICC1*ureC*4R


Chromatogram view (zoom in to see base calls)
T G C A
1 5 10 15 20 25 30 35 40 45 50 55 60 65 70 75 80 85 90 95 100 105 110 115 120 125 130 135 140 145 150 155 160 165 170 175 180 185 190 195 200 205 210 215 220

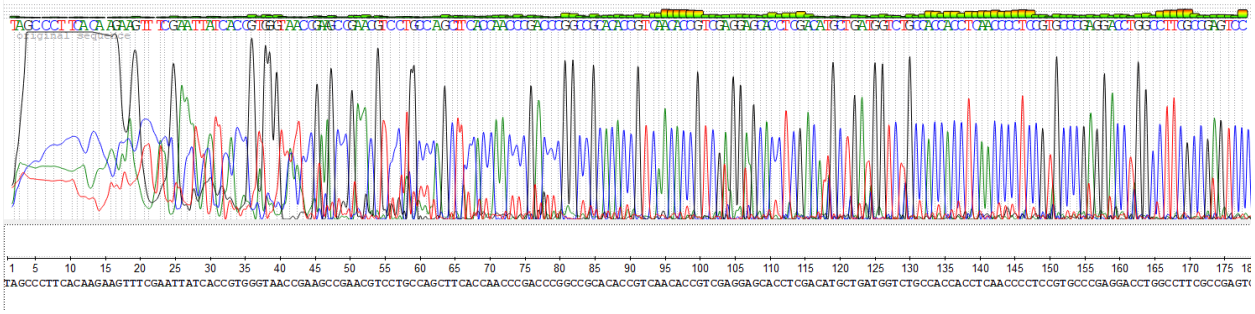

Chromatogram view (zoom in to see base calls)
T G C A
227 235 240 245 250 255 260 265 270 275 280 285 290 295 300 305 310 315 320 325 330 335 340 345 350 355 360 365 370 375 380 385 390 395 400 405 410 415 420 425 430 435 440 445

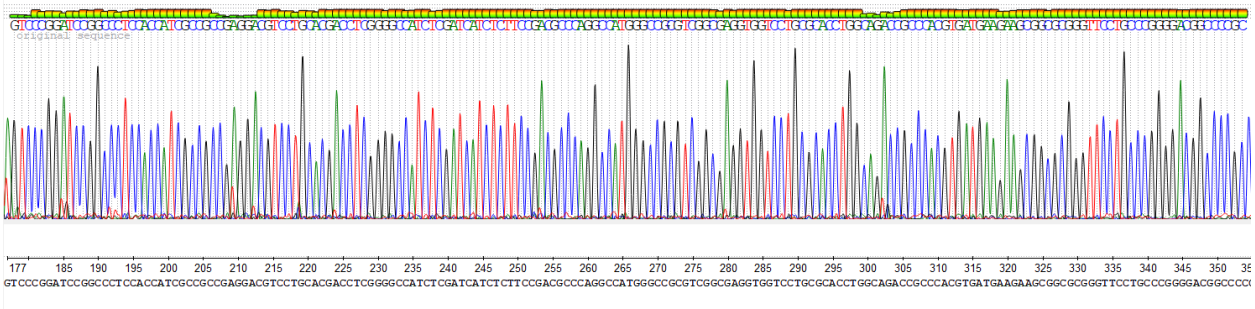

Chromatogram view (zoom in to see base calls)
T G C A
449 455 460 465 470 475 480 485 490 495 500 505 510 515 520 525 530 535 540 545 550 555 560 565 570 575 580 585 590 595 600 605 610 615 620 625 630 635 640 645 650 655 660 665 670

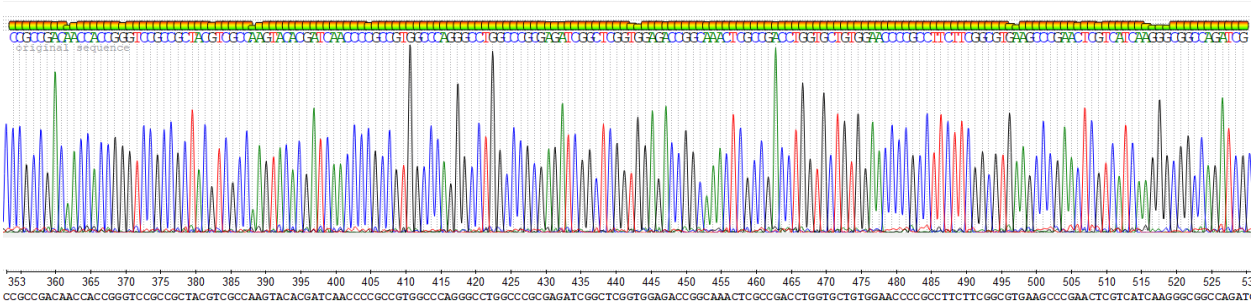

Chromatogram view (zoom in to see base calls)
T G C A
675 685 690 695 700 705 710 715 720 725 730 735 740 745 750 755 760 765 770 775 780 785 790 795 800 805 810 815 820 825 830 835 840 845 850 855 860 865 870 875 880 885 890 895

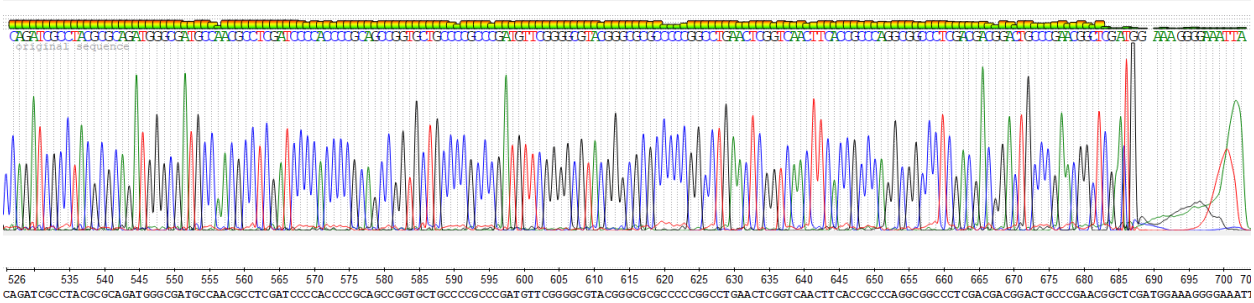

T G C A
895 900 905

ICC1*ureC*5F

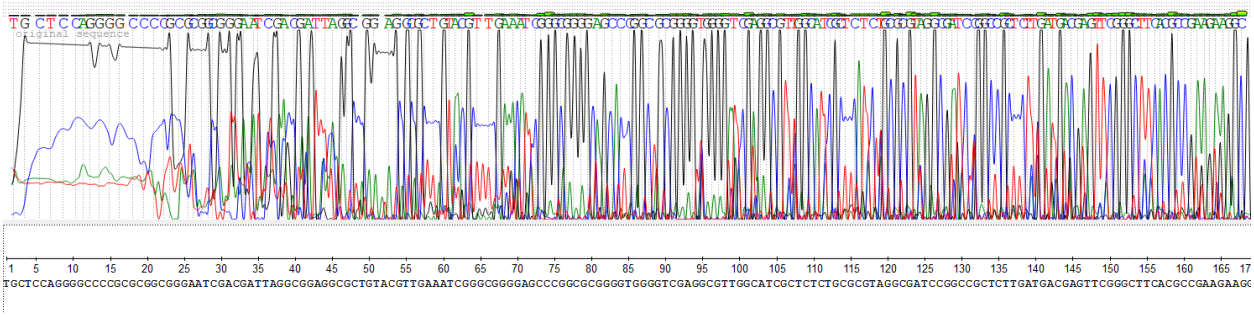

1 5 10 15 20 25 30 35 40 45 50 55 60 65 70 75 80 85 90 95 100 105 110 115 120 125 130 135 140 145 150 155 160 165 170 175
TAGCCCTTCACAAGAAGTTTCGAATTATCACCGTGGGTAACCGAAGCCGAACGTCCTGCCAGCTTCACCAACCCGACCCGGCCGCACACCGTCAACACCGTCGAGGAGCACCTCGACATGCTGATGGTCTGCCACCACCTCAACCCCTCCGTGCCCGAGGACCTGGCCTTCGCCGAGT

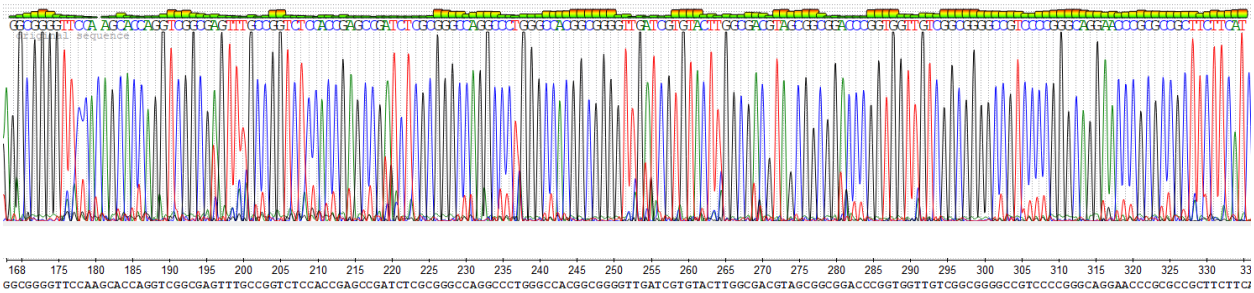

177 185 190 195 200 205 210 215 220 225 230 235 240 245 250 255 260 265 270 275 280 285 290 295 300 305 310 315 320 325 330 335 340 345 350
GTCCCGGATCCGGCCCTCCACCATCGCCGCCGAGGACGTCCTGCACGACCTCGGGGCCATCTCGATCATCTCTTCCGACGCCCAGGCCATGGGCCGCGTCGGCGAGGTGGTCCTGCGCACCTGGCAGACCGCCCACGTGATGAAGAAGCGGCGCGGGTTCCTGCCCGGGGACGGCCCC

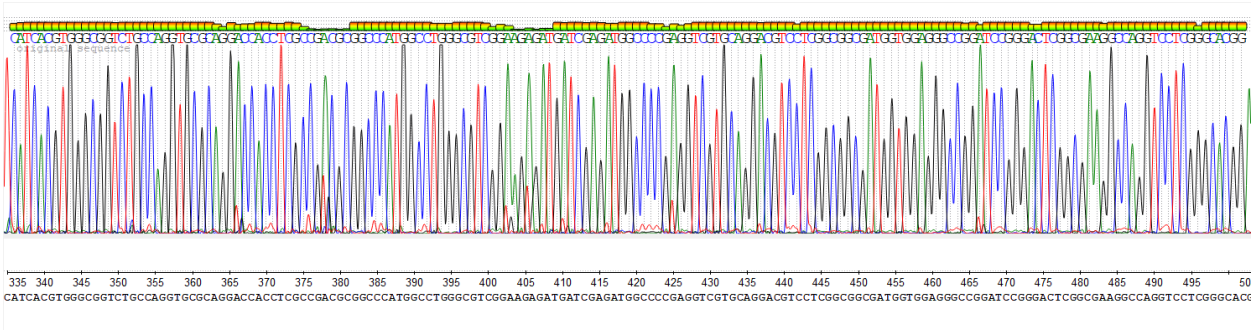

original sequence

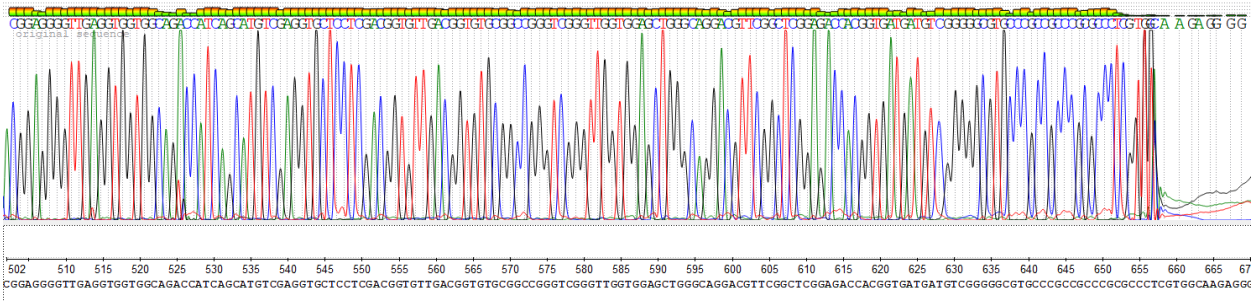

original sequence

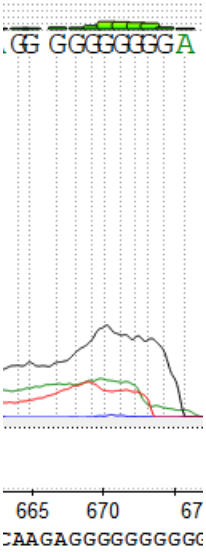
ICC1*ureC*5R


original sequence

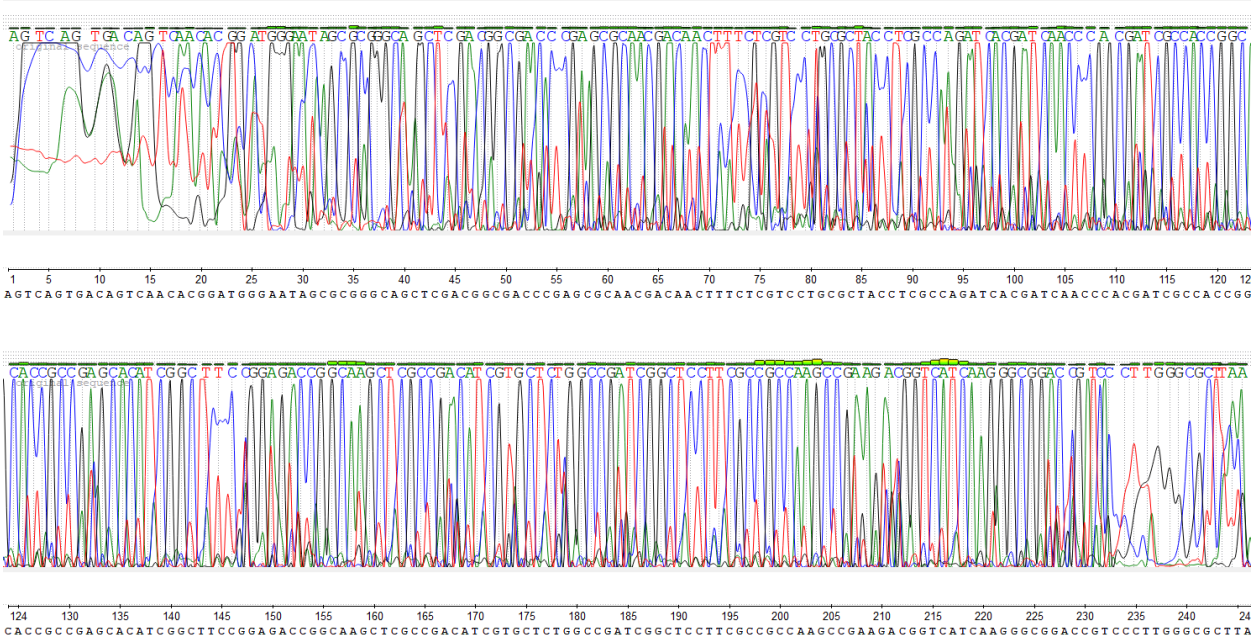

original sequence





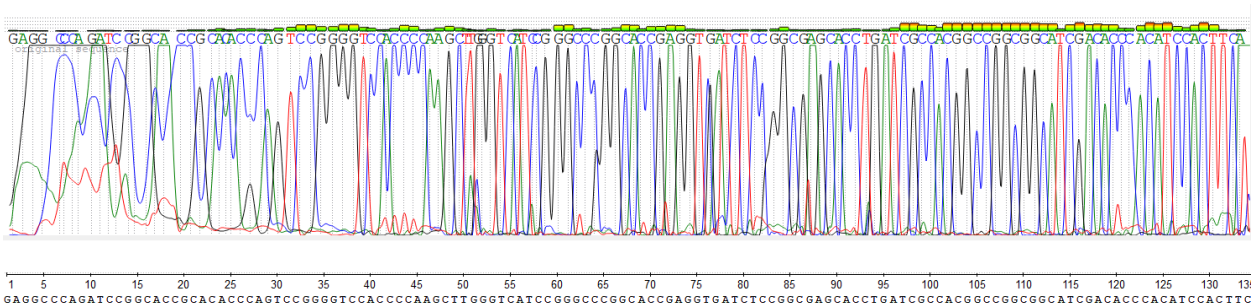


ICC4*ureC*1F

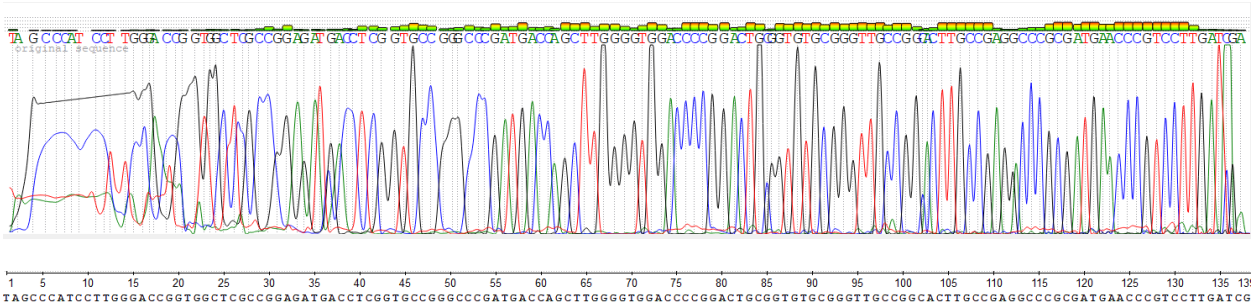


ICC4*ureC*1R

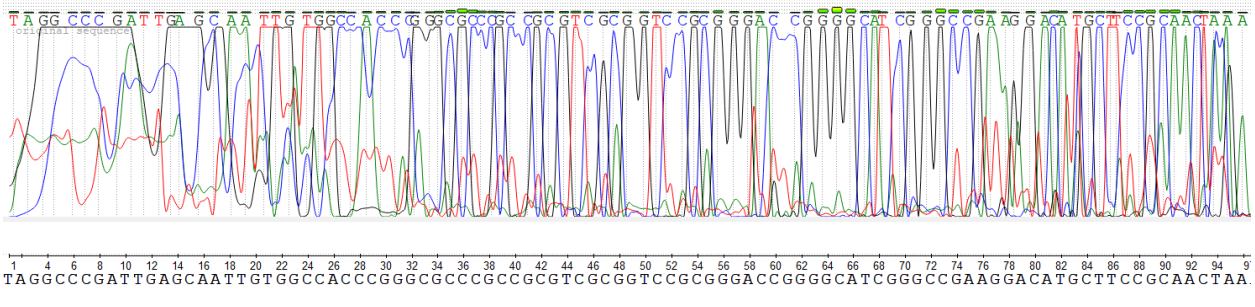
ICC4*ureC*2F



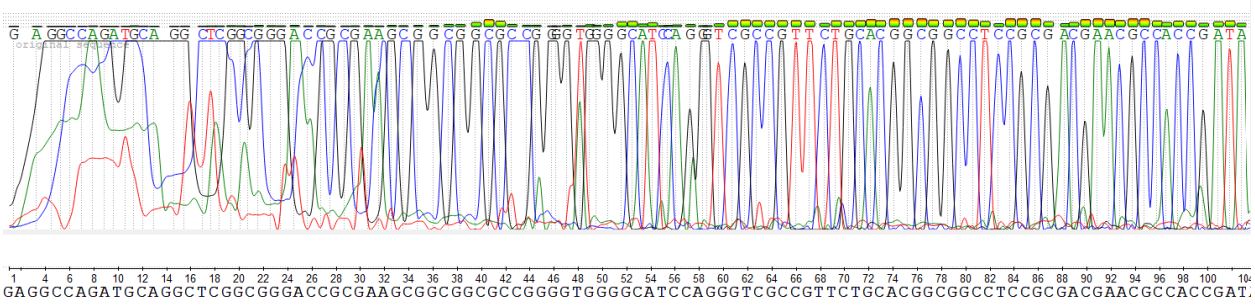
ICC4*ureC*2R



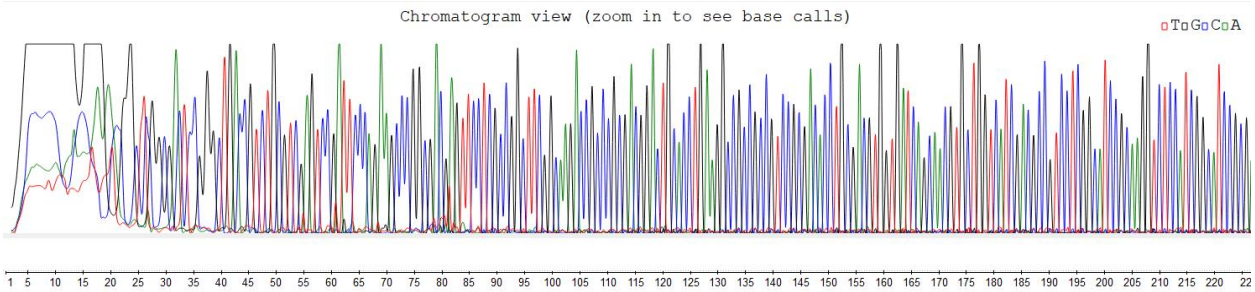
ICC4*ureC*3F

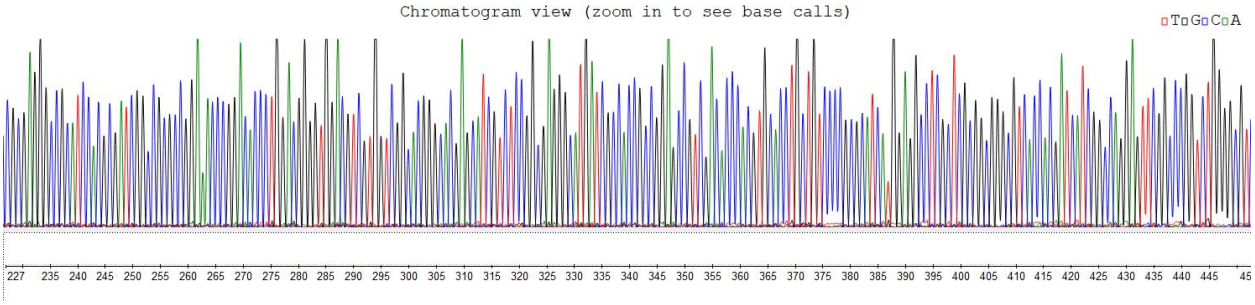


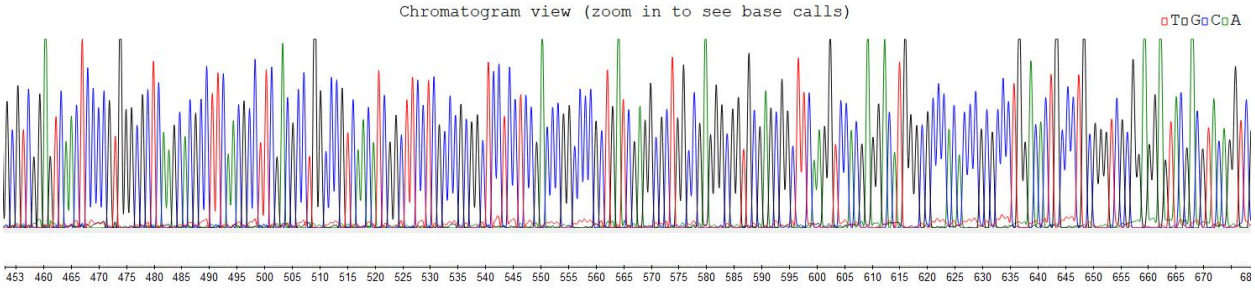
ICC4*ureC*3R

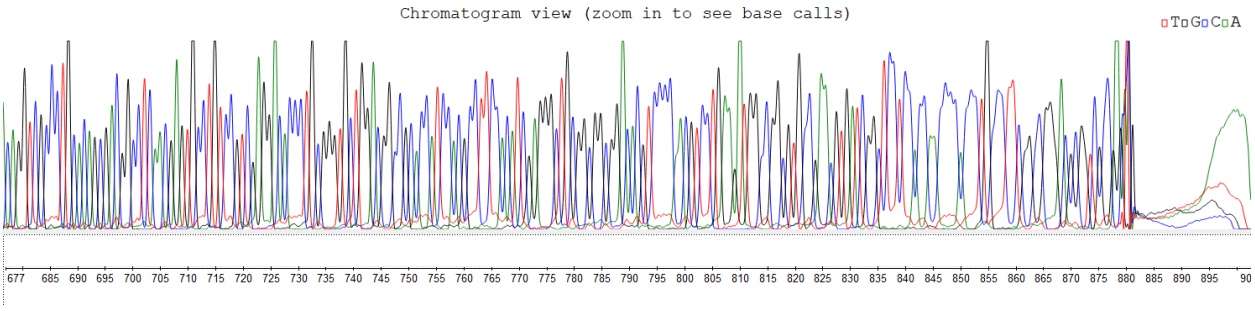


ICC4*ureC*4F

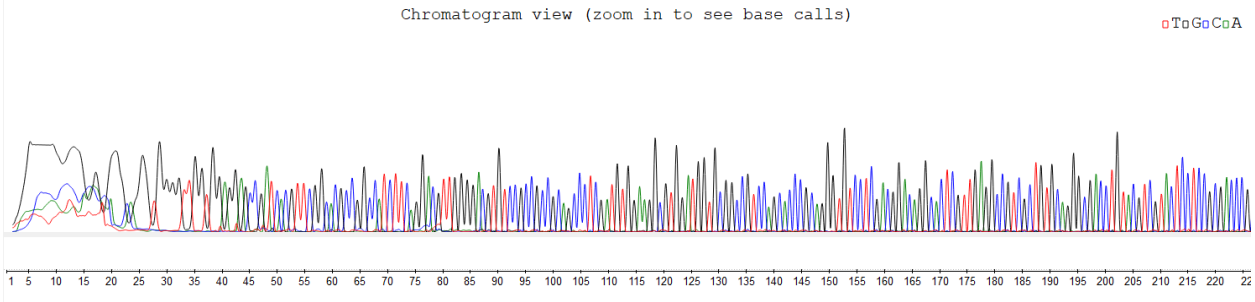

Chromatogram view (zoom in to see base calls)
T G C A
1 5 10 15 20 25 30 35 40 45 50 55 60 65 70 75 80 85 90 95 100 105 110 115 120 125 130 135 140 145 150 155 160 165 170 175 180 185 190 195 200 205 210 215 220 22

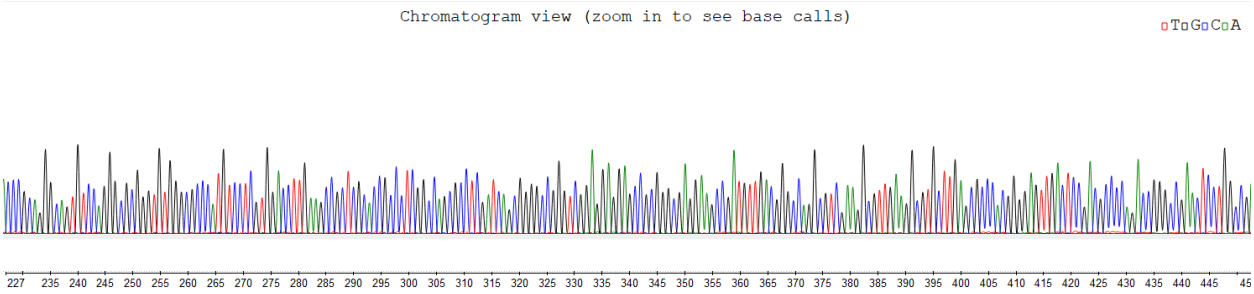

Chromatogram view (zoom in to see base calls)
T G C A
227 235 240 245 250 255 260 265 270 275 280 285 290 295 300 305 310 315 320 325 330 335 340 345 350 355 360 365 370 375 380 385 390 395 400 405 410 415 420 425 430 435 440 445 45

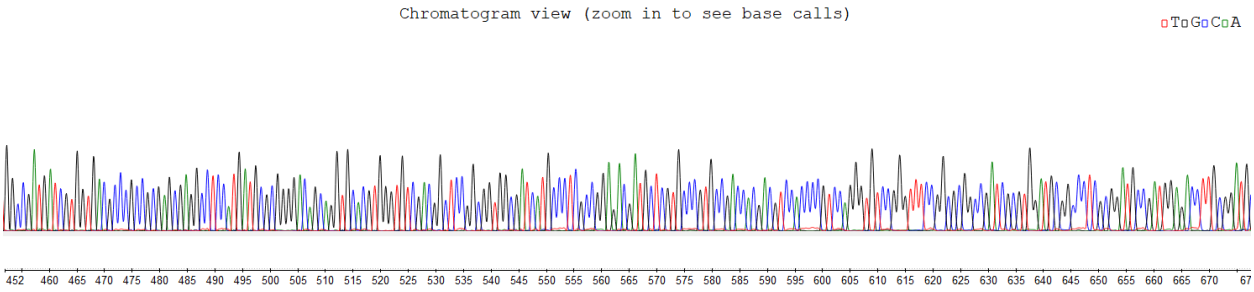

Chromatogram view (zoom in to see base calls)
T G C A
453 460 465 470 475 480 485 490 495 500 505 510 515 520 525 530 535 540 545 550 555 560 565 570 575 580 585 590 595 600 605 610 615 620 625 630 635 640 645 650 655 660 665 670 68

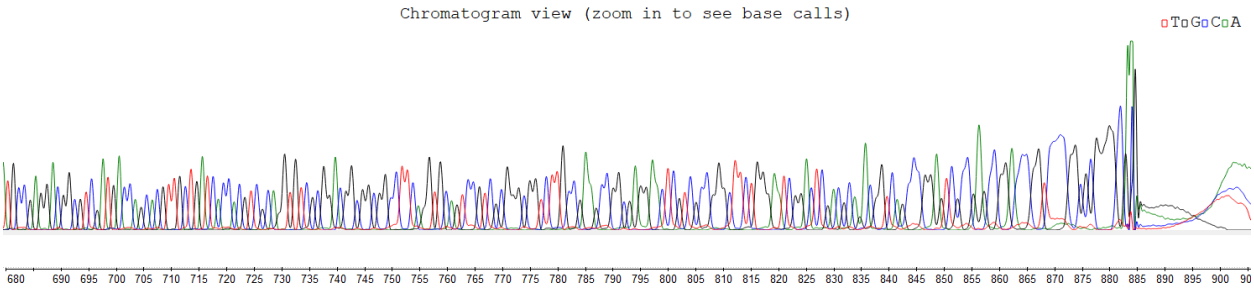

Chromatogram view (zoom in to see base calls)
T G C A
677 685 690 695 700 705 710 715 720 725 730 735 740 745 750 755 760 765 770 775 780 785 790 795 800 805 810 815 820 825 830 835 840 845 850 855 860 865 870 875 880 885 890 895 90

ICC4*ureC*4R

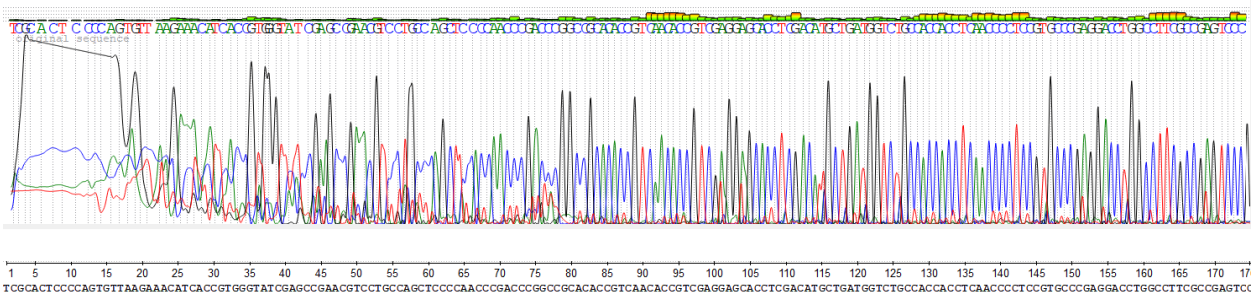

Chromatogram view (zoom in to see base calls)
T G C A
1 5 10 15 20 25 30 35 40 45 50 55 60 65 70 75 80 85 90 95 100 105 110 115 120 125 130 135 140 145 150 155 160 165 170 175 180 185 190 195 200 205 210 215 220 22

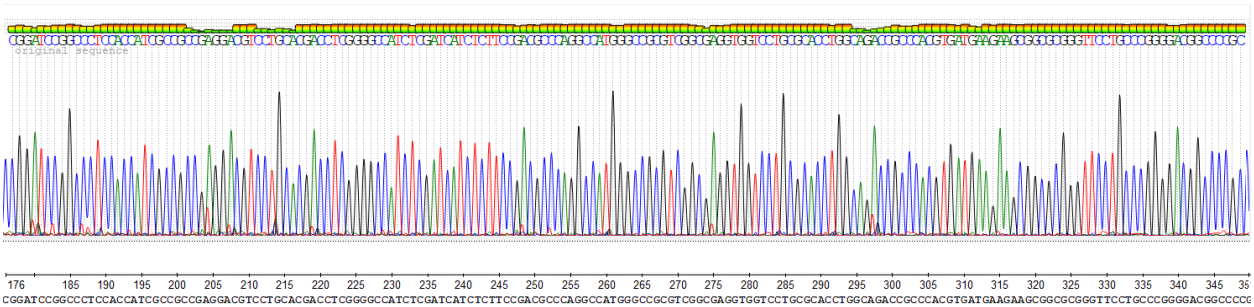

Chromatogram view (zoom in to see base calls)
T G C A
227 235 240 245 250 255 260 265 270 275 280 285 290 295 300 305 310 315 320 325 330 335 340 345 350 355 360 365 370 375 380 385 390 395 400 405 410 415 420 425 430 435 440 445 45

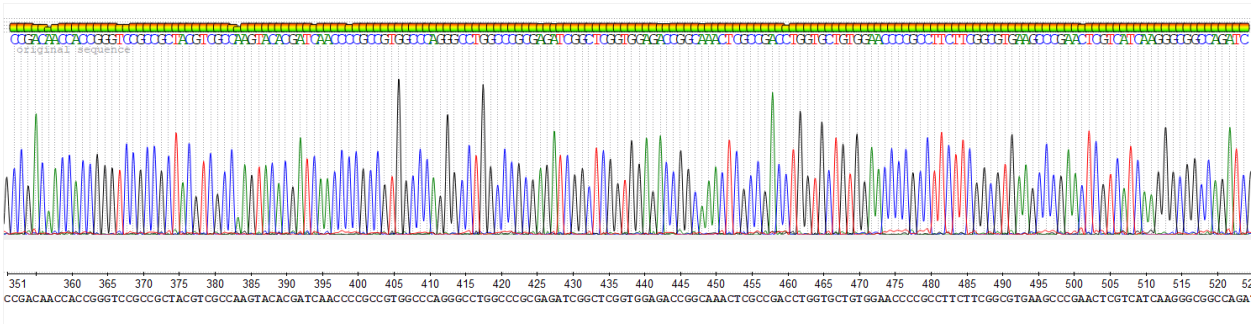

Chromatogram view (zoom in to see base calls)
T G C A
452 460 465 470 475 480 485 490 495 500 505 510 515 520 525 530 535 540 545 550 555 560 565 570 575 580 585 590 595 600 605 610 615 620 625 630 635 640 645 650 655 660 665 670 67

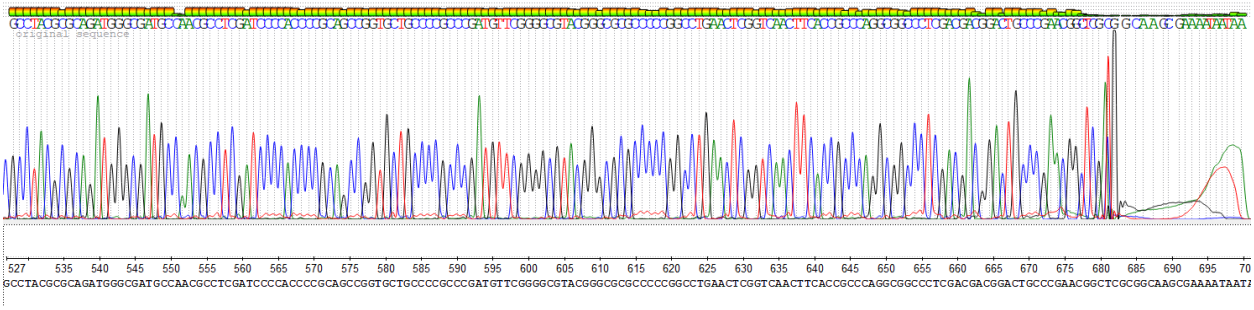

Chromatogram view (zoom in to see base calls)
T G C A
680 690 695 700 705 710 715 720 725 730 735 740 745 750 755 760 765 770 775 780 785 790 795 800 805 810 815 820 825 830 835 840 845 850 855 860 865 870 875 880 885 890 895 900 90

ICC4*ureC*5F

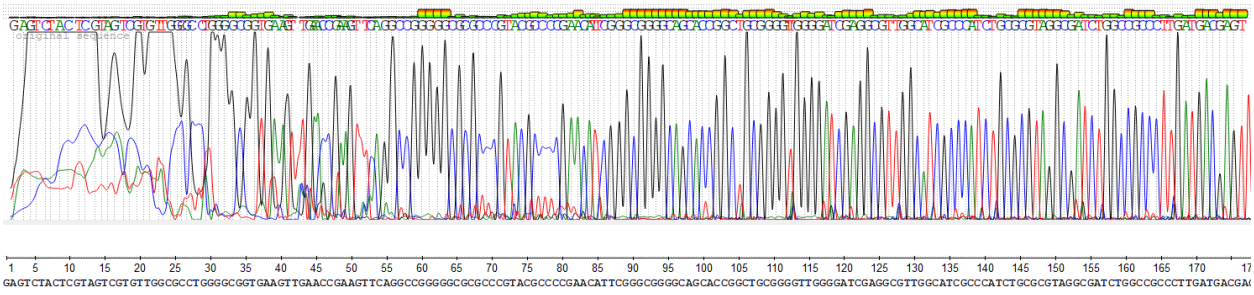

Original sequence
1 5 10 15 20 25 30 35 40 45 50 55 60 65 70 75 80 85 90 95 100 105 110 115 120 125 130 135 140 145 150 155 160 165 170 17
TCGCACTCCCCAGTGTTAAGAAACATCACCGTGGGTATCGAGCCGAACGTCCTGCCAGCTCCCCAACCCGACCCGGCCGCACACCGTCAACACCGTCGAGGAGCACCTCGACATGCTGATGGTCTGCCACCACCTCAACCCCTCCGTGCCCGAGGACCTGGCCTTCGCCGAGTCC

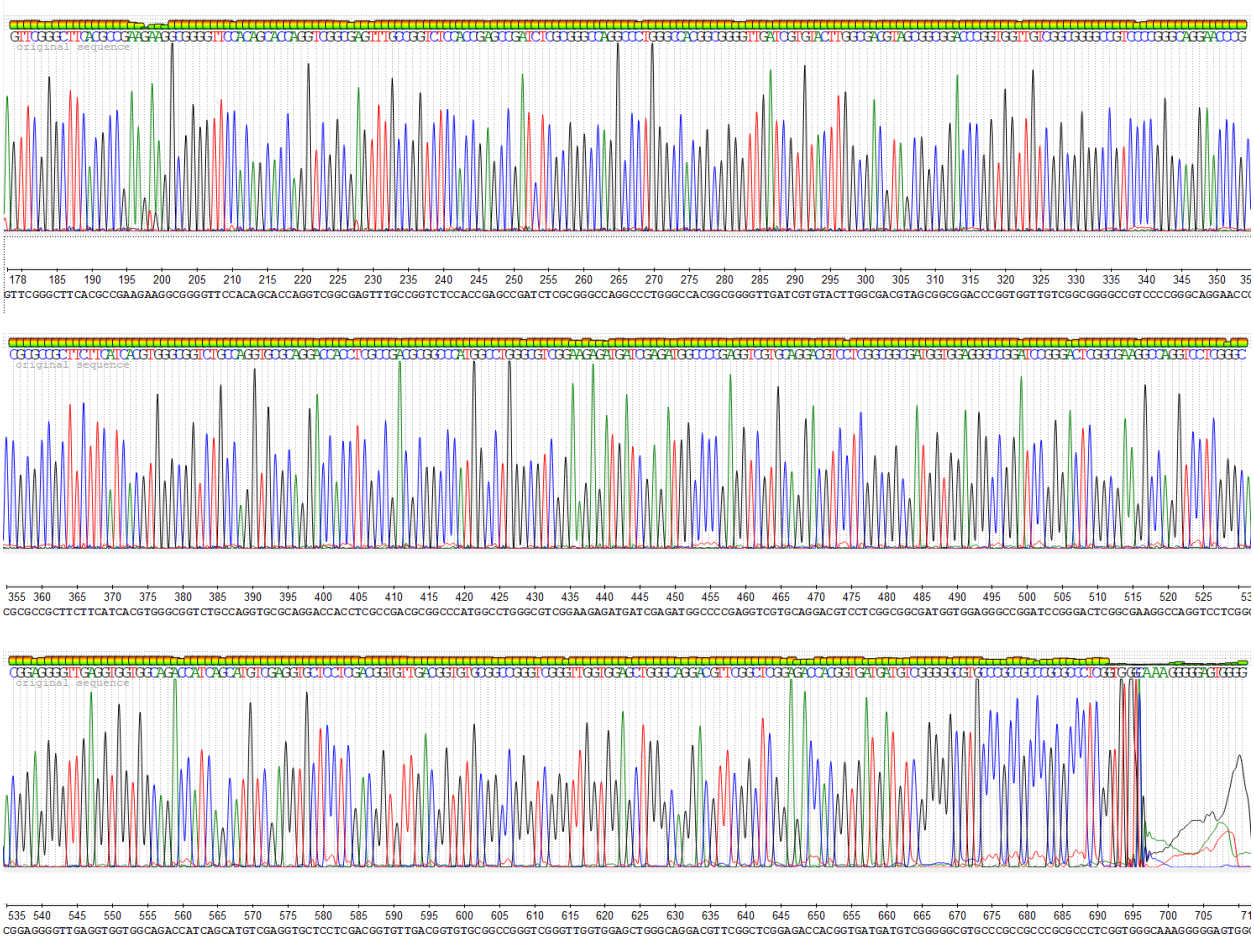






ICC4*ureC*5R




original sequence
178 185 190 195 200 205 210 215 220 225 230 235 240 245 250 255 260 265 270 275 280 285 290 295 300 305 310 315 320 325 330 335 340 345 350
GTTCGGGCTTCACGCCGAAGAAGGCGGGGTTCCACAGCACCAGGTCGGCGAGTTTGCCGGTCTCCACCGAGCCGATCTCGCGGGCCAGGCCCTGGGCCACGGCGGGGTTGATCGTGTACTTGGCGACGTAGCGGCGGACCCGGTGGTTGTCGGCGGGGCCGTCCCCGGGCAGGAACC
original sequence
355 360 365 370 375 380 385 390 395 400 405 410 415 420 425 430 435 440 445 450 455 460 465 470 475 480 485 490 495 500 505 510 515 520 525
CGCGCCGCTTCTTCATCACGTGGGCGGTCTGCCAGGTGCGCAGGACCACCTCGCCGACGCGGCCCATGGCCTGGGCGTCGGAAGAGATGATCGAGATGGCCCCGAGGTCGTGCAGGACGTCCTCGGCGGCGATGGTGGAGGGCCGGATCCGGGACTCGGCGAAGGCCAGGTCCTCGG
original sequence
535 540 545 550 555 560 565 570 575 580 585 590 595 600 605 610 615 620 625 630 635 640 645 650 655 660 665 670 675 680 685 690 695 700 705
CGGAGGGGTTGAGGTGGTGGCAGACCATCAGCATGTCGAGGTGCTCCTCGACGGTGTTGACGGTGTGCGGCCGGGTCGGGTTGGTGGAGCTGGGCAGGACGTTCGGCTCGGAGACCACGGTGATGATGTCGGGGGCGTGCCCGCCGCCCGCGCCCTCGGTGGGCAAAGGGGGAGTGG