BACTERIA BEWARE: LET THE PHAGE HUNT BEGIN!

A project by: Irma Komljenovic, Vito Menjak, Karolina Guzauskaite, Yi Ming Gan, and Cédric Lood

Introduction

The tree of life

The tree of life is a theoretical model connecting all species - both living and extinct - based on the hypothesis of a single origin of life (Last Universal Common Ancestor, or LUCA). It was proposed by Charles Darwin in his 1859 book "On the Origin of Species". If originally, species were grouped using morphological similarities with the obvious limitations that entails, today's tree of life has been largely redefined by using molecular information found in the genomes of the various species. Typically, the universal mechanisms encoded in deoxyribonucleic acid (DNA) for the ribosomal machinery, responsible the synthesis of proteins. Using similarities in this genomic encoding, all three domains of life, bacteria, archae, and eukaryotes, were connected in a single tree, rooted in this LUCA (see Figure 1).

Notoriously, viruses are missing on the tree of life. Viruses have a parasitic lifestyle and depend on infecting hosts to propagate, hence do not possess the ribosomal machinery. As such, they cannot be connected to the tree of life as a "Fourth domain", this is also fitting to their status as entities that are not considered to be "alive".

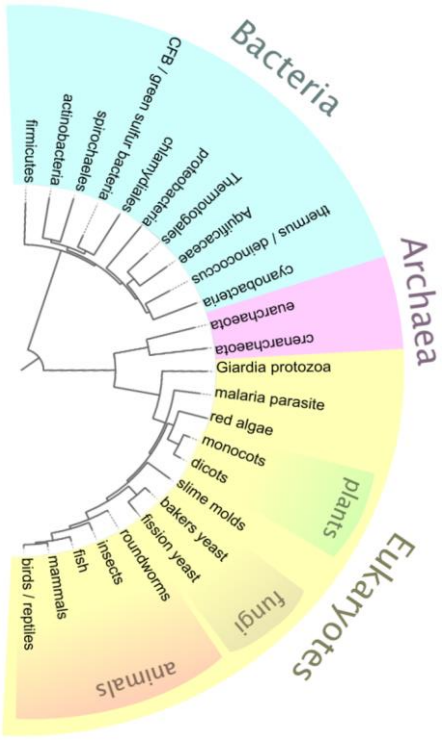
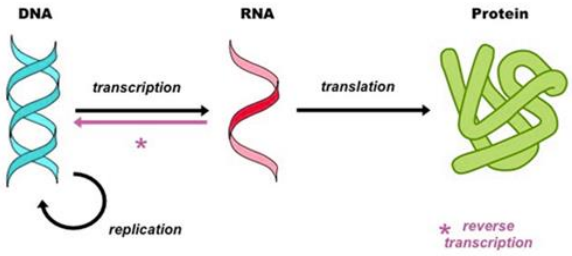


FIGURE 1: THE TREE OF LIFE CONNECTS ALL SPECIES LIVING TODAY

The molecules of life

Deoxyribonucleic acid (DNA) is often referred to as the key of life. It consists of monomers called nucleotides that are assembled using a sugar deoxyribose, a phosphate group, and a nitrogenous base. There are two main groups of nitrogenous bases: purines (adenine and guanine – or A and G) and pyrimidines (cytosine and thymine – or C and T). DNA is a double stranded helix where adenine pairs with thymine and cytosine pairs with guanine. The order of these bases is exactly what forms the coding instructions in the genome. However, the question is how does the interaction of bases enable cells to live, and how are the instructions used?

FIGURE 2: THE CENTRAL DOGMA OF MOLECULAR BIOLOGY



This is explained by central dogma of molecular biology which illustrates the flow of genetic information (see Figure 2). Looking into the

dogma, crucial steps for gene expression to occur are DNA transcription (copying of DNA encoded information to mRNA) and translation to proteins which are essential for cell to function as they perform variety of functions such as structural, storage, transport, cellular communications, movement, defense and catalytic. On the other hand, some viruses instead of DNA as a genetic material use RNA. Because of that these viruses (chiefly the members of retrovirus group, such as HUV) have an enzyme called reverse transcriptase that produces double stranded DNA from single stranded RNA genome. Later on, this virus DNA is integrated in host's genome and now is called provirus. This way, it undergoes usual transcription and translation processes to express the genes. Yet, it is very important to note that gene expression is unidirectional, meaning that it goes only from DNA to proteins. This highlights the importance of DNA as the hub of genetic information.

Bacteria and their viruses

Bacteria are unicellular organisms that are found in a wide range of habitats. For most of life's history on earth, bacteria have been the dominant form of life and have been shaping earth's geochemical cycles. They are extremely numerous, and a typical person carries about as many bacteria as human cells (40 trillion bacteria/30 trillion cells). Although some bacteria cause disease in humans, most them are harmless, and some are beneficial, helping us to degrade compounds we wouldn't otherwise be capable to degrade.

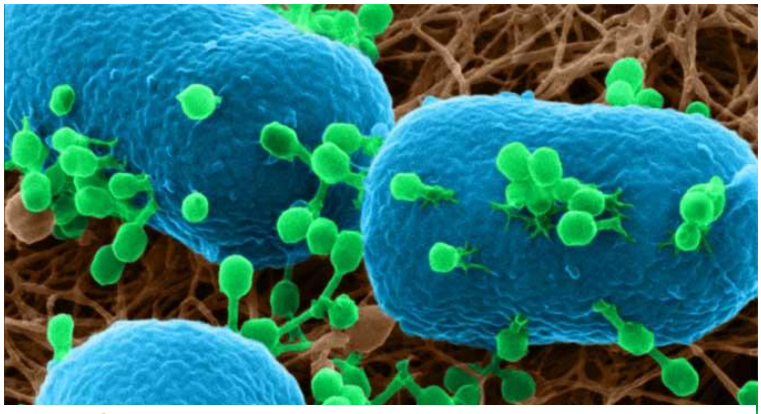


FIGURE 3: BACTERIOPHAGES INFECTING A BACTERIA

Viruses of bacteria, known as bacteriophages (bacteria eaters) are the most common organisms found on earth (estimated to be in the range of 10^31 phages). They can infect bacteria by injecting their genetic material into the cell, eventually using the cellular machinery to create copies of themselves as traditional viruses do.

Key objectives of our project

One of the goals of our work was to identify the species of bacteria found in the collected water samples and samples of isolated bacterial genomic DNA, place and locate the organisms on a phylogenetic tree, and observe the presence, activity, and efficacy of bacteriophages found in the samples. We identified the species by means of biochemical testing and nanopore DNA sequencing. We aimed to identify the bacteria with the biochemical tests by looking at their characteristics (e.g. Gram staining, determining the types of molecules they break down, etc.). We used two kinds of biochemical tests in the project: selective media and API20E strips. We used blood, chocolate and MacConkey agar plates as selective media to isolate Gram-negative bacteria. The selective media helped us isolate single colony forming units (CFU) and select the bacteria that would be tested with the API20E strips. The API20E strips are biochemical tests used to determine the species of bacteria from the Enterobacteriaceae genus. The genus and species classification is determined by the reaction (or lack thereof) of bacteria with characteristic sugar fermentation capabilities, metabolites, etc.

The sequencing of multiple bacterial genomic DNA was performed using nanopore technology, specifically, the MinION device from Oxford Nanopore Technology. The aim of the sequencing was to show the power of sequencing technologies to accurately determine the species (and, possibly, subtype) of bacteria, and place the organisms on a phylogenetic tree. Obtaining the genomic sequence of an organism is very rich in information, enabling many more types of analysis than biochemical identification would, and allowing us to understand the diversity at the population level of bacteria within a given species.

Materials & Methods

Preparation of media

We prepared hard LB agar by mixing 15 g of agar, 10 g of tryptone, 5 g of yeast extract and 10 g of Sodium chloride (NaCl) in a non-sterile bottle. After that, we added deionized water until the volume of the solution reached 1 L. We shook the bottles until the solid ingredients dissolved. We then uncapped the bottles slightly, covered the caps with autoclave tape and put the bottles in the autoclave. After the autoclaving process had finished, we fully closed the bottle caps and the medium was left to cool before we poured it into sterile Petri dishes. Soft agar was prepared in a similar fashion, using only 7g of agar, and after preparation the bottles were kept on the sand burner, whose temperature was between 50 and 60°C, due to the tendency of soft agar to solidify quickly (and high temperature would kill inoculated bacteria). Liquid LB was prepared similarly, omitting the agar.

Blood, MacConkey and chocolate agar plates that we used in the project were received finished from the public health institute in Zagreb. Blood agar was prepared by suspending 28 g of nutrient agar powder in 1 L of distilled water. The mixture was heated while stirred until the components were fully dissolved. Once dissolved, the mixture was autoclaved in the same manner as previously mentioned. After autoclaving, the nutrient agar was cooled to 45-50°C, after which 5% (vol/vol) sterile defibrinated blood warmed to room temperature was added and mixed gently. The mixture should be poured into sterile plates while liquid. MacConkey agar was prepared by mixing 17 g of pancreatic digest of gelatin (peptone), 3 g of proteose peptone (meat and casein), 10 g of lactose monohydrate, 1.5 g of bile salts, 5 g of Sodium chloride, 0.03 g of Neutral red, 0.001 g of Crystal Violet and 13.5 g of agar. Distilled water was added until the volume of the solution reached 1 L. The final pH at 25°C was set to 7.1 +/- 0.2. The solution was heated to boiling so that the medium was completely dissolved, after which it was autoclaved in the same manner as hard LB agar. The final product should be mixed well before being poured into sterile Petri dishes. Chocolate agar was prepared by heating the blood agar base after the blood (preferably of a horse or a sheep) has been added slowly to 56°C in a water bath.

Preparation of agarose gel for gel electrophoresis

50X TAE buffer was prepared by adding 242 g of Tris free base and 18.61 g of Disodium EDTA to 700 mL of dH_2O and stirred until completely dissolved. We added 57.1 mL of Glacial Acetic Acid and the amount of water needed to adjust the volume to 1 L to the solution. We made 500 mL of 1x TAE buffer by mixing 10 mL of 50x TAE buffer and 490 mL of dH_2O. We made the agarose gel for gel electrophoresis by mixing 0.21 g of agarose and 30 mL of 1x TAE buffer in a non-sterile bottle. The uncapped bottle was placed in the microwave at the highest possible temperature and 3-4 s after the contents of the bottle boiled, then 1 µL of SYBR Green was added, and the mixture was quickly

poured into a gel container previously sealed with tape. We inserted the combs and the gel was left to cool, time after which the tape was detached. We removed the combs when the gel was laid in the electrophoresis tank filled with 1X TAE buffer.

Collection of environmental samples

We looked for water we expected to have bacteria (therefore bacteriophages) from the river Orljava, the stream Veličanka, the sewer and two puddles found in the Stjepana Radica street. We used sterile falcon tubes and tried our best to open them as close to the water and close them after taking a sample to minimize outside contamination. We took a photo at every collection site to document the location and time of taking the sample. Each sample was marked with a number on the lid there and later received a line in the database table describing it (see Table 2).

Isolation of bacteria

Original samples were serially diluted before being spread on a petri dish. In short, we took one milliliter of the original sample, placed it in an Eppendorf tube and mixed it with nine milliliters of sterile water, thus creating the first serial dilution by the factor of 10 (SD 10^1). When we took one milliliter of SD 10^1 and mixed it with nine milliliters of sterile water in another Eppendorf we made the second serial dilution or SD 10^2. We made four dilution stages for every sample we worked with.

We then plated every serial dilution onto a plate with LB agar and left it incubate at room temperature (~30°C) . As you will in the database (table 2), the number of colony forming units (CFU) was reduced ten times with every dilution and that's how we calculated the number of colonies the original sample would have. The third and fourth serial dilution actually had no CFU so we decided to remove the fourth from the database table.

Except from LB agar which is non-selective, we also worked with MacConkey, Blood agar and Chocolate agar. MacConkey agar - selective differential feeding ground for isolating GRAM negative bacteria. Blood agar - non-selective, enriched with blood. It is used for isolating bacteria based on their hemolytic properties. Chocolate agar - non-selective feeding ground for isolating GRAM positive bacteria, enriched with cooked blood. It is used for isolation of pathogenic bacteria.

We placed (amount) of the original sample in one corner of a dish with one of these types of agar and diluted it by streaking. We decided to further research some colonies that stood out by reacting with the medium and those were the bacteria we identified using APi strips later on.

Isolation of phage

To separate phages from bacteria, we took 20 mL of water from the original sample into a syringe and added a filter to its tip. The pores of the filter were 220 nanometers in size so that only the phages and small molecules can pass through while the bacteria stayed behind in it along with any particles of dirt or plastic that might have been in the sample. We then put the extracted phages on the double agar overlay assay to see bacterial lysis.

Biochemical tests with API20E strip

We used the API20E strips according to the instructions received with the set. The strips were handled in sterile conditions. We transferred 1 bacterial CFU with a loop to a sterile falcon tube containing 10 mL sterile dH_2O. All of the colony samples were taken from different MacConkey agar

plates, due to the fact that all of the members of the Enterobacteriaceae family are Gram-negative. We pipetted the prepared samples in the testing chambers. Immediately after being pipetted in the chambers, the samples mixed with the reagents and changed colour. Most of the chambers required to be filled halfways, to the bottom of the opening. Those chambers that had lines under their names required a few drops of paraffin oil to be added to mimic anaerobic conditions and keep the byproducts of chemical reactions (e.g. gases) inside the chamber. The chambers that had bowl-like drawings under their names were supposed to be filled to the maximum capacity. After the samples were inside the chambers, the strips were placed inside a plastic container. We sprayed the bottom of the plastic container with 3 mL sterile deionized water in order to increase humidity in the chamber. The strips were incubated in the dark at room temperature (~30°C) for 18-24 h. TDA, IND and VP chambers required reagents to be added after the incubation. We added a drop of TDA reagent to the TDA chamber, a drop of James reagent to the IND chamber and a drop of each of the two VP reagent to the VP chamber. We left the reagents to react for 10 minutes and performed an additional oxidase test by transferring a generous bacterial sample on filter paper previously moistened with oxidase reagent. We transferred the colonies with a sterile pipette tip because a normal nickel loop would have given a false positive. If the sample turned blue within the next minute, the sample is oxidase positive. If its colour stayed the same, it's oxidase negative. We analyzed the results with the help of the manual and then wrote it down on API20E result sheets. We sent the sheets to a clinical biochemistry lab for further identification of the species.

FIGURE 4: API20E STRIPS AFTER ADDING THE SAMPLE TO ANALYZE

Isolation of bacterial genomic DNA

The genomic DNA of 12 bacteria was isolated using the Qiagen UltraClean Microbial DNA isolation kit, and the quality was assessed using a nanodrop spectrophotometer (ThermoFisher) and a Qubit device from Invitrogen. Additionally, a DNA gel was loaded with the samples in order to assess the integrity of the DNA.

Sequencing DNA library preparation

The aim of library preparation is to prepare the DNA into a form that is compatible with the sequencing (in this experiment nanopore sequencing) used. For library preparation we used Rapid Barcoding Sequencing (SQK-RBK004) kit as it provided some important features to our experiment: allowed sample multiplexing, was PCR -free, and had short preparation time.

One of the most important features of the chosen kit was that it included DNA barcoding which is very important for sample multiplexing. Multiplexing is a useful technique as it saves cost and labor because the physical sample that is actually run through the sequencer can be pulled together with other samples. Therefore, barcodes are needed to identify the source of an individual sample. We achieved this by transposase-based library preparation.

Transposons are genetic elements that can move within or between genomes by either replicative or 'cut-and-paste' mechanisms mediated by an enzyme called transposase. This enzyme recognizes the inverted repeats at the ends of the transposon and also recognizes the target sequence, in which it makes a double-strand break and inserts the transposon. In our case, barcoded tags were inserted (Fig. 5). During library preparation adapters are also ligated to genomic DNA fragments. These adapters facilitate strand capture and loading of a processive enzyme at the 5′-end of one strand. The enzyme is required to ensure unidirectional single-nucleotide displacement along the strand at a millisecond time scale.

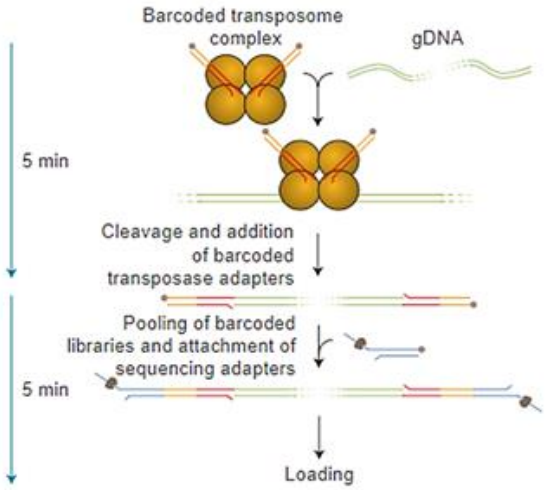


FIGURE 5: GDNA LIBRARY PREPARATION

In our project we used DNA from 12 different strains of bacteria to prepare library for sequencing, to prepare it we followed Rapid Barcoding Sequencing (SQK-RBK004) protocol (version: RBK_9054_v2_revB_23Jan2018) provided by Oxford Nanopore Technologies (ONT).

Genome sequencing with nanopore technology

For sequencing we used MinION by Oxford Nanopore Technologies (ONT) with the flow cell FLO-MIN106. Flow cells bear up to 2048 individually addressable nanopores that can be controlled in 4 groups of 512 which we checked before sequencing. MinKNOW software (by Oxford Nanopore Technologies (ONT)) controls the nanopore sequencing device as well as collects sequencing data in real time. After flow cell testing 1543 pores out of 2048 were detected (504, 468, 376, and 195 per group)

The working principle of nanopore sequencer is simple: DNA is ratcheted through the nanopore base-by-base. As the DNA passes through the pore, the sensor detects changes in ionic current caused by differences in the shifting nucleotide sequences occupying the pore.

To prime and load DNA on the sequencer we followed Rapid Barcoding Sequencing (SQK-RBK004) protocol (version: RBK_9054_v2_revB_23Jan2018) provided by Oxford Nanopore Technologies (ONT).

Just before the sequencing we had to set certain settings. This included selecting the device, indicating the flow cell as well as library preparation kit used. Additionally, we needed to choose whether to base-call data live on the instrument. The group decided not to. We had to select run

options (run time and starting voltage) as well. We chose not to change default values which are 48 hours and -180 mV. Finally, it is required to indicate the desired output. In this experiment, we decided that just .fastq should be produced.

Software tooling for bioinformatics analysis

Albacore (v2.1.10.): Albacore is Oxford Nanopore Technologies' official command-line base caller. To successfully use it we needed to indicate some parameters like flow cell type, library preparation kit, and whether barcoding was used.

Bandage (v1): software that was used for genome assemblies visualization.

FastQC (v0.11.7) – visualizes Illumina reads and provides some additional information about the reads such as quality scores, GC content, sequence length distribution.

NanoPlot (v1.13.0) – visualization of nanopore reads quality metrics.

PoreChop (v0.2.3) – removes the barcodes from the nanopore reads.

Prokka (v1.13) - whole genome annotation is the process of identifying features of interest in a set of genomic DNA sequences, and labelling them with useful information. Prokka is a software tool to annotate bacterial, archaeal and viral genomes quickly and produce standards-compliant output files.

Trimmomatic (v0.38) - performs a variety of useful trimming tasks for illumina paired-end and single ended sequencing data.

Unicycler (v0.4.4) - as DNA sequencing technology cannot read the whole genome at once, it does that in smaller pieces which size depends on the sequencing technology used. Therefore, these shorter fragments, called reads, is the result of DNA sequencing rather than the full genome. Sequence aligning includes aligning and merging those reads in order to reconstruct the original sequence. That is the reason why we used Unicycler. Unicycler is an assembly pipeline for bacterial genomes. Unicycler can: 1, assemble Illumina -only read sets, 2. long-read-only sets (Nanopore) or 3. both Illumina reads and long reads conducting a hybrid assembly. Illumina has the accuracy of around 99%. However, with Illumina there are two problems: 1. If in the DNA there is a homopolymer region, it's difficult for Illumina to distinguish between separate bases. 2. Illumina sequencing requires PCR reaction. However, different reads of DNA may have very different GC content, meaning that in some reads where GC content is exceptionally high PCR reaction may fail and therefore sequencing data may be misleading. On the other hand, nanopore sequencer overall accuracy compared to Illumina is way lower, only about 92 %. Therefore, in this project Unicycler was given both long and short reads as this way the we got best possible assemblies.

Double agar overlay assay

We mixed the samples of bacteria and phages in 6 mL of soft agar in a manner shown in the table 1.

Sample name	Volume of O/N (µL)	Name of Phage sample	Volume Phage sample (µL)
Negative control 1	NA	NA	NA
Negative control 2	200	NA	NA
Positive control (E. coli)	200	S30	100
MacConkey A	200	S6	100

MacConkey B	200	S6	100
MacConkey C	200	S6	100
E. coli (lab strain)	200	S31	100
E. coli (lab strain)	200	S32	100
E. coli (lab strain)	200	S33	100
E. coli (lab strain)	200	S34	100
E. coli (lab strain)	200	S35	100
E. coli (lab strain)	200	S36	100

Polymerase Chain Reaction and DNA Restriction

First of all, using PCR reaction we amplified lambda DNA with 5 pairs of different primers. PCR is a biochemical process capable of amplifying a single DNA molecule into millions of copies in a short time. Amplification is achieved by a series of three steps: (1) denaturation, in which double-stranded DNA templates are heated to separate the strands; (2) annealing, in which short DNA molecules called primers bind to flanking regions of the target DNA; and (3) extension, in which DNA polymerase extends the 3′ end of each primer along the template strands. These steps are repeated ("cycled") to exponentially produce exact copies of the target DNA).

To prepare 5 different tubes of DNA with different primers for PCR, we mixed 12.5µL of EmeraldAmp MAX PCR Master Mix (2X Premix), 0.5 µL of each forward and reverse primer, 0.5µL of lambda DNA and 11 µL of nuclease – free water. The device used for PCR was Lab-Line Programmable Thermal Blok II. Our program had 3 cycles with a hot start which is essential for the activation of polymerase (we used TaqGold polymerase), The program was: hostart at 98 °C for 2 minutes, denaturation step - 98 °C for 10 s, annealing step - 50 °C for 30 s and extension step - 72 °C for 1 min, with the total number of cycles being 30.

After the PCR reaction, we added 1.5µL of EcoRI restriction endonuclease to each of the tubes with the amplified DNA and incubated at 37°C for 1 hour so the enzyme would have enough time to cut the DNA. After that, we inactivated EcoRI by incubation at 65°C for 20 minutes.

To see if the reaction took place and EcoRI cut the DNA we ran DNA gel electrophoresis. To prepare the gel we mixed 30ml of TAE buffer with 0.21 g of agarose and heated in the microwave. While the gel was still hot, we added 3µL of SybrGreen to it. (SybrGreen is a cyanine dye that binds to DNA and stains it so later on the DNA running through gel could be detected). Before turning on the electric current in the first lane we added 10 µL digested lambda DNA, which was supposed to act as a ladder. To the next 8 lanes we added 10 µL of DNA in 4 pairs as follows: 1) amplified lambda DNA without EcoRI,2) amplified lambda DNA with EcoRI.

Results and Discussion

Environmental samples

TABLE 1: ENVIRONMENTAL SAMPLES DATABASE WITH LABELS, LOCATIONS, PH, TYPE OF SAMPLES

#	collect by	body of water	date	time	volume (mL)	pH	LB agar CFU - dilution				BAP	MAP
							0	1	2	3		
1	Vito	small puddle	July 30, 2018	09:10	30	9	null	null	null	null	null	null

2	Cedric	stream Veličanka	July 30, 2018	09:35	32	9	null	null	null	null	null	null
3	Woody	stream Veličanka	July 30, 2018	09:35	49	9.5	160	16	3	0	2	null
4	Vito	stream Veličanka	July 30, 2018	09:35	47	9.5	580	58	5	0	null	null
5	Vito	river Orljava	July 30, 2018	09:50	50	9.5	500	50	5	0	4	null
6	Irma	sewer	July 30, 2018	10:00	13	8.5	540	54	3	0	17	24
7	Karolina	Large puddle	July 30, 2018	10:05	20	9	40	4	1	0	1	3

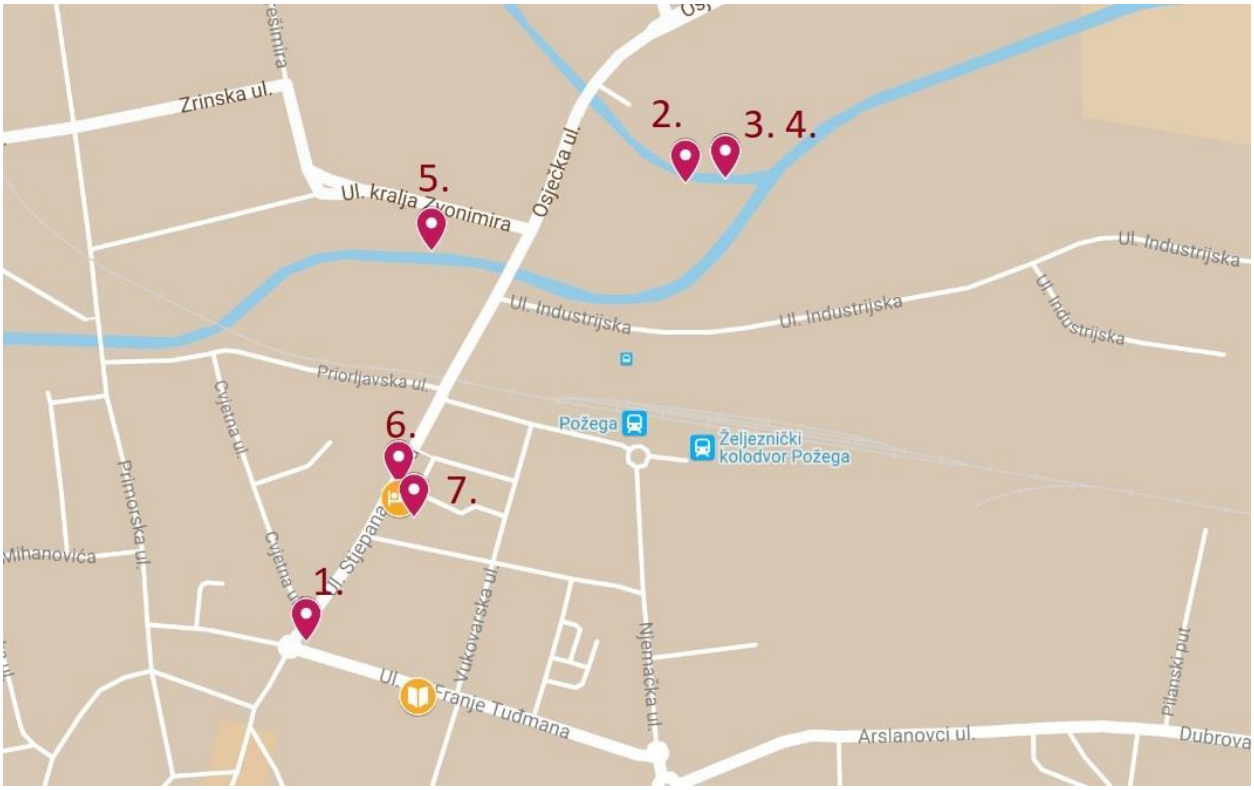


FIGURE 6: LOCALIZATION OF SAMPLES (POZEGA, CROATIA)

Identification of bacteria with biochemical tests

TABLE 2: RESULTS OF THE IDENTIFICATION OF BACTERIA WITH THE API20E STRIPS.

Name of the sample and plate it originated from	Identification: 1st possibility, percentage of accuracy	Identification: 1st possibility, percentage of accuracy
E. coli test sample	Klebsiella oxytoca, 96.0%	Raoultella terrigena, 2.4%
W.V.R. side 2; MacConkey agar	/	/
KG MAC 1 (VP NEG), MacConkey agar	Enterobacter cloacae, 94.3%	Enterobacter amnigenus 2, 2.7%
KG MAC 1 (VP POS), MacConkey agar	Enterobacter cloacae, 95.1%	Enterobacter sakazakii, 3.0%

V.M. SD1, MacConkey agar	Aeromonas hydrophila/caviae/sobria 2, 50.1%	Aeromonas hydrophila, caviae, sobria 1, 49.7%

The identification of the *E. coli* test sample hadn't resulted in E. coli, due to possible contamination or the inaccurate analysis of the results. The W.V.R. side 2 sample was deemed unacceptable by the person who identified the bacteria, possibly due to contamination or wrongful analysis of the strip results. There are two KG MAC 1 samples, the VP negative and positive ones, due to the fact that the colour of the VP chamber on the API20E strip had been somewhere between the positive and the negative results, so we wanted to be thorough. The identified bacterial species all belong to the *Enterobacteriaceae* family. *Enterobacter cloacae* is a bacteria common in the human gut flora, which is to be expected because the samples were deemed to have feces or other forms of organic matter. *Enterobacter amnigenus* 2, *Enterobacter sakazakii*, *Aeromonas hydrophila/caviae/sobria* 1 and 2 are all pathogenic bacterial species.

Gel electrophoresis of gDNA samples

The 12 samples passed quality control after gDNA isolation, with all values of spectrophotometry being above 1.8 as per the nanodrop. They also displayed limited amout of shearing as shown on figure 7.

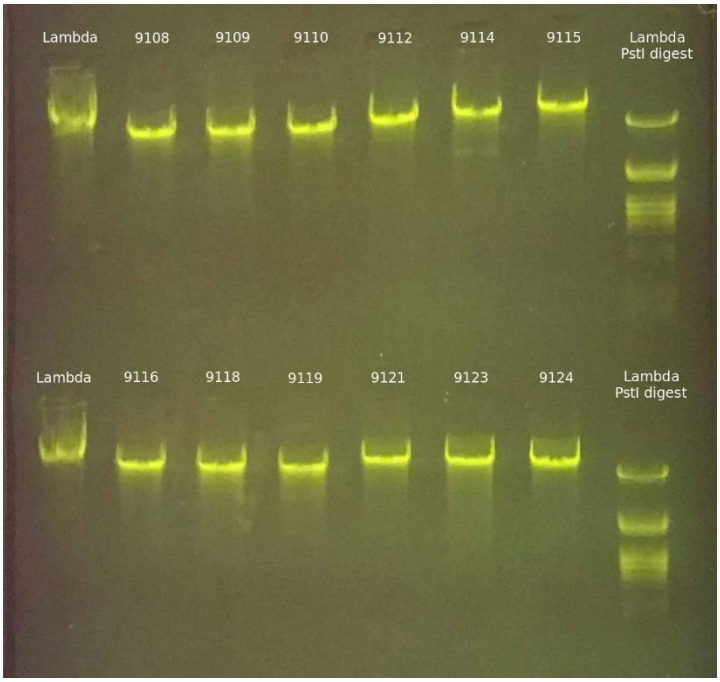


FIGURE 7: TWELVE SAMPLES OF THE UNKNOWN POPULATION OF BACTERIA (LADDER LEFT: LAMBDA DNA, LADDR RIGHT: LAMBDA DNA DIGESTED WITH PSTI)

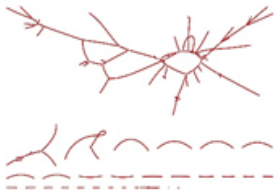
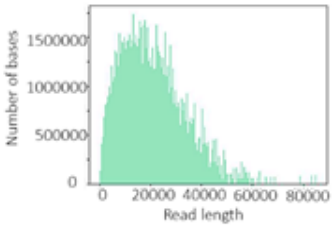
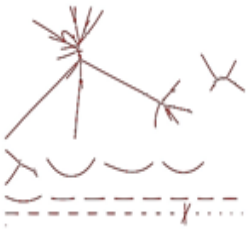
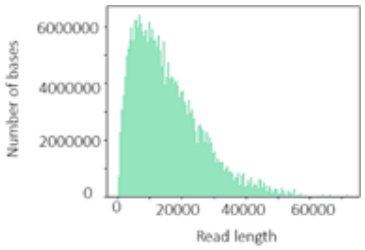
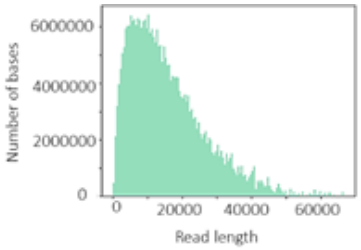
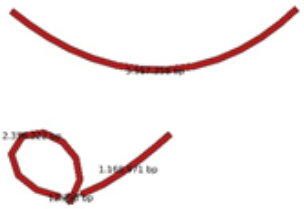
Identification of bacteria with nanopore sequencing

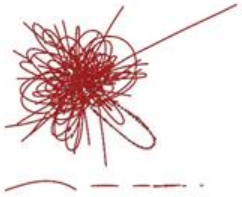
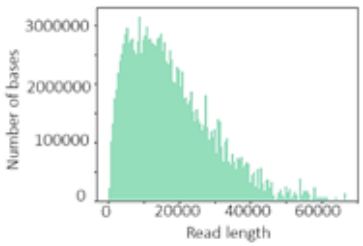
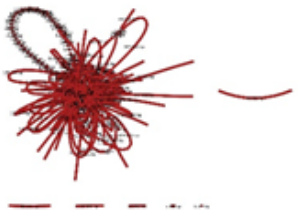
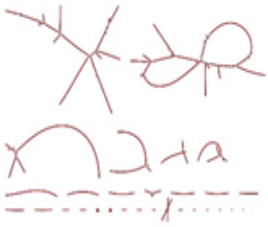
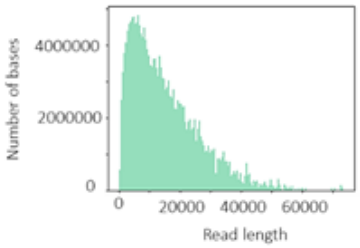
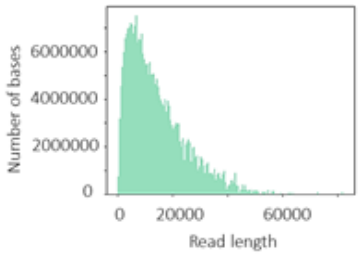
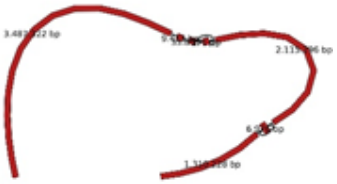
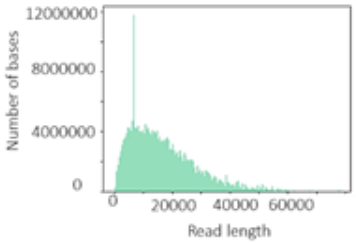
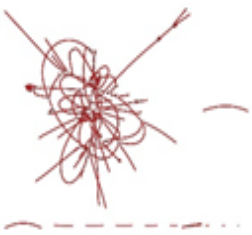
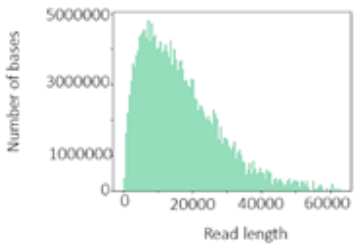
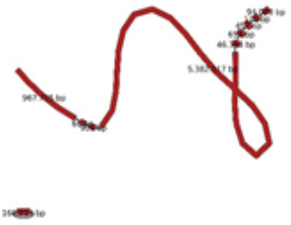
Nanoplot software provided us with multiple files, concerning reads quality, length, number of reads, total bases number. With ONT MinION sequencer we generated 269735 reads, totaling 0.945Gbp after base – calling (Table 3).

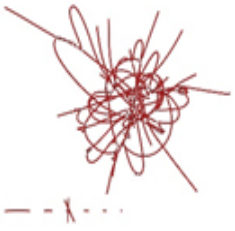
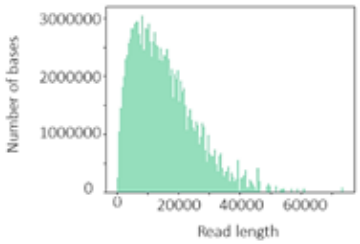
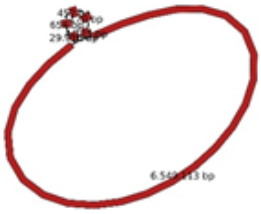
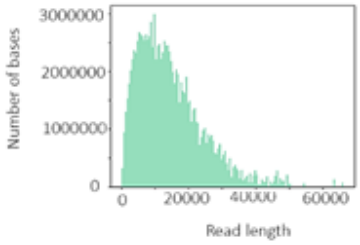
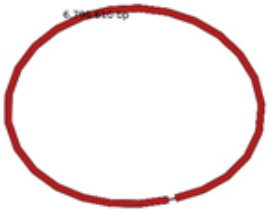
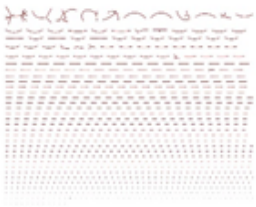
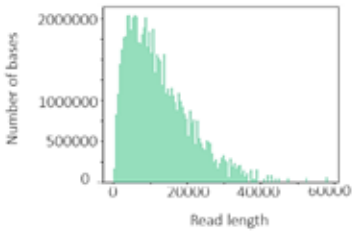
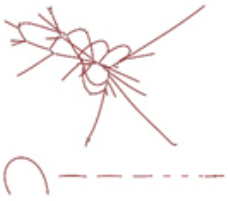
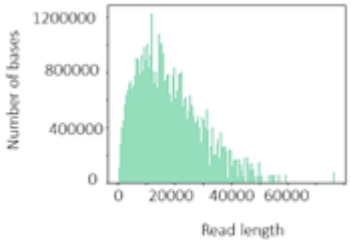
Having completed genome annotation, we can see can see while that Illumina - only assembly graphs contain multiple dead ends, with hybrid assemblies we managed to get complete circular chromosome sequence for 7 samples (Table 3). Therefore, it is obvious that combining Illumina and ONT MinION sequencer reads is beneficial to accurately recreate the genome and get the best possible assembly.

The next step we did was genome annotation which provided us with valuable information such as the name of the organism DNA came from, proteins and their coding places in the DNA, number of repeat regions, tmRNA, tRNA. We found that the DNA of 12 strains sequenced came from Pseudomonas aeruginosa. Additionally, some of the bacteria strains sequenced have repeat regions (strains: 9108, 9109, 9112, 9114, 9121, 9123, 9124) which are highly related to bacteria resistance to antibiotics.

TABLE 3: RESULTS OF THE ASSEMBLY OF THE ILLUMINA DATASET (PREVIOUSLY COLLECTED) AND THE NANOPORE SEQUENCING DATA.

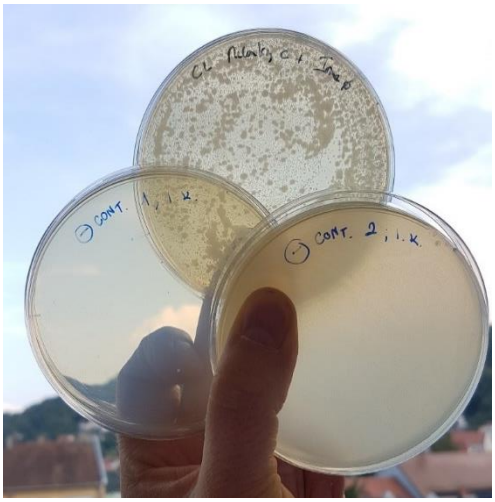
Sample	Illumina-only assembly graph	ONT reads	Hybrid assembly graph
9108			
9109			
9110			

9112			
9114			
9115			
9116			
9118			

9119			
9121			
9123			
9124			

Use of the double agar overlay to detect presence of phages in samples

After 24 h of incubation, we examined the double agar overlay assay plates in the hopes of finding viral plaques, see-through zones in the medium that indicate the presence of lysis by phages. The negative controls appeared not to be contaminated. The positive control had resulted in some viral plaques. The plate with the most plaques was the one which contained the MacConkey 3 bacterial and S6 (sewer) phage samples. The plaques were spread all over the plate, but some bacterial growth was still noticed, which would mean that there was more than one type or species of bacteria present, that the concentration of phages was insufficient or that the medium was contaminated. Both the bacterial



and the phage samples originated from the sewer samples, so a deduction could be made that the phages specific to certain bacteria do indeed appear in the same environment as the bacteria.

DNA Restriction of Lambda phage with EcoRI

From gel electrophoresis (Fig. 8), we can see that in the first lane digested lambda DNA segments of various length have travelled different distances, thus allowing us to approximately determine the length of the DNA in the following lanes. As they go in pairs as follows: amplified lambda DNA without EcoRI, lambda DNA with EcoRI, we can see that in each sample, our restriction enzyme EcoRI cut the DNA at specific sites, resulting in two distinctive lines on the gel, while there is only one in the DNA without an enzyme. (Fig 8).

Therefore, with this experiment we can demonstrate that R-M efficiently works in the bacteria as one of the protections against phages ways. Additionally, these results are important because based on the segments size or cutting site, we can validate that the DNA we had comes from lambda virus DNA. This technique can easily be used when searching for new viruses or for identifying the phage type.

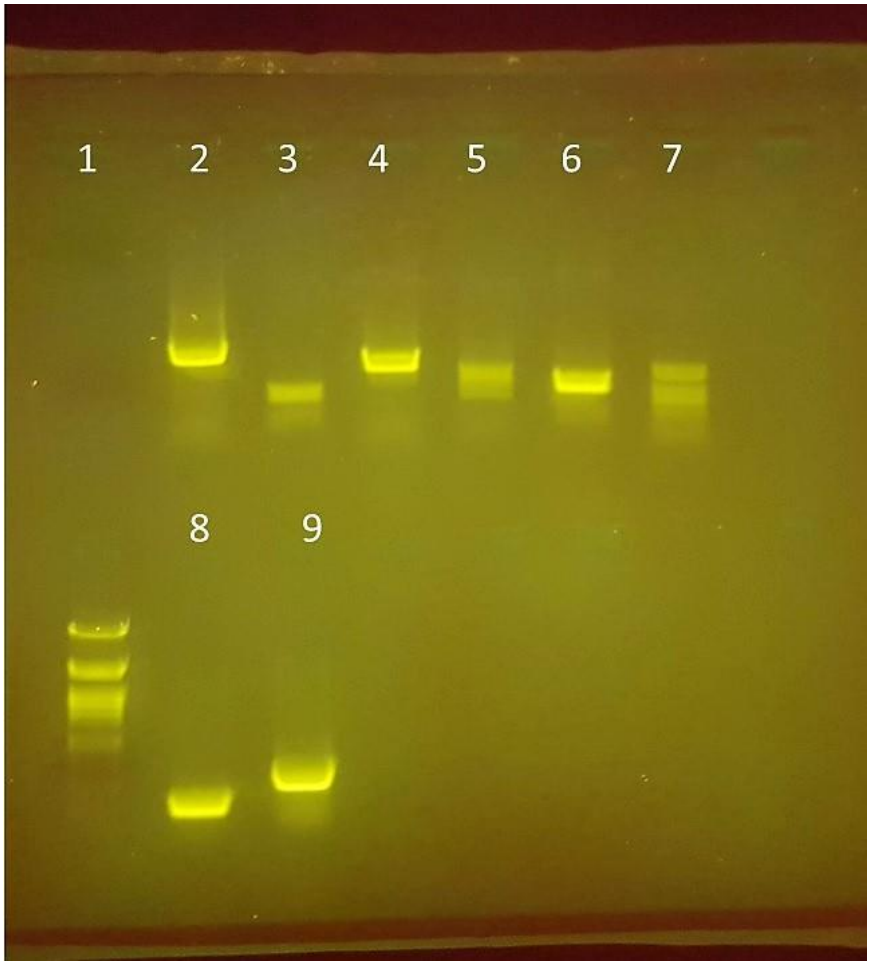


FIGURE 8: GEL ALLOWING THE VISUALIZATION OF PCR PRODUCTS (LANES 2, 4, 6, 8, AND 9) AS WELL AS THEIR RESPECTIVE RESTRICTION WITH ECORI (LANES 3, 5, AND 7 – NO RESTRICTION SHOWN FOR 8 AND 9). LANE 1 WAS LEFT EMPTY ON THE FIRST ROW, AND IS LAMBDA DNA/PSTI DIGESTED IN THE SECOND ROW

Phylogenetic placement of the isolates

Using both the biochemical characterization (API20E) and sequencing methods, we obtained information about the species that we isolated. These species have been placed on a phylogenetic tree using their 16S rDNA information (Fig 9). All species studied belonged to the taxonomy class of *Gammaproteobacteria*.

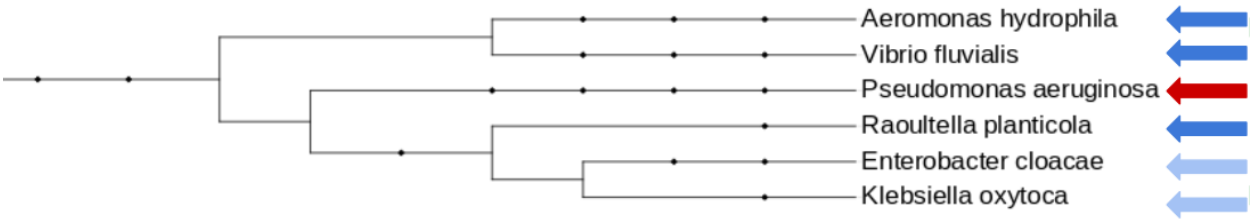


FIGURE 9: PHYLOGENY OF OUR SPECIES. IN RED, THE 12 ISOLATES FROM THE P. AERUGINOSA SPECIES CHARACTERIZED THROUGH SEQUENCING, IN BLUE ARE THE RESULTS FROM THE BIOCHEMICAL TESTS.

Conclusion and Perspectives

While biochemical tests for identification of bacteria provide some considerable advantages such as cost-effectiveness and are relatively easy to perform. At the same time, there are some drawbacks which include potential misidentification of bacteria (as multiple bacteria may share similar phenotypic or metabolic traits) as well as lesser discriminatory power. Additionally, to identify bacterial species multiple biochemical tests may be needed as well as bacteria need to be grown and pure colony need to be obtained.

However, with the advancement of technology, next–generation sequencing is becoming more and more available for usage in research and public health laboratories to identify bacteria. This approach brings some major advantages such as high accuracy compared to biochemical tests as well as it provides some additional information, for example about antibiotic resistance genes and finally is relatively quick as from a few reads bacterial species can already be identified. It is an invaluable tool for metagenomics when genetic information can be analyzed directly from environmental samples which is not possible using some biochemical tests such as API 20E strips.

One of the main reasons why accurate bacteria identification is needed is for medicine to provide the right treatment for the patients. However, because of heavy use of antibiotics, some bacteria that are capable of causing serious disease are becoming resistant to most commonly available antibiotics. Because of the rapid emergence of antibiotic resistant bacteria, we need to look for new ways to fight bacterial infection. One alternative is phage therapy. To put it simply, phage therapy is the application of bacteria–specific viruses (bacteriophages) to combat bacteria and treat diseases. This approach provides several major advantages, for example some of the antibiotics are bacteriostatic meaning that they stop bacteria from reproducing, while not necessarily killing them therefore favoring bacterial evolution towards resistance. Additional advantages include: applying phages to antibiotic resistant bacteria, low toxicity, easy discovery as well as phages and bacteria coevolution, meaning that while bacteria evolve and develop various ways to defend themselves from viruses, phages co-evolve as well and find new ways how to infect bacteria.

Acknowledgments

We would like to thank Dora Grbavac and Nikolina Šoštarić for the organization of the summer school of science. We also thank the Lab of Gene Technology from KU Leuven and the Faculty of Science of the the university of Zagreb for providing some of the equipment required for the project. We also thank all the participant of the summer school of science 2018 for sharing their knowledge and passion for science with us!