Master Thesis:

On the Prototrophy of *E. coli*

Give or take

Student: Hilkes T.F.J.

891214-338-040

Supervisors: Duur Aanen

Fons Debets

Marjon de Vos

Index

1. Introduction

1.1. History of E. coli

In 1884 Theodor Escherich isolated and identified a bacterium he called *Bakterium coli commune* from the stool samples of infants (Escherich 1885). Escherich was a paediatrician researching diarrhoea in young children. The contemporary mindset held the idea that diarhoea only had chemical causes, but Escherich was convinced of the possibility of infectious causes, which he had observed before with cholera in Italy (Friedmann 2006). Over the course of his works, Escherich described and named multiple bacteria, but special attention always went to coli. Escherich described coli as a slender slightly bent rod-shaped bacterium. He also noted differences in cultures based on the diets of the children.

In 1897 Shiga Kiyoshi isolated a toxin producing bacterium that also caused diarrhoea (Shiga 1898), which was later named as *Shigella dysenteriae*. With the genus (*Shigella*) and the toxin (Shiga) being named after him.

The reason why *coli* was so focussed upon at the start of associating microbes with intestinal diseases, was its seeming over abundant presence in infant faecal matter and how easy it was to grow in isolation on medium. Combined these meant it was of great interest to bacteriologists and was studied extensively in the years following Escherich's discovery.

In 1919 Castellani and Chalmers officially renamed *Bakterium coli commune* to *Escherichia coli*, after its discoverer and formally integrating it into the binominal nomenclature system.

In 1922 an *E. coli* strain was isolated from a diphtheria patient, designated K-12. K-12 was used in student lab work and thus cultivated on medium till the early 40's, when Tatum used it for his research (Tatum and Lederberg 1947) and from there became a standard in molecular biology. It was used as a model species to unravel the workings of the cell. *E. coli* being easy to grow and manipulate was also taken advantage of by biochemists, by having *E. coli* produce all kinds of substances by inserting the genes from other organisms, that can produce the substance, but are difficult to grow themselves, into *E. coli*.

Recent advances in sequencing techniques have shown that the intestines and excrement hold hundred times more different bacteria than what was isolated and grown in the past (Weinstock 2012). Where it was once believed *E. coli* was abundantly present and therefore of relevance, now it had turned out it was just one of the few that could have been identified with the limited techniques of the past. While replacements for *E. coli* are being found, it will take quite some years before these become the standard. Not just *E. coli*'s historical presence, but also its extreme ease of use, should allow it to remain a relevant standard and tool for quite some more time.

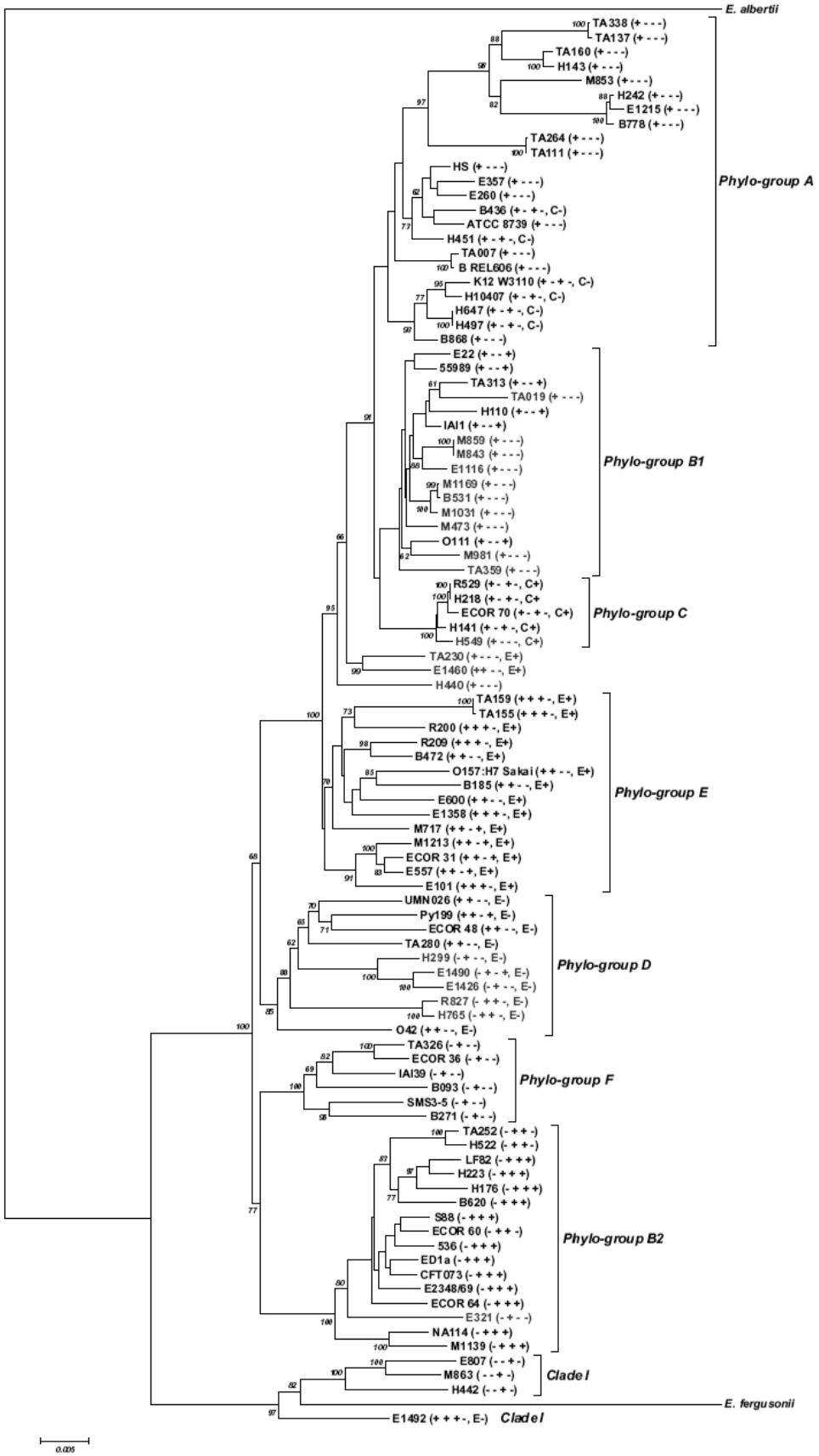
1.2. Definition of E. coli

E. coli is defined as a cylindrical rod-shaped bacterium, about 0.5 by 2.0 μm, gram negative, aerobic and facultative anaerobic, some capable of movement by flagellum (Ewing 1986). Sequencing suggests that only 20% of genes are shared between all *E. coli* strains (Lukjancenko et al. 2010), compared to the often mentioned 99.5% or 99.9% in humans. Traditionally *E. coli* and the four Shigella species are considered separate species and families, that under old nomenclature were considered part of the same "tribe" (Ewing 1986), but recent research has put the Shigella strains in the Escherichia genus (Zuo et al. 2013), while some researchers claim the four are just *E. coli* variations (Lan and Reeves 2002). Shigella strains do not usually have the flagella associated with some *E. coli*, but do have the genes and the capability to express these under very specific circumstances. Originally only Shigella strains were known/named

that produced Shigella toxins, but recent identified strains producing Shigella have been classified as *E. coli*.

E. coli is divided into four main phylogenetic lines A, B1, B2, and D (Clermont et al. 2000), four minor phylogenetic lines C, E, F and cryptic clade 1 (Clermont et al. 2013) and four more cryptic clade groups II-V that genetically resemble *E. coli* but have not been fully identified yet (Clermont et al. 2013). Most virulent strains of *E. coli* are either subgroup B2 or D, while e.g. K-12 is from subgroup A. The picture included shows strains ordered into the phylogenetic group they are a part of and how they are genetically related over the course of evolution. It shows the newly identified minor phylogenetic groups as fitting in-between the main *E. coli* lines.

Another way to group and identify *E. coli* is through identification of certain antigens in the membranes and the flagella, called serotyping. Serotypes are strongly connected to specific genotypes and are both used in research, but mostly in the medical field to quickly identify whether an isolated strain is of a type known to cause illness. The O and K antigens from the membrane and H antigens from the flagella are used for this and as such each *E. coli* strain can be named/numbered based in the version of the antigen it produces for each of the three. Ishii et al., 2007, strongly established that serotype antigens and phylogenetic groups are connected and serotype is thus also a good early indication of the phylogenetic group a strain belongs to.



Picture 1: Phylogenetic Tree of *E. coli* lines (Cleremont et al. 2013)

Historically, the main usage of *E. coli*, be it for research or production, has been based on it being easier to grow than other bacteria, on account of it being prototrophic and facultatively anaerobic. This means it can create all it needs to survive from the most basic building blocks of life, both in the presence and absence of oxygen. It is able to grow properly at temperatures ranging from just below 20°C to slightly over 40°C, with an optimum around 37°C. *E. coli* were always selected in the past on being able to grow well on whatever medium was available, from a limited range of sources, like strains surviving in human gut material. The limited section and growth methods used, might have limited the initial view on *E. coli* and in turn might have strongly influenced the dogma that *E. coli* is always prototrophic. The wide range of temperatures *E. coli* can be active at, also leads to the question whether its prototrophy might be temperature dependant.

The first strains of *E. coli* were isolated from human excrements (Escherich 1885), leading to the early and long held belief that the natural, and only, environment for *E. coli* is the gut and intestinal track of many animals, including humans. A large part of the strains carried by a host, are acquired at birth from the mother (Bettelheim et al. 1974) and the presence of *E. coli* even has an effect on the development of the host. The generation time of *E. coli* in the gut has been determined to be about 40 hours, which is more than the 20 minutes observed in a laboratory setting on optimal medium (Savagau 1983). Traditionally the gut is considered rich in all sorts of nutrients, suggesting there is no inherent advantage to being a prototrophic lifeform in that environment. Competition with the many species present in the microbiome and the actual availability of certain resources might give an advantage to being prototrophic. Meaning that if a certain required nutrient is actually scarce instead of abundant and there is competition over it, the organism that can produce it itself through more readily available substances, will not have to waste energy participation in the competitive struggle and will have energy left over it can spend on reproduction or otherwise.

Morris et al., 2012, proposed a theory they call the “Black Queen Hypothesis”, named after and tangentially related to the “Red Queen Hypothesis” (Valen van 1973) The original Red Queen Hypothesis proposes that evolution is always to facilitate reproduction alongside other species that are also evolving and thus has to be population wide. The Black Queen Hypothesis argues that in a population of microorganisms, it can be advantageous to the population if only small subsets maintain the ability to be able to produce certain substances, while the rest of the population loses the ability. This would result in a division of energy and labour that could be advantageous to the population as a whole. The theory postulates the possibility of a division across the population, making it a mutual exchange of essential nutrients, but also the possibility of a small sub-population maintaining all abilities and becoming the suppliers for the entire population.

1.3 Ecology of E. coli

In gut ecology a distinction is made between resident and transient strains. Research like Caugant et al., 1981, and Nowrouzian et al., 2005, show clear distinctions between strains that show up in samples from the same source only for a few days at a time and strains that are present for more than three weeks during sample testing. Statistical analysis by the researchers indicated that B2 strains are more likely to be or become resident strains than strains from the other phylogenetic groups, independently from whether strains were virulent or not.

Long has the belief been held that all *E. coli* present in the environment, outside of the bodies of other organisms, is there because of contamination by faecal matter. One of the tests used to test for faecal contamination was even to test for the presence of *E. coli*. Recent research has shown that there are strains of *E. coli*, that are completely adapted to life outside of a host

organism and have lived out in the environment for many generations, (Ishii and Sadowsky 2008). Still, most research has been done on *E. coli* acquired from animal faecal and descendants of such strains, mainly from humans. Most if not all of our knowledge on "natural" *E. coli* is based on intestinal *E. coli* dependant on a host. Like mentioned before, most strains used in the lab are either K-12 or descended from K-12, placing most *E. coli* that have been used for research in just one of the four subgroups of *E. coli*.

1.4 Prototrophy of E. coli

In short, there seem to be a lot of assumptions and paradigms surrounding *E. coli* that have been maintained since its discovery. Ideas that in the last few years have slowly been changing. Combined with the other assumptions mentioned before, this does bring into doubt the central assumption about *E. coli*. *E. coli* is assumed to only occur prototrophic in nature, but is mainly known from a source where this is most likely not an advantage. At the same time, we learned that that same environment holds more microbial life, than we always believed, discovered only because of modern sequencing techniques. This thesis seeks to actively question this assumption and find evidence either in favour or against *E. coli* being naturally prototrophic.

To try and answer that question , one would take a strain of *E. coli* that is known to be prototrophic and well-studied and grow it for a long period of time, across thousands of generations in a situation without evolutionary pressure and test samples from set intervals by growing them under nutrient poor conditions. Over generations, one would expect the strain to lose the ability to produce certain substances that are abundantly present in the main growth medium. The mechanism by how such a mutation could occur, has two prevailing theories for it. The first, as proposed by Medawar 1952, mutation accumulation states that mutations that have no detrimental effects on the organism under the current circumstances persist and accumulate over generations and only become lethal when those circumstances change to a situation where they would be relevant. The second, as proposed by Williams 1957, antagonistic pleiotropy states that when genes encode for multiple effects, the gene can survive generations if at least one of the beneficial effects outweighs the negatives of the other effects, or the negative effects occur in a later stage of life, after reproduction. In short, the question is whether the auxotrophy is a result of individual mutations to a gene over multiple generations that eventually change the functioning of a gene or a disadvantageous mutation becoming noticeable after the more positive effect it was connected to becoming inactive or changing due to a further mutation.

2. Hypothesis

Prototrophy will disappear over time in nutritionally rich environments.

3. Research Questions

1. Are all natural isolates of *E. coli* prototrophic?
2. How stable is the prototrophy of prototrophic *E. coli* samples on rich medium?
3. What is the mechanism by which prototrophy is lost from *E. coli*, when it is lost?

4. Materials and Methods

4.1 E. coli Evolution Experiment

For the evolution experiment, Ara+ strains from the Lenski experiment were used (Lenski et al. 1991). Lenski grew 6 Ara- and 6 Ara+ strains on minimal medium over a long time. The Ara- strain is a mutant they made that is auxotrophic for Arabinose, the Ara+ strain was a spontaneous back mutation they found.

Two 15 ml tubes of LB were inoculated with Ara+ and Ara- from glycerol back-ups, used by the university to make the stocks for course lab experiments based on the Lenski experiments. These grew in a rotating incubator overnight at 37˚C and were used the next day to inoculate TA-Arabinose plates. These were again grown overnight in a 37˚C oven. To start each culture a single colony was picked from each plate and put into 10 ml of LB in a 30 ml bottle. These were allowed to grow for 24 hours at 37˚C in a rotating incubator. Both inoculations were done aseptically.

After 24 hours, 0.1 ml of the Ara+ culture was added to 6 separate sterile bottles containing 9.9 ml of LB. These were grown at 37˚C in a rotating incubator. While shaking, these bottles were never tightly closed, allowing oxygen to move into the bottle while shaking. From this point onward, after a day of growth, 0.1 ml from each bottle was added to a fresh sterile bottle containing 9.9 ml of LB and allowed to grow overnight again. These transfers were performed in sterile inoculation rooms next to an open flame.

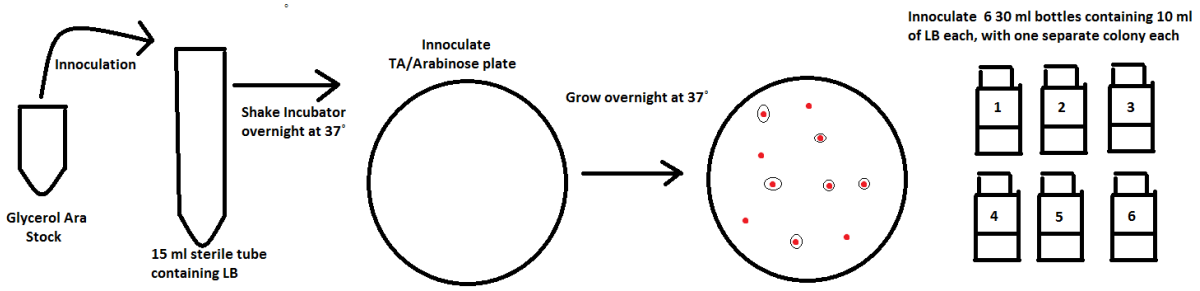


Fig 4.1: Inoculation of the six lines from the Ara Stock

After four daily transfers, back-ups of the six lines were made in glycerol and the bottles themselves stored in the refrigerator, for the lines to be transferred again on the next cycle starting day. This pattern of 4 growth days and 2-3 storage days was kept for 13 weeks.

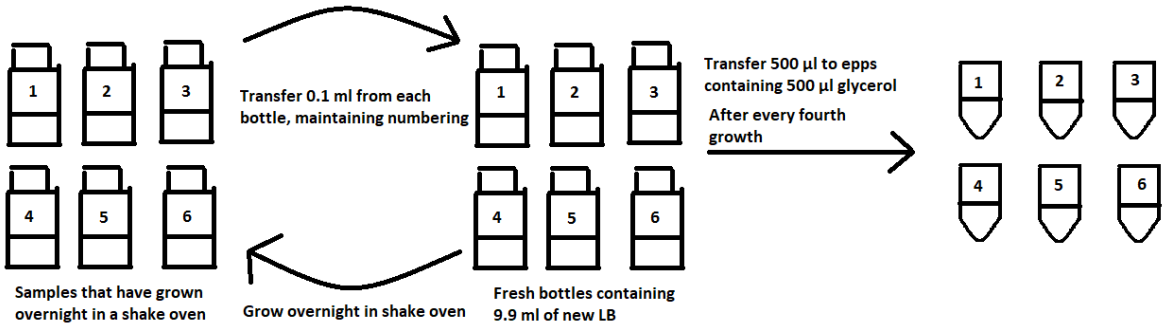


Fig 4.2: 4 Day Growth Pattern Set-up

Using the glycerol back-ups of week 13, the six lines were put through the same protocol as before, for 1 week, once allowed to grow at 28˚C daily and once allowed to grow at 42˚C daily.

4.2 Screening for auxotrophic mutants

TA/Arabinose Plates were used to distinguish between the well documented Ara- and Ara+ lines. Ara- colonies appear a dark red and Ara+ white on these plates.

Chromagar (Becton Dickinson GmbH. 2011.) plates were used to test the purity of cultures before making glycerol stocks. *E. coli* appears dark pink on these plates. Any other colour colonies appearing implies contamination by other bacteria or other organisms. A dilution of 10^{-3} of the cultures was utilized to get proper, visibly separate results on the plates.

To test auxotrophy a nutrient rich medium, LB and a nutrient poor medium, Davis Broth, were used. Liquid Davis Broth and Davis Agar plates were made according to Lenski et al (1991) adjustments of the Davis recipe from Carlton and Brown, 1981. For this a dilution of an end-of-week stock culture was plated on an LB plate and allowed to grow overnight at 37˚C, dilutions were selected and tested on the largest density of visible, but clearly separate number of colonies. These would then be stamped over by velvet stamping, first to a Davis plate and then to another LB plate and grown again at 37˚C until colonies are properly visible again. A colony that appears on both LB plates, but not the Davis plate, is possibly auxotrophic and therefore interesting. These colonies would be picked and grown in sterile bottles with 10ml LB for later testing.

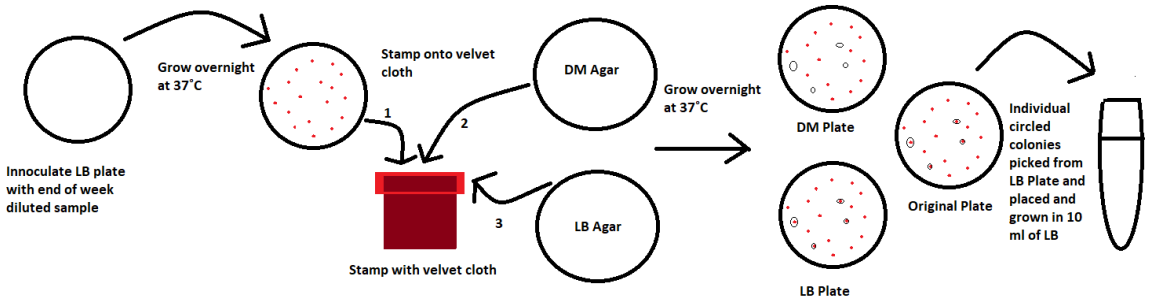


Fig 4.3: Finding and Growing Auxotrophic Colonies

4.3 Identification of auxotrophic mutants

Utilizing "sticks with cotton", these cultures would be placed on LB plates in a 4 by 4 grid, each culture in duplo, in a sterile room, near a flame. These were grown at 37˚C and used as master plates for stamping onto plates for auxotrophy testing purposes.

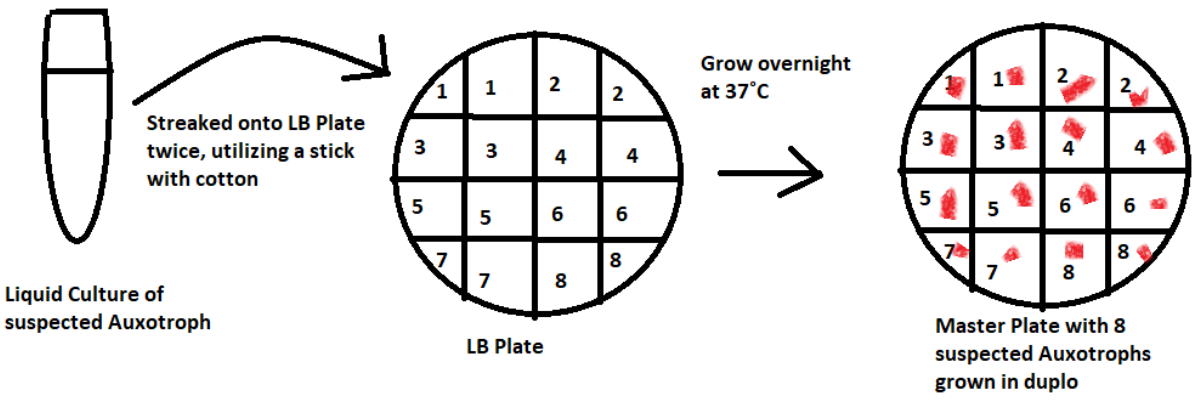


Fig 4.4: Making Auxotrophy Testing Master Plates

Davis Agar plates with vitamins, amino acids or nucleotides added were used to test for specific deficiencies. For nucleotides, 3 µg of RNA were added to 100 ml of DM agar. For amino acids, 500 µl of a solution prepared from Casaminoacids was added to 100 ml of DM agar. For vitamins, 500 µl of a premade vitamin solution provided by the lab was added to 100 ml of DM agar. This vitamin mixture contained per 100 ml of demi water: 10 mg Thiamine (B1), 125 mg Riboflavin (B2), 10 mg Para-Aminobenzoic Acid (B10), 100 mg Nicotinamide (Niacinamide) (B3), 50 mg Pyridoxine HCL (B6), 0.2 mg Biotin (B7) and 10 mg Pantothenic Acid (B5).

Before pouring plates, 80 µl $MGSO_4$, 80 µl thiamine and 100 µl of a 10% glucose solution were added to all three agar mixtures. Solutions added replace equal parts of the distilled matter, to maintain the ratios for DM. Each master plate was stamped DMA, DMA+AA, DMA+VIT, DMA+NT and LBA, in that order, in triplicate. Allowing the stamped plate sets from each master-plate to be grown at 37˚C, 28˚C and 42˚C separately. Colonies not appearing on the DMA plate, but on one of the augmented plates and the LB plate suggests that type of deficiency. The triplicate at three different temperatures allows to verify whether any deficiencies are temperature based. The consistent order of stamping is for internal consistency and to minimalize false positives, by always making the last stamping to an optimal plate.

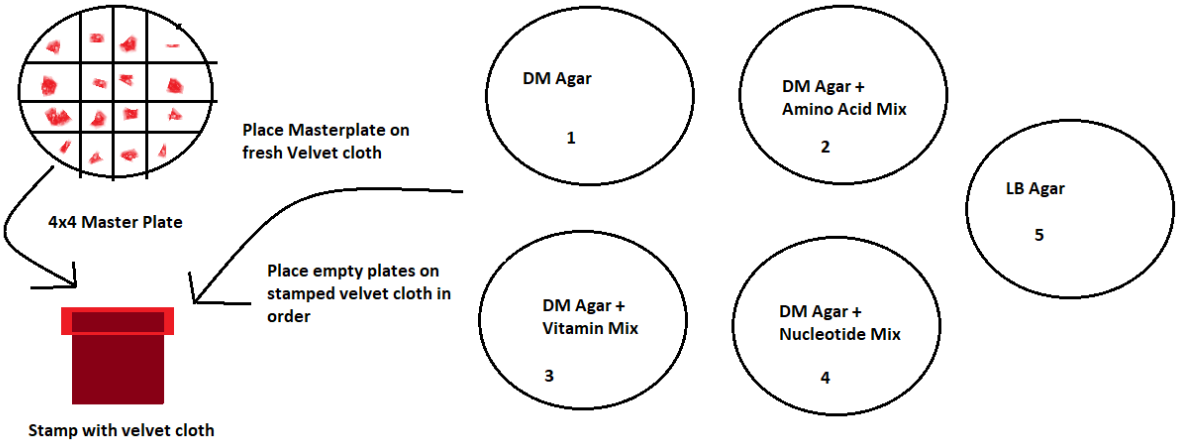


Fig 4.5: Testing for Auxotrophies

5. Results

The end results of 13 weeks of 4 growth days were each stored in glycerol at -20˚C, giving measurable points amongst intervals over an estimated 4000 generations of *E. coli*. This excludes line 6, as at the end of the second week it showed signs of contamination by other bacteria and thus week 3 was grown from the stock made at the end of week 1. As such line 6 is always an estimated 300 generations behind the others. No other contaminations occurred to the main lines. The sets numbered 28 and 42 represent an additional ~300 generations grown at either 28˚C or 42˚C after the end point of set 13.

Over the course of testing the weekly 15 sets of 6 cultures, colonies of interest were found. Singular colonies from 1.2, 1.3, 11.1, 11.2, 11.4, 11.5, 11.6, 13.5 and many from the 28 and 42 sets appeared to be auxotrophic, but only the isolated colonies from 11.4 and 11.6 held up under further scrutiny.

The colony from 11.4 was found to have an amino-acid deficiency.

The colony from 11.6 had a deficiency not solved by the amino acid mix, vitamin mix or nucleotide mix, so its deficiency remains unidentified.

Table 5.1: Results of presumed auxotrophs testing

Sample ID.	Line	Week	Generations	Auxotrophy Temperature	Auxotrophy Testing
1-2.1	2	1	300	42°	None
1-2.2	2	1	300	42°	None
1-3.1	3	1	300	37°, 42°	None
11-1.1	1	11	3300	28°, 37°	None
11-2.1	2	11	3300	28°, 37°	None
11-4.1	**4**	**11**	**3300**	**28°, 37°, 42°**	**Amino Acid**
11-5.1	5	11	3300	28°, 37°, 42°	None
11-6.1	**6**	**11**	**300**	**28°, 37°, 42°**	**Unidentified**
13-5.1	5	13	3900	37°	None
13-5.2	5	13	3900	42°	None
28-1.1	1	14 (28°)	4200	37°	None
28-1.2	1	14(28°)	4200	28°	None
28-1.3	1	14(28°)	4200	42°	None
28-3.1	3	14(28°)	4200	28°	None
28-3.2	3	14(28°)	4200	28°	None
28-3.3	3	14(28°)	4200	42°	None
28-3.4	3	14(28°)	4200	42°	None
28-4.1	4	14(28°)	4200	42°	None
42-1.1	1	14(42°)	4200	37°	None
42-1.2	1	14(42°)	4200	37°	None
42-1.3	1	14(42°)	4200	37°	None
42-1.4	1	14(42°)	4200	37°	None
42-1.5	1	14(42°)	4200	28°	None
42-1.6	1	14(42°)	4200	42°	None
42-2.1	2	14(42°)	4200	42°	None
42-3.1	3	14(42°)	4200	37°	None
42-3.2	3	14(42°)	4200	37°	None
42-3.3	3	14(42°)	4200	37°	None
42-3.4	3	14(42°)	4200	37°	None
42-5.1	5	14(42°)	4200	37°	None
42-5.2	5	14(42°)	4200	37°	None
42-5.3	5	14(42°)	4200	37°	None
42-5.4	5	14(42°)	4200	28°	None
42-5.5	5	14(42°)	4200	42°	None
42-5.6	5	14(42°)	4200	42°	None
42-6.1	6	14(42°)	3900	37°	None
42-6.2	6	14(42°)	3900	37°	None
42-6.3	6	14(42°)	3900	28°	None
42-6.4	6	14(42°)	3900	42°	None
42-6.5	6	14(42°)	3900	42°	None

6. Discussion

6.1 Arisal of auxotrophic mutations in nutrient rich E. coli cultures

After 4000 generations of *E. coli* at 37˚C and some 300 generations more at 28˚C and 42˚C, 2 auxotrophs were picked up amongst a multitude of false positives. These false positives also increased in number amongst the 28˚C and 42˚C sets. It is possible that these "false positives", were simply not pure colonies and the auxotrophs lost out to the prototrophs present upon further growth and culturing for the purpose of testing. A way to deal with this would be to grow these suspected auxotrophs further on plates and not in liquid. The auxotrophs that did show up, did so in week 10 and 11 of the growth phase. None of these were observed again in any of the later weeks. Meaning these did not increase in size in the population and suggests they held no benefits that would allow them to perform better under those specific circumstances.

For the colony from 11.4 it was deduced its deficiency is an amino acid related one. To further identify the amino acid deficiency, further screenings with media containing only a subset of the amino acids found in the original mixture could be done, but was not done for it having gone far over lab scheduled time.

For the colony from 11.6 the set-up did not yet deduce the finer details of the auxotrophy. This is most likely because of the limitations of the set-ups used. Since the vitamin solution used only held a limited selection of B vitamins and the minimal broth as suggested by Lenski contains thiamine, selection for vitamin deficiency was very limited. For the purposes of auxotrophy testing, the purpose of thiamine in the broth might need to be revaluated, its inclusion seems to only be supported by the idea of "nicer growth". While B-Vitamins is the larger group, the other vitamins should also be included in any further testing.

Fatty acids were not considered as possible deficiencies in the original set-up and might be relevant to consider in later studies. Deficiencies not identified in the first preliminary set-ups might also have more than one deficiency, that is not fully covered by one of the selections. Combining selections would be how to identify these, despite these being even less likely than singular deficiencies which already were rare, but it remains a distinct possibility.

While the alternate temperatures the bacteria were grown at for one week, gave no immediate results, it might still be interesting to test the temperature dependence of *E. coli* at more varied temperatures. Particularly temperatures around the edges of what *E. coli* has been found to be able to survive at normally would be of interest to test. Since the set-up was only aerobic, testing under anaerobic conditions might reveal auxotrophies related to pathways only activated in such circumstances.

6.2 Presence of auxotrophic colonies

Returning to the research questions, the first question "Are all natural isolates of *E. coli* prototrophic?", was not covered by the set-up, but is the dogma our understanding of *E. coli* hinges upon, it is discussed further under sub point 4. The second question: "How stable is the prototrophy of prototrophic *E. coli* samples on rich medium?" could be answered as: Over 4000 generations, very stable. Auxotrophies do rarely show up but have disappeared again by the next sample time. Suggesting the metabolic cost of prototrophy over auxotrophy is so small, it is not an evolutionary pressure in this timespan. It might also be that the set-up used introduces an artificial bottleneck upon culture transfer that negatively affects auxotrophs, as it is possible that upon transfer to fresh medium and the following regrowth, the prototrophs can grow faster since they are less dependent on getting specific substances from the medium and thus might be able to overtake auxotrophs in the early transfer stages. The third question: "What is the

mechanism by which prototrophy is lost from *E. coli*, when it is lost?" is not fully answered by the current results, but since the auxotrophs disappeared rather quickly again and thus held no large enough advantages, suggests it is just through sheer mutation accumulation.

6.3 Further testing and mutation experiments

Deducing the specific auxotrophies of the two isolated auxotrophic mutants and then running competition experiments against the ancestral strain at different starting concentrations of mutant to ancestor on different media and under differing conditions like temperature and anaerobic, could lead to a better understanding of why these mutants disappeared again over time between two measurement points. With the auxotrophies identified, it can be tested how much the substances occur in each medium and in what amounts the *E. coli* not auxotrophic for the substance produce it in excess.

For a more direct follow-up or continuation, the existing lines can always be continued under the same protocols, to test further generations and to see any later changes if any. The glycerol stocks, both of the week points and the isolated auxotrophs, can be used for a genome sequencing, to see how much the *E. coli* changed genetically over the 4000 generations and how different the auxotrophies are and what causes them metabolically. Sequencing weekly non-auxotrophs could be done to better estimate the rate of genome wide mutations against generation times, in part since recent studies and papers have put classically held ideas, of mutation rates into question. Drake et al., (1998) quantified and summarized mutation rates, amongst others for *E. coli*, that would have suggested noticeable rates over the course of this experiment. Krašovec et al., (2017), found a negative relationship between mutation rate and population density, which suggest the possibility that in highly dense populations, like the situation in this experiment, mutation rates are reduced.

6.4 Further research on the natural environment of E. coli

As our observations do not fit with hypotheses we hold because of our expectations of the environment of the gut. In the case that the gut is nutritionally rich, and that nutrition is freely available, a purely prototrophic organism would not be expected to be overrepresented. In the case that not all nutrients are freely available, a division of labours like the one proposed by Morris's (2012) Black Queen Hypothesis would be expected, which would be shown by having different autotrophies present in the community. To test these ideas, the investigation of the gut environment as mentioned before is important, but also warrants a closer look at the communities of *E. coli* located in the gut. In particular it would be important to verify the uniformity of the *E. coli* communities, whether any auxotrophies are present and whether these differ within and between communities or not.

It would also be relevant to look at the actual gut environment aside from *E. coli*. Are all nutrients freely available in gut systems, or are they located in sub-niches, potentially monopolized by specific bacterial colonies? How strict is the competition between different species in the environment and does it rely on cross-species syntrophy and if so, how much? Answers to any of those could lead to a better understanding as to why the form of *E. coli* we are most familiar with is a prototrophic one and how that reflects on our base line expectation of *E. coli* being prototrophic in general.

Concerning the second and third research questions as they were formulated in the original proposal "Can naturally occurring auxotrophic *E. coli* be observed? If so, how much do these differ in genotype from prototrophic strains?" and "Is there an observable difference between resident and transient strains in a host organism? If so, can these explain why certain strains are permanent and others are not?", the aim of the proposal was overzealous in its time allocation

and no time was available to do any of the tests assigned. The questions remain relevant to deduce what role the prototrophy and potential auxotrophy of *E. coli* performs in their relationships with their hosts. These could be tested in the ways outlined in the previous paragraphs, which would include genome sequencing.

6.5 Summary of Follow-Up Experiments

1. Continuation of Lines

All the lines are stored and can be continued as is, they could also be restarted from the Ara stocks in case there is a need to change part of the set-up. Regardless, the weekly samples could be sequenced.

2. Identify Auxotrophies of found Auxotrophs

The currently identified auxotrophs are not yet fully identified and could be identified further. Since the original set-up was missing vitamins, thymine etc, the testing mixtures used should either be altered or replaced, to include the substances named to be missing. Sequencing the genomes of identified auxotrophs could also be done.

3. Competition Experiments

To test the viability of the isolated auxotrophs, they should be placed in competition experiments against the ancestral strain. These tests would be done at different starting concentrations of ancestor and auxotroph, different temperatures, different feeding media and both aerobically and anaerobically.

4. Sample E. coli natural environment

The contents of taken gut samples would be identified. This would focus on the material components of the sample, but also the concentrations and genomes of *E. coli* and other inhabitants present. With specific attention to differences between *E. coli* and the identification of transient and non-transient strains.

7. References

Becton Dickinson GmbH. 2011. BD CHROMagar Orientation Medium: INSTRUCTIONS FOR USE –READY-TO-USE PLATED MEDIA.

Bettelheim, K., Breadon, A., Faiers, M. C., O'Farrell, S. M. and Shooter, R. (1974). "The origin of O serotypes of Escherichia coli in babies after normal delivery." Epidemiology & Infection 72(1): 67-70.

Carlton, B. C., and Brown B. J.. (1981) Gene mutation. Pp. 222-242 in P. Gerhardt (ed.), Manual of methods for general bacteriology. American Society for Microbiology, Washington, D.C. 222-242

Castellani, A. and Chalmers, A. (1919). Manual of Tropical Medicine, 3rd ed., Williams Wood and Co., New York.

Caugant, D. A., Levin, B. R. and Selander, R. K. (1981). "Genetic diversity and temporal variation in the E. coli population of a human host." Genetics 98(3): 467-490.

Clermont, O., Bonacorsi, S. and Bingen, E. (2000). "Rapid and simple determination of theEscherichia coli phylogenetic group." Applied and environmental microbiology 66(10): 4555-4558.

Clermont O., Christenson J.K., Denamur E. and Gordon D.M.. "The ClermontEscherichia coliphylo-typing method revisited: improvementof specificity and detection of new phylo-groups phylo-groups". Environmental Microbiology Reports (2013)5(1), 58–65.

Drake, J. W., Charlesworth, B., Charlesworth, D., & Crow, J. F. (1998). "Rates of spontaneous mutation". Genetics, 148(4), 1667-1686.

Escherich, T. (1885). "Die Darmbacterien des Neugeborenen und Säuglings."

Ewing, W. H. (1986). "Edwards and Ewing's identification of Enterobacteriaceae." Edwards and Ewing's identification of Enterobacteriaceae.(Edition 4).

Friedmann, H. C. (2006). "Escherich and Escherichia." Advances in applied microbiology 60: 133-196.

Ishii, S., Meyer, K. P., & Sadowsky, M. J. (2007). Relationship between phylogenetic groups, genotypic clusters, and virulence gene profiles of Escherichia coli strains from diverse human and animal sources. *Appl. Environ. Microbiol.*, *73*(18), 5703-5710.

Ishii, S. and Sadowsky, M. J. (2008). "Escherichia coli in the environment: implications for water quality and human health." Microbes and Environments 23(2): 101-108.

Krašovec R, Richards H, Gifford DR, Hatcher C, Faulkner KJ, et al. (2017). "Spontaneous mutation rate is a plastic trait associated with population density across domains of life". PLOS Biology 15(8): e2002731.

Lan, R. and Reeves, P. R. (2002). "Escherichia coli in disguise: molecular origins of Shigella." Microbes and infection 4(11): 1125-1132.

Lenski, R. E., Rose, M. R., Simpson, S. C. and Tadler, S. C. (1991). "LONG-TERM EXPERIMENTAL EVOLUTION IN ESCHERICHIA-COLI .1. ADAPTATION AND DIVERGENCE DURING 2,000 GENERATIONS." American Naturalist 138(6): 1315-1341.

Lukjancenko, O., Wassenaar, T. M. and Ussery, D. W. (2010). "Comparison of 61 sequenced Escherichia coli genomes." Microbial ecology 60(4): 708-720.

Medawar, P. B. "An Unsolved Problem of Biology." London, UK: H. K. Lewis, 1952.

Merlino, J., Siarakas, S., Robertson, G. J., Funnell, G. R., Gottlieb, T. and Bradbury, R. (1996). "Evaluation of CHROMagar Orientation for differentiation and presumptive identification of gram-negative bacilli and Enterococcus species." Journal of clinical microbiology 34(7): 1788-1793.

Morris, J. J., Lenski, R. E. and Zinser, E. R. (2012). "The Black Queen Hypothesis: evolution of dependencies through adaptive gene loss." MBio 3(2): e00036-00012.

Nowrouzian, F. L., Wold, A. E. and Adlerberth, I. (2005). "Escherichia coli strains belonging to phylogenetic group B2 have superior capacity to persist in the intestinal microflora of infants." The Journal of infectious diseases 191(7): 1078-1083.

Savageau M.A. (1983) *Escherichia-Coli* Habitats, Cell-Types, and Molecular Mechanisms of Gene-Control. Am Nat. 1983;122(6):732–744.

Shiga, K. (1898). "Ueber den dysenteriebacillus (Bacillus dysenteriae)." Zentbl. Bakteriol 24: 817-828.

Tatum, E. and Lederberg, J. (1947). "Gene recombination in the bacterium Escherichia coli." Journal of bacteriology 53(6): 673.

Valen van, L. (1973). "A new evolutionary law." Evol Theory 1: 1-30.

Weinstock, G. M. (2012). "Genomic approaches to studying the human microbiota." Nature 489(7415): 250-256.

Williams, G. C. (1957). “PLEIOTROPY , NATURAL SELECTION, AND THE EVOLUTION OF SENESCENCE” Evolution, Vol. 11, No. 4 (Dec., 1957), pp. 398-411

Zuo, G., Xu, Z. and Hao, B. (2013). "Shigella strains are not clones of Escherichia coli but sister species in the genus Escherichia." Genomics, proteomics & bioinformatics 11(1): 61-65.