ISF 弘立

BAUHINIA

The Student Research Journal of the Independent Schools Foundation Academy
1 Kong Sin Wan Road, Pokfulam, Hong Kong

書院

Volume VI Issue 2, 2020

ISSN 2409-4064

Contents

Editor's Note

Building upon the success of the first five editions, The Editorial Board is proud to present you with the culmination of our students' high-level research in this, sixth edition of Bauhinia. Coming full circle, as expressed by the Chinese concept of Wuxing 五行, we revisit the first part of this cycle and showcase how, despite the global upheaval, our research efforts remain constant.

The stated purpose of this student research journal is to honor excellent scholarship, demonstrated either as part of our distinguished curriculum and elective courses, or as a product of our unique extracurricular programs. The *Eight plus One* Virtues are the foundational values of the ISF community and, while the works herein manifest many of these guiding principles, we wish to highlight in particular our students' pursuit of 'Zhi' (Intelligence & Wisdom) and 'Ai' (Passion for Learning & Life) during this process.

To accommodate the proliferation of research produced by our students, especially during this era of COVID-19, when many other forums for publication were either cancelled or moved online, this volume is divided into two parts: the first issue contains articles that are more Humanities-focused, whereas the theme for the second issue is STEM (Science, Technology, Engineering, and Mathematics).

Please enjoy the diversity of topics Bauhinia VI has to offer, which include the following: how the ghostly figure of the courtesan (Su Xiaoxiao) became a theme in Chinese poetry; why Hellenic Atomism never emerged in ancient China; how the cosmological models of Aristotle and Zhang Heng differed; in what way bacteriophages drive the survival and virulence of *E. Coli* in a soil environment; and how horizontal gene transfer enables *E. Coli* to acquire halotolerance in seawater.

The Editorial Board invites the reader to join this academic community by delving into the topics that sparked such passion in our students. We welcome all engagement and responses: sy_team@isf.edu.hk.

編者的話

在前五期成功的基礎上，編輯部很自豪地將我們學生的高水平研究成果呈現在第六期《紫荊花》中。如中國的「五行」概念所表達的那樣，在完成了一個週期循環後，我們重新進入一個週期的開端，在全球動蕩的情況下，向讀者展示展示我們的研究工作如何保持進行。

這本學生研究期刊的目的是為了表彰優秀的學術成果——無論是作為我們的傑出課程或選修課的一部分，還是作為我們獨特的課外活動的一部分。「八德一智」是弘立大家庭的基本價值觀，雖然這裡的作品展示了許多這些價值原則，但我們希望強調學生在這個過程中對「智」（智識見與智慧）和「愛」（對學習和生活的熱情）的追求。

為了適應我們學生研究工作的發展，特別是在 2019 冠狀病毒流行的時期，許多其他出版被取消或轉移到網上，本卷分為兩期：第一期的文章更注重人文，而第二期的主題是 STEM（科學、數學、工程和數學）。

請欣賞《紫荊花》第六期的多元主題：蘇小小的鬼魂形象如何發展為歷代中國詩歌的一個主題？亞里士多德和張衡的宇宙模型有何不同？古希臘所推崇的原子論在為何在中國古代從未出現？細菌噬菌體（感染細菌的病毒）如何推動大腸桿菌在土壤環境中的生存和毒力？水平基因轉移如何使大腸桿菌在海水中獲得耐鹽性？

編輯部邀請讀者加入到這個學術共同體中來，深入探究引發我們學生激情的話題。我們歡迎大家的參與和回應：sy_team@isf.edu.hk

Artist: Christina Chiu, G12

Title: *In Red*

Medium: Oil paint and fabric on board

Description: Inspired by the work of Tao Aimin, this piece is a reflection of gender identity in Chinese culture through an anthropological lens. While the sense of tranquility portrays how female figures embody history, it is a reminder of how culture is a heritage that travels through time and space and will always remain as part of Chinese women's character.

A Note about Style

Articles included in this publication are written for many different purposes. Any differences in style are due to the need to adhere to the format required for that purpose. Generally, the Modern Language Association (MLA) citation and format style (8th Ed.) is used for articles written in English as part of the Oxford University *Shuyuan* Classics Summer Program or the NRI Scholar's Retreat (Needham Research Institute, at Cambridge University). However, articles writte in the STEM fields were offen adapted from posters students prepared for the American Microbiology Society conference (ASM Microbe) and they adhere to the American Psychological Association (APA) citation and format style (7th Ed.). Articles written in Chinese use footnotes following the style outlined in the Bulletin of the Institute of Chinese Literature and Philosophy. However, articles that were originally submitted as partial fulfillment of the International Baccalaureate (IB) programmes, such as the Middle Years Programme's (MYP) Personal Project or the Diploma Programme's (DP) Extended Essay, have followed the specific requirements as outlined by the student's supervisor, and they are published in this journal as they were originally submitted. A footnote under each article indicates the program from which each piece of work was culled.

關於文體的說明

本出版物中的文章是為許多不同目的而寫的。任何風格上的差異都是由於需要遵守該目的所需的格式。一般來說，牛津大學書院經典暑期班或劍橋大學 NRI 研究所（Needham Research Institute）暑期班的英文文章，採用現代語言協會（MLA）的引文和格式（第 8 版）。然而，STEM 範疇的論文大多引用自美國微生學會研究會上的文章，採用美國心理學會（APA）的引文和格式（第 7 版）。中文撰寫的文章採用中研院《中國文哲研究集刊》的腳注樣式。但是，如果是作為國際文憑課程（IB）的部分內容而提交的文章，如中學課程（MYP）的個人項目或文憑課程（DP）的擴展論文，則按照學生導師提出的具體要求，按原樣在本刊發表。每篇文章下的腳注都注明了文章入選前所屬的項目。

Multidrug-resistant *Enterobacter* sp. and *Serratia Marcescens* from a School Campus in Hong Kong

Cheuk Yiu Allison Cheung

Abstract

Antibiotic resistant microbes pose an increasing worldwide risk to human health. This project examined bacteria extracted from a freshwater stream at a school campus in Hong Kong to determine their susceptibility to antibiotics and to identify the genes responsible for any resistance. Two isolates, which were identified as strains of *Enterobacter* sp. and *Serratia marcescens*, were characterised using draft genomic sequences. Both isolates proved to be multi-drug resistant and were shown to carry many antibiotic resistance genes (ARGs) as well as plasmids linked to virulence. Surprisingly, phylogenetic analysis of the *Enterobacter* sp. found it to be most closely related to *Enterobacter mori*, a plant pathogen.

1. Introduction

Overuse of antibiotics, or antimicrobial cleaning products in the livestock industry and clinical settings, as well as inappropriate prescription of antibiotics are leading causes of multidrug resistance (López and Quirós, 2019). In addition, the fact that most antibiotics are taken orally explains why resistant strains are often found in the intestine and in sewage (Zhang *et al*., 2013; Guo *et al*., 2017). Recent sampling has even shown resistance to be increasing to the newest antibiotics, such as carbapenems, which are typically the 'last line of defence' in many cases of Multi Drug-Resistant infections (López and Quirós, 2019).

Intensive animal husbandry is believed to be a major contributor to the increased environmental burden of antibiotic resistance genes (ARGs). Recent studies show that an abundance of ARGs correlates directly with antibiotic and heavy metal ion concentrations (a non-specific inducer), indicating the importance of the environment in the selection of resistant bacteria (Zhu *et al*., 2013). Diverse, abundant, and potentially mobile ARGs in farm samples suggest that unmonitored use of antibiotics, antimicrobial cleaning agents and heavy metals is linked to the proliferation of ARGs (Davies and Wales, 2019). Between 2000 and 2018, the proportion of drugs to which bacteria have become resistant almost tripled in chickens and pigs, and doubled in cattle, suggesting a likely increase in the levels of 'superbugs' that are resistant to key medicines (Harvey, 2019).

In urban environments, another cause of the increasing levels of antibiotic resistance is the use of pharmaceuticals in clinical settings. Hospitals are crucial hotspots for the dissemination of antibiotic resistant microorganisms so that hospital wastewaters can contribute significantly to the antibiotic resistance profile (Al Salah *et al*., 2020).

Resistance genes can be spread by horizontal gene transfer (the direct uptake of additional genetic material by bacterial cells via plasmids), bacteriophages and even cell-free DNA and then selected for by high levels of antimicrobial agents in the environment. (Zu *et al*., 2013). Plasmids carrying resistance genes (AMR plasmids) are increasingly thought to be responsible for the spread of AMR phenotypes. These mobile forms of self-replicating DNA can adapt rapidly to carry any genes that can offer them and their bacterial hosts a competitive advantage (Madec & Haenni, 2018).

The research for this article was conducted in the *Shuyuan* Molecular Biology Laboratory and a poster about this work was prepared for the ASM Microbe Conference in 2020.

2. Method

2.1 Sample Collection and Screening

Samples were collected from a freshwater stream in Telegraph Bay, Hong Kong. Aliquots were then serial diluted and screened for *E. coli* and coliforms using 3M Petrifilm™ plates.

2.2 Isolation and Passaging

Selected colonies were passaged 10 times on LB agar, which was incubated aerobically at 37 °C degrees. Streaking on MacConkey agar showed that the isolates were likely to be Gram-negative microbes with different abilities to ferment lactose. Isolates were also tested for antibiotic susceptibility via the Kirby-Bauer disk diffusion method.

2.3 DNA Extraction and Tests

Genomic DNA was extracted from two isolates, using a PureLink® Genomic DNA Mini Kit, and shotgun sequencing was performed via the Illumina MiSeq platform to generate draft sequences. NCBI BLAST and autoMLST were used for identification (Johnson *et al*., 2008). PATRIC (Wattam *et al*., 2016) and GhostKoala for KEGG analysis (Kanehisa *et al*., 2016) enabled functional prediction, while PHASTER (Arndt *et al*., 2016) and PlasmidFinder (Carattoli *et al*, 2014) identified putative phage and plasmid sequences. Antimicrobial resistance and virulence genes were determined using the RGI 5.1.0 at CARD (Alcock *et al*., 2020), ResFinder (Zankari *et al*., 2012) and Pathogen Finder (Cosentino *et al*., 2013).

3. Results and Discussion

3.1 Characterisation of the *Serratia marcescens* isolate 32

3.1.1 Identification

Isolate 32 was identified as a strain of *Serratia marcescens*, a Gram-negative bacillus that occurs naturally in soil and water and produces the red pigment prodigiosin at room temperature. *S. marcescens* is the most common species of *Serratia* found in hospitals and the only species reported as pathogenic. Strains have often been found resistant to several antibiotics (Buckle, 2015).

Environmental isolates of *S. marcescens* characteristically produce a red pigment, prodigiosin, found in various ecological niches, including soil, water, air, plants and animals (Hejazi & Falkiner, 1997). The pigment has a role in respiration and has some antibiotic properties, acting as a protective mechanism in unfavourable conditions when the growth of cells is delayed.

	Isolate 31	**Isolate 32**
No. of reads	212368	87940
Average depth	12	5
No. of large contigs	547	2184
No. of bases (large)	4686543	3996156
No. of all contigs	850	2488
N50 contig size	16801	2286
Estimated genome size	5.4 MB	5.4 MB
Homology %	*(E. asburiae* < 94%)	100%
Identity	*Enterobacter sp.*	*Serratia marcescens*

Table 1. Draft genomic sequencing identified isolates 31 and 32 as a species of *Enterobacter* and *Serratia marcescens* respectively. The *Enterobacter* sp. is a poor match for *E. asburiae*

Figure 1. As previously reported by Haddix and Shanks (2018), *Serratia marcescens* does not produce its characteristic pigmentation when incubated at 37 ℃ (left) but prodigiosin synthesis is clearly evident at 27 ℃ (right).

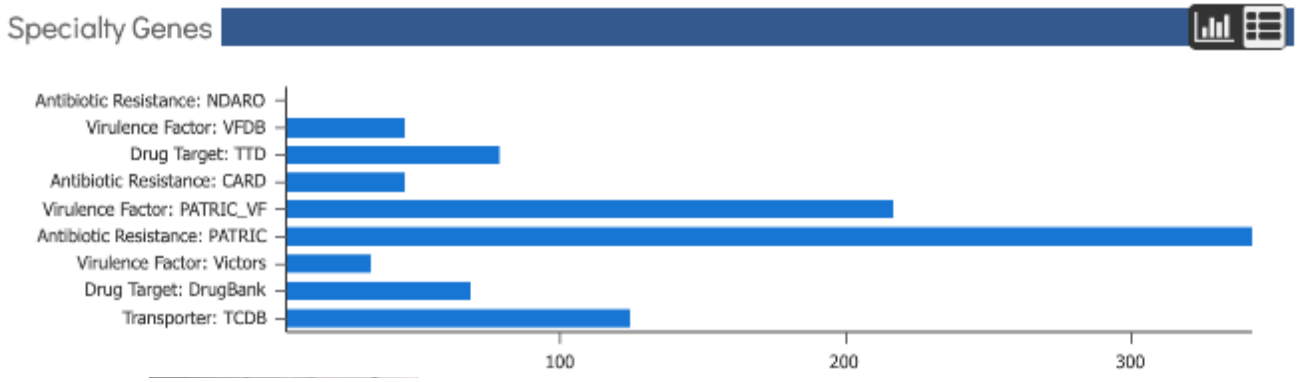


Figure 3. Analysis of specialty genes by PATRIC reveals the predominance of resistance and virulence genes in this strain of *S. marcescens*.

3.1.2 Antimicrobial Resistance

The *S. marcescens* is resistant to the antibiotics Cephalothin (a cephalosporin), Clindamycin (a lincosamide), Rifampicin (an ansamycin) and Vancomycin (a glycopeptide), as shown from functional test results using the Kirby-Bauer disk diffusion assay. The isolate also contains the *Klebsiella pneumoniae* AMR gene KpnH, which may account for the resistance to cephalosporin (Ekwanzala *et al*., 2019). In addition, it also contains a mutant version of the CRP regulator, which is likely to increase the activity of multidrug efflux pumps that help the bacterium to eliminate a range of antibiotics (Nishino *et al*., 2008).

3.1.3 Virulence and Plasmids

3.1.4 Prophages

The *Serratia* carries the prophage Klebsi_phiKO2 which contributes to the genome plasticity of *K. pneumoniae* (Casjens *et al*., 2004). This is a linear plasmid phage that can stimulate the assimilation of virulence and resistance genes (Shen *et al*., 2019).

Figure 2. *Serratia marcescens* isolate 32 is resistant to Cephalothin (a cephalosporin)

3.2 Characterisation of the *Enterobacter sp.* isolate 31

3.2.1 Identification

Isolate 31 was identified as an *Enterobacter sp*., although with less than 94% homology with its closest relative, *Enterobacter asburiae* str. AEB30. *Enterobacter* are straight, Gram-negative bacilli that naturally inhabit the human intestinal tract, soil, water, animals, plants and are common in foods. Nevertheless, they have been reported as emerging pathogens, especially in immunocompromised patients within the hospital setting, where they are responsible for a variety of different infections that can be challenging to treat due to intrinsic and acquired resistance to antibiotics (Kus, 2014).

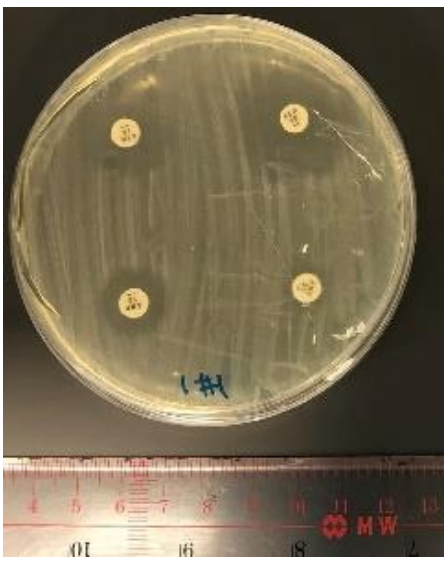


Figure 4. *Enterobacter* isolate 31 is resistant to Augmentin – a next-generation penicillin preparation that is commonly prescribed in Hong Kong.

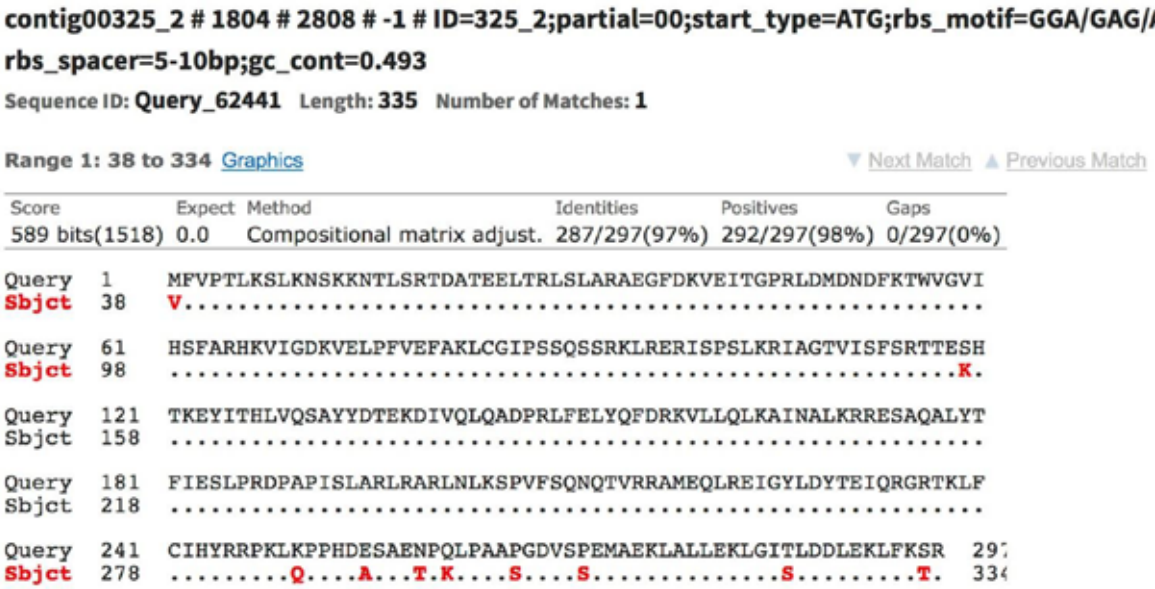


Figure 5. BLAST (protein) results for contig 00325 of isolate 31 shows alignment with IncFIB reference sequence.

3.2.2 Antimicrobial resistance

The *Enterobacter spp.* is resistant to Ampicillin, Augmentin, Cephalothin, Clindamycin, Rifampicin and Vancomycin.

A number of AMR genes is likely to be responsible for the broad range of resistance in this isolate. For example, the presence of the KpnE-KpnF efflux system can give rise to multidrug resistance, including to Rifampicin (Srinivasan & Rajamohan, 2013). The isolate also has several genes that act specifically on penicillins and cephalosporins, including PBP3 and two beta-lactamases, $bla_{\text{ACT-29}}$ and *ampH*.

Interestingly, this isolate is not only resistant to Ampicillin (a penicillin) and Cephalothin (a cephalosporin), but also to Augmentin, which contains clavulanic acid to counteract resistance due to the beta-lactamase enzyme (Drawz and Bonomo, 2010). The two beta-lactamases, $bla_{\text{ACT-29}}$ and *ampH* are both implicated, but a similar $bla_{\text{ACT-29}}$ has been reported in an isolate of *Enterobacter mori* from Austria (Hartl *et al*., 2019) and in this case it was also found resistant to carbapenems, which are penicillin-like molecules specifically developed to be resistant to degradation by beta-lactamases (Papp-Wallace *et al*., 2011). Beta-lactamases that are able to deactivate penicillins, cephalosporins and carbapenems are known as *extended-spectrum β-lactamases* (ESBLs) and these have been increasingly associated with multi-drug resistant infections throughout the world (Murray and Peaper, 2015).

3.2.3 Virulence and Plasmids

The *Enterobacter* contains the plasmid-encoded *fosA2*, *merA* and *qnrE2* (a qnrE-member of plasmid-borne (fluoro)quinolone-resistance genes). Similar isolates bearing these genes have been found in water samples from Hangzhou (Zhou *et al*., 2017).

Sequence analysis suggests that the *Enterobacter* also contains the plasmids IncFIB(pB171) and IncFII(pECLA), self-replicating, mobile genetic elements which can encode both virulence factors and antimicrobial resistance genes. Similar plasmids have also been found in hospital settings in China (Zhou *et al*., 2020).

The image below shows protein BLAST results from an alignment of contig 00325_2 and Ref(erence) FIB. The V at the start of the sequence indicates the less common GTG start codon is used and the close alignment confirms that the *Enterobacter* is likely to contain a variant of the IncFIB plasmid.

D-alanyl-D-alanine-carboxypeptidase/endopeptidase AmpH [Enterobacter mori]

Sequence ID: WP_100167146.1 **Length:** 386 **Number of Matches:** 1

See 2 more title(s) Identical Proteins

Range 1: 1 to 386 GenPept Graphics ▼ Next Match ▲ Previous Match

Score	Expect	Method	Identities	Positives	Gaps
796 bits(2056)	0.0	Compositional matrix adjust.	385/386(99%)	386/386(100%)	0/386(0%)

```
Query  1    MKRCLLSFAALCAVSFSTAQAAQPLTAPVLASDIADRYANLIYYGSGGTGMALVVIDGNQ  60
Sbjct  1    ............................................................  60

Query  61   RVFRSFGETRPGNNIHPQLDSVIRIASITKLMTSEMLVKLLDQGVVKLDDPLSKYAPPGA  120
Sbjct  61   ..............V.............................................  120

Query  121  RVPTYQGAPIRLVNLATHTSALPREQPGGAAHRPVFVWPTREQRWNYLSTATLKTTPGSQ  180
Sbjct  121  ............................................................  180

Query  181  AAYSNLAFDLLADALSTASGKPYPQLFEEQITRPLGMKDTTFTPSPDQCRRLMVAEKGAS  240
Sbjct  181  ............................................................  240

Query  241  PCNNTLAAIGSGGVYSTPGDMMRWMQQFLSSDFYSRTHQADRMQTLIYQRTQLTRVIGMD  300
Sbjct  241  ............................................................  300

Query  301  VPGKADALGLGWVYMAPKDGRPGIIQKTGGGGGFITYMAMIPQSNVGAFVVVTRSPNTRF  360
Sbjct  301  ............................................................  360

Query  361  VNMSDGINNLVAELSANKAQVLTASN  386
Sbjct  361  .........................  386
```

Figure 6. The AmpH amino acid sequence of Isolate 31 aligns well with the *E. mori* protein.

3.2.4 Prophages

The *Enterobacter* was found to carry an active version of the prolific temperate prophage phi80 "Lula". Its cryptic lysogenic productivity and stealthy infectivity facilitates its inconspicuous spread that has been widely reported (Rotman *et al*., 2012).

3.2.5 Enzymes

The *Enterobacter* possesses the chitinase EC 3.2.1.14, which is a potential agent for biological control of plant diseases and pathogens. There is also urease EC 3.5.1.5, mainly found in plant seeds which catalyzes the hydrolysis of urea. It also has EC 3.2.1.122 and invertase EC 3.2.1.26, which hydrolyse 6-phospho-alpha-D-glucosides and beta-D-fructofuranosides respectively. It also contains trehalase EC 3.2.1.28 in contig 157, which hydrolyzes α,α'-trehalose as well as beta-glucosidase, EC 3.2.1.21 and endoglucanase/ cellulase (EC 3.2.1.4) that are essential for the degradation of cellulose. Given the extent of its ability to handle a wide range of carbohydrates, it is possible that – like the closely-related *Enterobacter mori* – the bacterium is ideally suited to degrading plant material.

3.2.6 Identifying the *Enterobacter* isolate 31

Sequence comparison

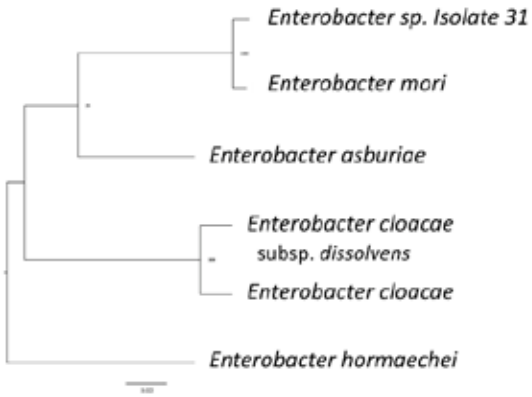


Figure 7. Phylogenetic tree showing distance between *Enterobacter* isolate 31 and other *Enterobacter* species, generated by PATRIC based on 1000 genes (Wattam *et al*., 2016)

Initial BLAST results indicated that *Enterobacter* isolate 31 was closest to *Enterobacter asburiae*, although at < 94% it seemed a poor fit. On the other hand, some of the isolate's proteins – such as the beta-lactamase AmpH, align very well with the plant pathogen *Enterobacter mori* (see Figure. 7).

Using PATRIC (Wattam *et al*., 2016), it was possible to construct a phylogenetic tree, based on a much larger number of genes, to find out how isolate 31 is related to the other *Enterobacteraceae*: it is clearly closer to *Enterobacter mori* than to other known species (see Figure. 8).

Functional comparison

If Isolate 31 is a sub-species of *Enterobacter mori*, it would be expected to behave in a similar way in terms of biochemical or metabolic functions. Using KEGG analysis (Kanehisa *et al*., 2016) and observations in culture, it is possible to compare the isolate with *E. mori* and other *Enterobacteriaceae* (see Table 2).

When isolate 31 is compared to the other Enterobacteraceae, including E. mori, it shows a distinctive set of functions. It is possible that either the isolate is an atypical example or sub-species of E. mori or should be designated as a close but separate species.

1. *E. asburiae*; 2. *E. cloacae subsp. cloacae*; 3. *E. cloacae* subsp. dissolvens; 4. *E. hormaechei*; 5. *E. mori*
[+, present or shown; –, absent; v, variable]

	1	**2**	**3**	**4**	**5**	**Isolate 31**	*notes*
Aesculin hydrolysis	+	v	+	–	+	+	3.2.1.21 β-glucosidase present
Arginine dihydrolase	v	+	+	v	+	–	
Lysine decarboxylase	–	–	–	–	+	–	
Methyl red test	+	–	–	–	–	+	
Motility	v	+	(–)	v	+	+	
Ornithine decarboxylase 4.1.1.17	+	+	+	+	+	+	
Voges-Proskauer reaction	+	+	+	+	+	+	
Utilisation of:							
1-O-Methyl- α-galacto-pyranoside	(+)	+	+	–	v	–	3.2.1.22 α-galactosidase absent
Citrate	+	+	+	+	+	–	
D-Arabitol 1.1.1.11/250/287	–	v	–	(–)	+	–	
D-Sorbitol 1.1.1.14	+	+	+	(–)	+	–	
L-Fucose 5.3.1.25	–	–	v	+	v	–	
Melibiose 3.2.1.22	–	(+)	+	(–)	+	–	3.2.1.22 α-galactosidase absent
Mucate (D-galactarate)	v	v	+	+	+	–	4.2.1.42 & 5.5.1.27 absent
Putrescine	(-)	+	+	–	+	+	2.6.1.82 present
Sucrose	+	+	+	+	+	+	3.2.1.20 present
α-L-Rhamnose	(+)	+	(+)	+	+	–	
Urea	–	–	–	(+)	+	+	urease present
D-gluconate	–	–	–	+	–	(–)	

Table 2. The *Enterobacter* isolate 31 shows a distinct set offunctions and is a poor match for *Enterobacter mori.*

Conclusion

Serratia marcescens and *Enterobacter* are two multidrug resistant potential human pathogens isolated from a freshwater stream near a school that flows past a hospital and medical school. They are common bacteria in the environment, but *S. marcescens* isolate 32 and *Enterobacter* isolate 31 show resistance to a fairly broad range of antibiotics.

Bacteria become adapted to the places they inhabit so that the genes they contain give them an advantage over others that are less well-adapted. According to its biochemical functions, the *Enterobacter* isolate 31 seems well-suited to digesting plant material, just like *Enterobacter mori*. On the other hand, it carries extended spectrum beta-lactamases suggesting that being resistant to Augmentin and cephalosporins is helpful to it. The nearby hospital is certainly a possible source of high antibiotic concentrations and this will need further investigation.

Although the *Enterobacter* at first appeared similar to *E. mori*, functional analysis finds significant differences. And if it has evolved from a similar plant pathogen, the high level of antimicrobial resistance and plasmid-encoded virulence suggest that it may now represent a human pathogen. Could this mean that over-exposure to antibiotics might be helping to turn plant bacteria into human pathogens?

Next Steps

An extension of this work will be to obtain the full sequences of the *Enterobacter sp*. and *Serratia marcescens* genomes. Whole genome sequencing will help to characterise plasmids, which often carry AMR and virulence genes. It will also be useful to locate the ESBL genes so as to better understand why the *Enterobacter* has similar plasmids to isolates from livestock or clinics.

Further investigation into the environment where the isolates were found could also be helpful in identifying why these traits may have arisen. For example, it may be possible to collect samples from freshwater locations near the initial spot of discovery for analysis. This can help pinpoint other possible sources and hotspots that may be encouraging the evolution of multidrug resistant 'superbugs'.

Acknowledgements

Many thanks to Dr Simon Griffin, Professor Fred Leung and Ms Grace Lai for their continuous support in this project. Special thanks to the *Shuyuan* Science team for providing equipment and the means to complete this project. This work is funded under The ISF Foundation.

References

Agrios, G.N. (2009) Plant Pathogens and Disease: General Introduction. In Schaechter, M. (Ed*.), Encyclopedia of Microbiology* (Third Edition), Academic Press; vol. 4, pp. 613-646. doi: 10.1016/B978-012373944-5.00344-8

Al Salah, D. M. M., Ngweme, G. N., Laffite, A., Otamonga, J.-P., Mulaji, C., Poté, J. (2020) Hospital wastewaters: A reservoir and source of clinically relevant bacteria and antibiotic resistant genes dissemination in urban river under tropical conditions. *Ecotoxicology and Environmental Safety*, 200: 110767. doi: 10.1016/j.ecoenv.2020.110767

Alcock, B. P., Raphenya, A. R., Lau, T. T. Y., Tsang, K. K., Bouchard, M., Edalatmand, A., … McArthur, A. G. (2020) CARD 2020: antibiotic resistome surveillance with the comprehensive antibiotic resistance database. *Nucleic Acids Research*, 48(D1), D517–D525. doi: 10.1093/nar/gkz935

Arndt, D., Grant, J. R., Marcu, A., Sajed, T., Pon, A., Liang, Y., Wishart, D. S. (2016). PHASTER: a better, faster version of the PHAST phage search tool. *Nucleic Acids Research,* 44(W1), W16–W21. doi: 10.1093/nar/gkw387

Bouvet, O. M. M., Lenormand, P., Grimont, P. A. D. (1989). Taxonomic Diversity of the D-Glucose Oxidation Pathway in the *Enterobacteriaceae. International Journal of Systematic Bacteriology*, 39(1), 61–67. doi: 10.1099/00207713-39-1-61

Buckle, J. (2015). Infection. *Clinical Aromatherapy: Essential Oils in Healthcare* (3rd Edn.), Elsevier (St. Louis, MO); Ch. 7, pp. 130–167. doi: 10.1016/b978-0-7020-5440-2.00007-3

Carattoli, A. (2009). Resistance Plasmid Families in *Enterobacteriaceae. Antimicrobial Agents and Chemotherapy*, 53(6), 2227–2238. doi: 10.1128/aac.01707-08

Carattoli, A., Zankari, E., García-Fernández, A., Voldby Larsen, M., Lund, O., Villa, L., … Hasman, H. (2014). In *SilicoDetection* and Typing of Plasmids using PlasmidFinder and Plasmid Multilocus Sequence Typing. *Antimicrobial Agents and Chemotherapy,* 58(7), 3895–3903. doi: 10.1128/aac.02412-14

Casjens, S. R., Gilcrease, E. B., Huang, W. M., Bunny, K. L., Pedulla, M. L., Ford, M. E., … Hendrix, R. W. (2004). The pKO2 linear plasmid prophage of *Klebsiella oxytoca. Journal of Bacteriology*, 186(6), 1818–1832. doi: 10.1128/jb.186.6.1818-1832.2004

Cosentino, S., Voldby Larsen, M., Møller Aarestrup, F., Lund, O. (2013) PathogenFinder - Distinguishing Friend from Foe Using Bacterial Whole Genome Sequence Data. *PLoS ONE*, 8(10): e77302. doi: 10.1371/journal.pone.0077302 URL:cge.cbs.dtu.dk/services/PathogenFinder/

Davies, R., Wales, A. (2019) Antimicrobial Resistance on Farms: A Review Including Biosecurity and the Potential Role of Disinfectants in Resistance Selection*, Comprehensive Reviews in Food Science and Food Safety*, 18, 753-774. doi: 10.1111/1541-4337.12438

D-BITE: Database of Biochemical Tests of Pathogenic *Enterobacteriaceae* Family (2015). *Enterobacter hormaechei.* Retrieved June 10, 2020, from: bioinfo.bisr.res.in/cgi-bin/project/docter/get_details.cgi?organism=enterobacter&&species=Enterobacter%20hormaechei

Drawz, S. M., Bonomo, R. A. (2010). Three decades of beta-lactamase inhibitors. *Clinical Microbiology Reviews*, 23(1), 160–201. doi: 10.1128/CMR.00037-09

Ekwanzala, M. D., Dewar, J. B., Kamika, I., Momba, M. N. B. (2019). Tracking the environmental dissemination of carbapenem-resistant *Klebsiella pneumoniae* using whole genome sequencing. *Science of The Total Environment*, 691, 80–92. doi: 10.1016/j.scitotenv.2019.06.533

Glycopeptide antibiotic. (2020, April 13). Retrieved June 12, 2020, from Wikipedia website: en.wikipedia.org/wiki/Glycopeptide_antibiotic

Guo, J., Li, J., Chen, H., Bond, P., Yuan, Z. (2017) Metagenomic analysis reveals wastewater treatment plants as hotspots of antibiotic resistance genes and mobile genetic elements, *Water Research*, 123, 468-478. doi: 10.1016/j.watres.2017.07.002

Haddix, P. L., Shanks, R. M. Q. (2018). Prodigiosin pigment of *Serratia marcescens* is associated with increased biomass production. *Archives of Microbiology*, 200(7), 989–999. doi: 10.1007/s00203-018-1508-0

Hartl, R., Kerschner, H., Gattringer, R., Lepuschitz, S., Allerberger, F., Sorschag, S., Ruppitsch, W., Apfalter, P. (2019). Whole-Genome Analysis of a Human *Enterobacter mori* Isolate Carrying a blaIMI-2 Carbapenemase in Austria. *Microbial Drug Resistance*, 25(1), 94–96. doi: 10.1089/mdr.2018.0098

Harvey, F. (2019). Superbug hotspots emerging in farms across globe - study. *The Guardian* (UK), Thurs. 19 Sep. URL: www.theguardian.com./environment/2019/sep/19/superbug-hotspots-emerging-in-farms-across-globe-study

Hejazi, A., Falkiner, F. R. (1997) *Serratia marcescens*. *Journal of Medical Microbiology*, 46(11), 903–912. doi: 10.1099/00222615-46-11-903

Hoffmann, H., Stindl, S., Ludwig, W., Stumpf, A., Mehlen, A., Monget, D., … Schleifer, K. H. (2005). *Enterobacter hormaechei* subsp. *oharae* subsp. nov., *E. hormaechei* subsp. *hormaechei* comb. nov., and *E. hormaechei* subsp. *steigerwaltii* subsp. nov., Three New Subspecies of Clinical Importance. *Journal of Clinical Microbiology,* 43(7), 3297–3303. doi: 10.1128/jcm.43.7.3297-3303.2005

Holmes, A.H., Moore, L.S.P., Sundsfjord, A., Steinbakk, M., Regmi, S., Karkey, A., Guerin, P.J., Piddock, L.J.V. (2016) Understanding the mechanisms and drivers of antimicrobial resistance, *The Lancet*, 387(10014), 176-187. doi: 10.1016/S0140-6736(15)00473-0

Jeske, L., Placzek, S., Schomburg, I., Chang, A., Schomburg, D. (2019) BRENDA in 2019: a European ELIXIR core data resource, *Nucleic Acids Research*, 47: D542–D549. doi:10.1093/nar/gky1048. URL: www.brenda-enzymes.org/index.php

Joensen, K. G., Scheutz, F., Lund, O., Hasman, H., Kaas, R. S., Nielsen, E. M., Aarestrup, F. M. (2014) Real-time whole-genome sequencing for routine typing, surveillance, and outbreak detection of verotoxigenic *Escherichia coli*. *Journal of Clinical Microbiology*, 52(5), 1501–1510. doi: 10.1128/JCM.03617-13

Johnson, M., Zaretskaya, I., Raytselis, Y., Merezhuk, Y., McGinnis, S., Madden, T. L. (2008). NCBI BLAST: a better web interface. *Nucleic Acids Research*, 36(Web Server), W5–W9. doi: 10.1093/nar/gkn201

Johnson, T. J., Nolan, L. K. (2009). Pathogenomics of the Virulence Plasmids of *Escherichia* coli. *Microbiology and Molecular Biology Reviews,* 73(4), 750–774. doi: 10.1128/mmbr.00015-09

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

Kanehisa, M., Sato, Y., Morishima, K. (2016) BlastKOALA and GhostKOALA: KEGG Tools for Functional Characterization of Genome and Metagenome Sequences. *Journal of Molecular Biology*, 428(4), 726–731. doi: 10.1016/j.jmb.2015.11.006

Konieczna, I., Zarnowiec, P., Kwinkowski, M., Kolesinska, B., Fraczyk, J., Kaminski, Z., Kaca, W. (2012). Bacterial Urease and its Role in Long-Lasting Human Diseases. *Current Protein and Peptide Science*, 13(8), 789–806. doi: 10.2174/138920312804871094

Kus, J. V. (2014) Infections due to *Citrobacter* and *Enterobacter*, *Reference Module in Biomedical Sciences*, Elsevier. doi: 10.1016/b978-0-12-801238-3.05089-3

López Romo, A., Quirós, R. (2019). Appropriate use of antibiotics: an unmet need. *Therapeutic Advances in Urology*, 11, 9-17. doi: 10.1177/1756287219832174

Madec, J.-Y., Haenni, M. (2018). Antimicrobial resistance plasmid reservoir in food and food-producing animals. *Plasmid*, 99, 72–81. doi:10.1016/j.plasmid.2018.09.001

Mathys, D. A., Mollenkopf, D. F., Feicht, S. M., Adams, R. J., Albers, A. L., Stuever, D. M., … Wittum, T. E. (2019). Carbapenemase-producing *Enterobacteriaceae* and *Aeromonas* spp. present in wastewater treatment plant effluent and nearby surface waters in the US. *PLOS ONE*, 14(6), e0218650. doi:10.1371/journal.pone.0218650

McGinnis, S., Madden, T. L. (2004) BLAST: at the core of a powerful and diverse set of sequence analysis tools. *Nucleic Acids Research*, 32, W20–W25. doi: 10.1093/nar/gkh435. URL: blast.ncbi.nlm.nih.gov/Blast.cgi

MerckSource. (2009, March 22). Dorlands Medical Dictionary:cephalosporin. Retrieved June 12, 2020, from web.archive.org website: web.archive.org/web/20090322002521/http://www.mercksource.com/pp/us/cns/cns_hl_dorlands_split.jsp?pg=/ppdocs/us/common/dorlands/dorland/two/000019430.htm

Microbiology Info. (2018a, June 12). Voges–Proskauer (VP) Test-Principle, Reagents, Procedure and Result. Retrieved June 1, 2020, from Microbiology Info.com website: microbiologyinfo.com/voges-proskauer-vp-test-principle-reagents-procedure-and-result/

Microbiology Info. (2018b, October 25). Methyl Red (MR) Test-Principle, Procedure and Result Interpretation. Retrieved June 1, 2020, from Microbiology Info.com website: microbiologyinfo.com/methyl-red-mr-test-principle-procedure-and-result-interpretation/

Murray, T. S., Peaper, D. R. (2015). The contribution of extended-spectrum β-lactamases to multidrug-resistant infections in children. *Current Opinion in Pediatrics*, 27(1), 124–131. doi: 10.1097/MOP.0000000000000182

Nishijima, K. A. (1993, January). *Enterobacter cloacae*. Retrieved June 10, 2020, from www.extento.hawaii.edu website: www.extento.hawaii.edu/kbase/crop/Type/e_cloac.htm#:~:text=Enterobacter%20cloacae%20is%20a%20gram

Nishino, K., Senda, Y, Yamaguchi, A., (2008). CRP Regulator Modulates Multidrug Resistance of *Escherichia coli* by Repressing the mdtEF Multidrug Efflux Genes. *J. Antibiot.* 61(3): 120–127

Octavia, S., Lan, R. (2014). The Family *Enterobacteriaceae*. *The Prokaryotes,* 225–286. doi: 10.1007/978-3-642-38922-1_167

O'Hara, C. M., Steigerwalt, A. G., Hill, B. C., Farmer, J. J., Fanning, G. R., Brenner, D. J. (1989*) Enterobacter hormaechei, a new species of the family Enterobacteriaceae* formerly known as Enteric Group 75. *Journal of Clinical Microbiology*, 27, 2046-2049.

Papp-Wallace, K. M., Endimiani, A., Taracila, M. A., Bonomo, R. A. (2011). Carbapenems: past, present, and future. *Antimicrobial Agents and Chemotherapy*, 55(11), 4943–4960. doi: 10.1128/AAC.00296-11

Rotman, E., Kouzminova, E., Plunkett, G., Kuzminov, A. (2012) Genome of Enterobacteriophage Lula/phi80 and insights into its ability to spread in the laboratory environment. *Journal of Bacteriology*, 194(24), 6802–6817. doi: 10.1128/JB.01353-12

Shen, J., Zhou, J., Xu, Y., Xiu, Z. (2019) Prophages contribute to genome plasticity of *Klebsiella pneumoniae* and may involve the chromosomal integration of ARGs in CG258. *Genomics*.112(1), 998-1010. doi: 10.1016/j.ygeno.2019.06.016

Srinivasan, V. B., Rajamohan, G. (2013). KpnEF, a New Member of the *Klebsiella pneumoniae* Cell Envelope Stress Response Regulon, Is an SMR-Type Efflux Pump Involved in Broad-Spectrum Antimicrobial Resistance. *Antimicrobial Agents and Chemotherapy*, 57(9), 4449–4462. doi: 10.1128/AAC.02284-12

VirtualUnknownTM Microbiology. (2012a) Arginine dihydrolase Test. Retrieved June 1, 2020, from www.vumicro.com website: https://www.vumicro.com/vumie/help/VUMICRO/Arginine_dihydrolase_Test.htm

VirtualUnknownTM Microbiology. (2012b). Lysine Decarboxylase Test. Retrieved June 1, 2020, from www.vumicro.com website: https://www.vumicro.com/vumie/help/VUMICRO/Lysine_decarboxylase_Test.htm

Wattam, A. R., Davis, J. J., Assaf, R., Boisvert, S., Brettin, T., Bun, C., … Stevens, R. L. (2016). Improvements to PATRIC, the all-bacterial Bioinformatics Database and Analysis Resource Center. *Nucleic Acids Research*, 45(D1), D535–D542. doi: 10.1093/nar/gkw1017

Wehrli, W., Staehelin, M. (1971). Actions of the rifamycins. *Bacteriological Reviews*, 35(3), 290–309. doi: 10.1128/mmbr.35.3.290-309.1971

Wu, W., Zong, Z. (2019). Genome analysis-based reclassification of *Enterobacter* tabaci Duan *et al*. 2016 as a later heterotypic synonym of *Enterobacter mori* Zhu *et al*. 2011. *International Journal of Systematic and Evolutionary Microbiology*, 70(2). doi: 10.1099/ijsem.0.003871

Wunderlichová, L., Buňková, L., Koutný, M., Jančová, P., Buňka, F. (2014). Formation, Degradation, and Detoxification of Putrescine by Foodborne Bacteria: A Review. *Comprehensive Reviews in Food Science and Food Safety*, 13(5), 1012–1030. doi: 10.1111/1541-4337.12099

Yang, Q.-E., Sun, J., Li, L., Deng, H., Liu, B.-T., Fang, L.-X., … Liu, Y.-H. (2015). IncF plasmid diversity in multi-drug resistant *Escherichia coli* strains from animals in China. *Frontiers in Microbiology*, 6: 964 doi: 10.3389/fmicb.2015.00964

Zankari, E., Hasman, H., Cosentino, S., Vestergaard, M., Rasmussen, S., Lund, O., … Larsen, M. V. (2012). Identification of acquired antimicrobial resistance genes. *Journal of Antimicrobial Chemotherapy*, 67(11), 2640–2644. doi: 10.1093/jac/dks261

Zhang, L., Huang, Y., Zhou, Y., Buckley, T., Wang, H. H. (2013) Antibiotic administration routes significantly influence the levels of antibiotic resistance in gut microbiota. *Antimicrobial Agents and Chemotherapy*, 57(8), 3659–3666. doi: 10.1128/AAC.00670-13

Zhou, H., Zhang, K., Chen, W., Chen, J., Zheng, J., Liu, C., … Cao, X. (2020). Epidemiological characteristics of carbapenem-resistant *Enterobacteriaceae* collected from 17 hospitals in Nanjing district of China. *Antimicrobial Resistance & Infection Control*, 9(1). doi: 10.1186/s13756-019-0674-4

Zhou, H.-W., Zhang, T., Ma, J.-H. Fang, Y., Wang, H.-Y., Huang, Z.-X., Wang, Y., Wu, C., Chen, G.X. (2017). Occurrence of Plasmid- and Chromosome-Carried mcr-1 in Waterborne *Enterobacteriacea*e in China. *Antimicrobial Agents and Chemotherapy*, 61(8). doi: 10.1128/AAC.00017-17

Zhu, Y.-G., Johnson, T. A., Su, J.-Q., Qiao, M., Guo, G.-X., Stedtfeld, R. D., … Tiedje, J. M. (2013). Diverse and abundant antibiotic resistance genes in Chinese swine farms. *Proceedings of the National Academy of Sciences*, 110(9), 3435–3440. doi: 10.1073/pnas.1222743110

An Analysis and Comparison of *Micrococcus Luteus* Isolates from Soil and Indoor Air

Charlotte A. C. M. Wong

Abstract

Micrococcus luteus (formerly *M. lysodeikticus*) is a Gram-positive, coccus shaped, tetrad-arranged bacterium which is found in a selection of environments such as soil, water, mammalian skin, and air. It is very widespread and therefore highly adaptable to a variety of conditions. Surprisingly then, for such a versatile bacterium, most isolates sequenced have small genomes, not much larger than 2.5 Mb. The genomic comparison of isolates drawn from two different environments, air and soil, may help to elucidate some reasons for its versatility.

Two isolates were recovered using the IUL Spin Air basic sampler; a third isolate was recovered from nearby soil after suspension in PBS. Phylogeny by autoMLST places the air isolates closer than the soil isolate, although all three share a branch with *M. luteus* strains SK58 and O'Kane. In addition, the three isolates share many functional characteristics: for example, their antibiotic profiles and their strong caseinolytic activity in culture associated with the identical extracellular S8 serine protease, *vpr*, bearing the same MKAEA signal peptide.

Despite their similarities, the estimated genome size of the soil isolate was over 30% larger (3.5 Mb), likely due to the higher number of plasmids. These included pSGAir0127, a plasmid recently identified in an airborne isolate in Singapore, which was present in all three isolates.

1. Introduction

Microbes are able to inhabit a huge variety of environments due to their ability to adapt to different conditions through manipulation of their physiology and/or their immediate surroundings. If a bacterium enters an unfamiliar environment but requires additional genes in order to thrive, it may be able to acquire them from external sources such as bacteriophages (viruses), as plasmids from related bacteria nearby or even as cell-free DNA fragments from the medium (*see* Mosig and Calendar, 2002; Harrison and Brockhurst, 2012; Sun, 2018) .

Micrococcus luteus is an extraordinarily versatile bacterium because it can be found in environments as diverse as seawater (Cabaj and Kosakowska, 2009) and freshwater (Zothanpuia *et al*., 2018), in soil (Tuleva *et al*., 2009), air (Kutmutia *et al*., 2019) and on surfaces such as human skin (Hanafy *et al.*, 2016; Souak *et al*., 2020). All of these environments require specialisation and in a number of cases *M. luteus* isolates have even shown potential for bioremediation because they are able to metabolise specialist substrates (*e.g.* Jampasri *et al*., 2020; Olowomofe *et al*., 2019; Yavas and Icgen, 2018; Hanafy *et al*., 2016).

Given that *M. luteus* is so versatile metabolically, its genome might be expected to contain a very large number of genes to allow so many possible biochemical pathways. However, its genome size is typically around 2.6 Mb and the reference genome *Micrococcus luteus NCTC 2665* encodes only 2,200 proteins from 2,331 genes. For comparison, the genome size of *Escherichia coli* str. K-12 substr. MG1655 is 4.64 Mb an d encodes only 4,242 proteins from 4,566 genes.[1] The small genome size of *M. luteus* probably accounts for the ability of air isolates to rely on natural air flow as a form of transport because it allows them to maintain a lighter physique

The research for this article was conducted in the *Shuyuan* Molecular Biology Laboratory and a poster about this work was prepared for the ASM Microbe Conference in 2020.

[1] *Micrococcus luteus* NCTC 2665: NCBI Reference Sequence NC_012803.1
Escherichia coli str. K-12 substr. MG1655: NCBI Reference Sequence NC_000913.3

in comparison to regular-sized bacteria – most of which (~60%) have genomes greater than 3.0 Mb (diCenzo and Finan, 2017).

Given its small genome size and the likelihood that a large proportion of *M. luteus* is airborne, this bacterium must be quick to adapt to a new location when it lands – perhaps by rapidly acquiring genes from its new surroundings.

In this work, two isolates of *Micrococcus luteus* were retrieved from the air inside a classroom and a third isolate was recovered from the soil. A comparison of their genomes may help to reveal something about this organism's versatility.

2. Methods

Two isolates of *M. luteus* were recovered from indoor air at a Hong Kong school using the IUL Spin Air basic sampler passing 100 L/min for 5 minutes; a third isolate was recovered from nearby soil after suspension in PBS. Selected colonies were passaged ten times on LB agar. DNA was subsequently extracted using a PureLink® Genomic DNA Mini Kit. Shotgun sequencing was performed using the Illumina MiSeq NGS platform and draft genomes were assembled using Prinseq Lite and Newbler.

Antibiotic susceptibility was determined by the Kirby-Bauer disk diffusion method. NCBI BLAST (Johnson *et al*., 2008) was used to determine homology, identity, protein sequences, and alignment of isolates and plasmid sequences. KASS was used to map KEGG pathways (Kanehisa and Goto, 2000) and autoMLST used to investigate phylogeny (Alanjary *et al*., 2019). Resistance genes were found using CARD (Alcock *et al*., 2020) and plasmids were identified using PLSDB (Galata *et al.,* 2019). Genome annotation and comparative gene analysis was conducted using PATRIC (Wattam *et al*., 2017).

3. Results

3.1 Genomic Assembly Table

	Isolate 19	**Isolate 21**	**Isolate 22**
No. of Reads	709,308	512,127	253,045
Peak Depth	79	56	22
No. of Large Contigs	335	356	471
No. of Bases	2,526,961	2,529,469	2,546,006
No. of All Contigs	411	415	8,592
N50 Contig Size	12,050	10,761	529
Genome Size	2.5 MB	2.7 MB	3.4 MB
Homology	*Micrococcus luteus*	*Micrococcus luteus*	*Micrococcus luteus*
Identity	100%	100%	78%

3.2 Phylogenetics

Figure 1. A phylogenetic tree (PATRIC) places air isolates closer to the human skin isolate M. luteus SK58

The phylogenetic tree (Figure 1) distinctly shows a closer clustering of the *M. luteus* strains 19 and 21, positioned on a closer branch than isolate 22. This finding may reflect the adaptation of genomes and resultant phenotypes depending on the external conditions. All three isolates share a branch with *M. luteus* strains SK58 and O'Kane, both of which have been isolated from human skin (Timm *et al*., 2020; Hanafy *et al*., 2016), although isolates 19 and 21 have a closer phylogenetic relationship with these two strains than isolate 22.

It is notable that the soil isolate has an estimated genome size of 3.4 Mb, which is around 30% larger than the airborne isolates, which have genome sizes of 2.5 Mb and 2.7 Mb. A difference of 700-900 Kb may indicate the presence of at least one additional plasmid.[3]

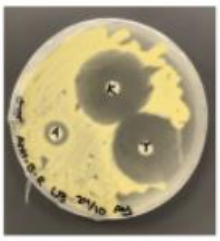


Figure 2

Figure 3

Figure 4

Figure 2.,3.,4. Agar diffusion essay has determined antibiotic resistance to ampicillin and susceptibility to kanamycin and tetracycline in isolates 19, 21, and 22.

3.3 Antibiotic Resistance

Agar diffusion assays showed that all three *M. luteus* isolates are susceptible to kanamycin and tetracycline but are resistant to ampicillin. Ampicillin is normally more active against Gram-positive species (*e.g.* Goldstein *et al.*, 2005), and this suggests that these strains of *M. luteus* are able to deactivate it.

Resistance to ampicillin, a penicillin-type antibiotic, is likely to be due to the beta-lactam resistance pathway identified in each isolate by KEGG. Ampicillin can be deactivated by beta-lactamase, an enzyme that cleaves the beta-lactam ring. CARD results have identified a number of beta-lactamase genes in each of the isolates. One hit of the ARO *NmcR* and *GOB-18* gene in each air isolate, while two hits of the NmcR and one hit of a *GOB-18* and *GOB-7* gene were found in the soil isolate. The *ARO NmcR* gene is a LysR regulator which contributes to the regulation of NmcA beta-lactamase. The *GOB-18* gene is a member of the subclass B3 GOB family of beta-lactamases, which is fully active against a broad range of beta-lactam substrates using a single Zn (II) ion in its active site, while the *GOB-7* gene is a class B beta-lactamase gene. These genes allow all three isolates to have a resistance mechanism against beta-lactam antibiotics.

3.4 Caseinolytic Activity

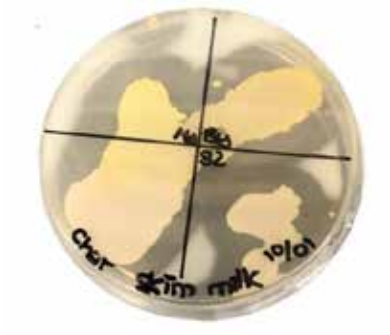


Figure 5. Skim milk plate testing demonstrates all isolates abilities to digest casein, leaving a clear LB plate with microbial growth.

All three isolates excrete an extracellular protease that digests casein. Skim milk agar plate testing has shown the growth of the isolates and production of casease. BLAST results have found the S8 serine protease, *vpr*, containing the same MKAEA signal peptide.

3.5 Identification of Plasmids

PLSDB identified 5 plasmids in isolate 19, 3 in isolate 21, and 7 in isolate S2, all of which match either 90% or above. The plasmid closest in sequence is compared to the three isolates below.

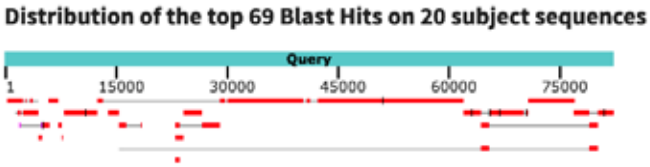


Figure 6. NCBI aligns the *Micrococcus sp*. A7 plasmid pLMA7 sequence with isolate 19, and shows a fragmented pattern of alignment.

Figure 7. Snapgene maps the regions of possible areas to cut the *Micrococcus sp*. A7 plasmid pLMA7 sequence in order to be manipulated for further tests.

Isolate 19: *Micrococcus sp. A7 plasmid pLMA7* is the closest match, with a size of 86 KB. This linear plasmid was detected in 2016 (Dib *et al*., 2010), with alignment mainly occurring in the 30,000 and 60,000 base pair regions.

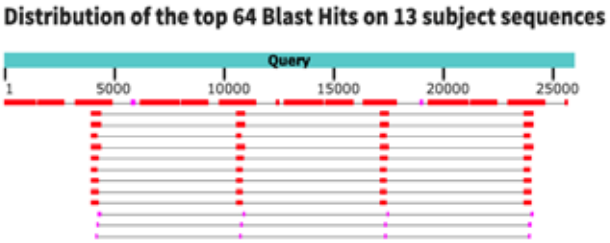


Figure 8. NCBI aligns the Arthobacter crystallopoietes strain DSM 20117 sequence with isolate 21, and shows a uniform pattern of alignment.

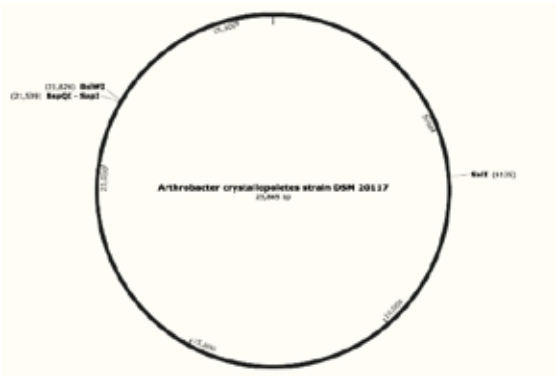


Figure 9. Snapgene maps the locations where the *Arthrobacter crystallopoietes* strain DSM 20117 sequences could be separated for further analysis.

Isolate 21: *Arthrobacter crystallopoietes strain DSM 20117* is the closest match, with a size of 29 KB. This circular plasmid was detected in 2018 in the USA (Romaniuk *et al*., 2018), with a consistent pattern of alignment.

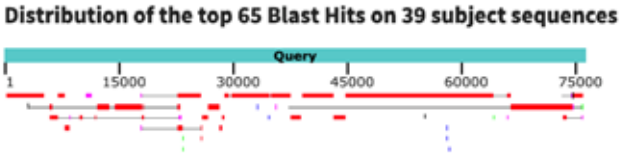


Figure 10. NCBI aligns the *Micrococcus sp.* MG-2010-D12 sequence with isolate 22, and like isolate 19, shows a fragmented pattern of alignment.

Figure 11. Snapgee indicates areas where the *Micrococcus sp*. MG-2010-D12 sequence could be slashed and isolated for future research.

Isolate 22: *Micrococcus sp. MG-2010-D12 plasmid pJD1* is the closest match, with a size of 79 KB. This linear plasmid was detected in 2015 in Argentina (Dib, *et al*., 2015).

These findings suggest that although all isolates have obtained plasmids similar in genomic composition to the matched sequences identified by PLSDB (Galata *et al*., 2019), the fragmented alignment occurring between Isolate 19 and plasmid pLMA7 as well as Isolate 22 and plasmid pJD1 suggests that the plasmids may be rather novel and unfamiliar. On the other hand, the regular alignment of Isolate 21 and the *Arthrobacter* DSM 20117 plasmid suggest the opposite.

3.6 Gene +6 Trait Comparison

Genes Repeated in PATRIC analysis

19	*GdpD*	Glycerophosphoryl diester phosphodiesterase
21	*dxr*	1-deoxy-D-xylulose 5-phosphate reductoisomerase
S2	*folP*	Dihydropteroate synthase
	rho	Transcription termination factor Rho

Traits Unique to the Isolate

19	*J*	Phage tail tip, host specificity protein J
21	*AMT*	Aminomethyltransferase
S2	*R*	Phage endopeptidase Rz
	R	Phage lysozyme R

The *GdpD* gene relates to lipid-related chemical reactions and pathways. This may facilitate the isolate in its cell membrane structure, energy storage, and function as signaling molecules (Shi *et al*., 2008).

The *dxr* gene is responsible for the catalysis in the first step of the MEP pathway in isoprenoid biosynthesis, the synthesis of ether-linked prenyl-lipids which constitute the isolate's plasmid membrane (Lange, 2000).

The *folP* gene is responsible for catalysing the condensation of pABA with DHPt-PP to form H2Pte (The Uniprot Consortium, 2019). In other words, the production of vitamin B, as well as resistance against sulfonamide antibiotics.

The *Rho* gene is a crucial transcription protein in prokaryotic organisms, responsible for marking a section of nucleic acid sequences at the end of a gene in genomic DNA during transcription (The Uniprot Consortium, 2019).

The phage tail tip with host specificity to protein J facilitates the attachment of virion to the host's LamB receptors, a family of outer membrane proteins.

The hypothetical, supposedly aminomethyl-transferase (AMT), protein is a gene that is responsible for encoding one of the four critical components of the glycine cleavage system, a major pathway responsible for catalysing glycine degradation (Kikuchi *et al*., 2008) .

The R gene referring to Phage endopeptidase Rz is responsible for the rupture of cell membranes and loss of cytoplasm, which may allow for bacteriophage to enter the isolate (Chatterjee and Rothenberg, 2012).

The R gene referring to Phage lysozyme R facilitates the peptidoglycan catabolic process, a component of the bacterial cell wall that stimulates immune responses (The Uniprot Consortium, 2019).

The identification of phages in isolate 22 may indicate the possibility of gene transfer, which is highly likely to occur in soil environments. In this case, the phage tail tip in isolate 19 may have previously been acquired when landing on a soil surface.

3.7 Antibiotic Resistance Genes in Plasmids

Isolate 19	*oleB*	Antibiotic Target Protection
	tet(V)	Antibiotic Efflux
Isolate 19, 22 and S2	*qaCA*	Antibiotic Efflux
	acrR	Antibiotic Target Alteration, Antibiotic Efflux

qaCA is a subsection of the *qac* multidrug efflux pump. This function mediates resistance towards a broad spectrum of monovalent and divalent antimicrobial cations, and primarily targets fluoroquinolone antibiotics (Mitchell *et al*., 1998).

TetV is a tetracycline efflux protein, meaning that it catalyses the efflux of metal-tetracycline complex in exchange for protons. This resistance gene specifically targets tetracycline antibiotics (Krulwich *et al*., 2001).

oleB codes for an ABC-F subfamily protein in *Streptomyces antibioticus* that directly affects oleandomycin secretion, a macrolide antibiotic that targets and reversibly binds to the 50S subunit of bacterial ribosomes, preventing translocation of peptidyl tRNA, inhibiting protein synthesis. Although this protein mainly targets macrolide antibiotics, it affects a range of others including lincosamide and streptogramin antibiotics (Champney *et al*., 1998).

AcrR is a repressor of the AcrAB-TolC multidrug efflux complex. AcrR mutations result in high level antibiotic resistance (Du *et al*., 2014). This gene also may target a number of antibiotics such as fluoroquinolone antibiotics, cephalosporin, and glycylcycline.

Conclusion

Certain genomic traits remain consistent no matter the source. This may be due to them aiding survival in a range of environments.

However, *M. luteus* may also obtain different plasmids depending on their surroundings. Plasmids allow for gene transfer to occur between microbes more efficiently, which may assist mechanisms necessary to the isolate's survival. Complete genome sequencing will help to fully characterise the plasmids to understand their role in the flexibility of this highly adaptable microbe.

Reference

Alanjary, M., Steinke, K., Ziemert, N. (2019) AutoMLST: an automated web server for generating multi-locus species trees highlighting natural product potential. *Nucleic Acids Research*, 47(W1), W276–W282. doi: 10.1093/nar/gkz282 URL:automlst.ziemertlab.com/.

Alcock, B.P., Raphenya, A.R., Lau, T.T.Y., *et al*. (2020) CARD 2020: Antibiotic resistome surveillance with the Comprehensive Antibiotic Resistance Database, *Nucleic Acids Research,* vol. 48, D517-D525 doi: 10.1093/nar/gkz935 URL: card.mcmaster.ca/

Cabaj, A., Kosakowska, A. (2009) Iron-dependent growth of and siderophore production by two heterotrophic bacteria isolated from brackish water of the southern Baltic Sea. *Microbiological research,* 164(5), 570-577. doi: 10.1016/j.micres.2007.07.001

Champney, W.S., Tober, C.L., Burdine, R. (1998) A comparison of the inhibition of translation and 50S ribosomal subunit formation in *Staphylococcus aureus* cells by nine different macrolide antibiotics. *Current Microbiology*, 37(6), 412–417. doi: 10.1007/s002849900402. [Oleandomycin CAS 3922-90-5]

Chatterjee, S., Rothenberg, E. (2012) Interaction of Bacteriophage λ with Its *E. Coli* Receptor, LamB. *Viruses*, 4(11), 3162-3178, doi: 10.3390/v4113162

Dib, J. R., Angelov, A., Liebl, W., Döbber, J., Voget, S., Schuldes, J., Gorriti, M., Farías, M. E., Meinhardt, F., Daniel, R. (2015) Complete Genome Sequence of the Linear Plasmid pJD12 Hosted by *Micrococcus sp*. D12, Isolated from a High-Altitude Volcanic Lake in Argentina. *Genome Announcements*, 3(3), e00627-15. doi: 10.1128/genomeA.00627-15

diCenzo, G.C., Finan, T.M. (2017) The Divided Bacterial Genome: Structure, Function, and Evolution. *Microbiology and Molecular Biology Reviews*, 81(3): e00019-17. doi: 10.1128/MMBR.00019-17

Du, D., Wang, Z., James, N.R., Voss, J.E., Klimont, E., Ohene-Agyei, T., Venter, H., Chiu, W., Luisi, B.F. (2014) Structure of the AcrAB–TolC multidrug efflux pump. *Nature,* 509(7501), 512-515, doi: 10.1038/nature13205

Galata, V., Fehlmann, T., Backesand, C., Keller, A. (2019) PLSDB: a resource of complete bacterial plasmids, *Nucleic Acids Research*, 2019, 47, D195–D202. doi: 10.1093/nar/gky1050. URL:ccb-microbe.cs.uni-saarland.de/plsdb/plasmids/search_form/seq/

Goldstein, E. J., Citron, D. M., Merriam, C. V., Warren, Y. A., Tyrrell, K. L., Fernandez, H. T., Bryskier, A. (2005). Comparative *in vitro* activities of XRP 2868, pristinamycin, quinupristin-dalfopristin, vancomycin, daptomycin, linezolid, clarithromycin, telithromycin, clindamycin, and *ampicillin against anaerobic gram-positive species, actinomycetes, and lactobacilli. Antimicrobial Agents and Chemotherapy,* 49(1), 408–413. doi: 10.1128/AAC.49.1.408-413.2005

Hanafy, R. A., Couger, M. B., Baker, K., Murphy, C., O'Kane, S. D., Budd, C., French, D. P., Hoff, W. D., Youssef, N. (2016). Draft genome sequence of *Micrococcus luteus* strain O'Kane implicates metabolic versatility and the potential to degrade polyhydroxybutyrates. *Genomics Data*, 9, 148-153. doi: 10.1016/j.gdata.2016.08.006

Harrison, E., Brockhurst, M. A. (2012). Plasmid-mediated horizontal gene transfer is a coevolutionary process. Trends in *Microbiology*, 20(6), 262-267. doi: 10.1016/j.tim.2012.04.003

Jampasri, K., Pokethitiyook, P., Poolpak, T., Kruatrachue, M., Ounjai, P., Kumsopa, A. (2020). Bacteria-assisted phytoremediation of fuel oil and lead co-contaminated soil in the salt-stressed condition by *Chromolaena odorata and Micrococcus luteus. International Journal of Phytoremediation,* 22(3), 322–333. doi: 10.1080/15226514.2019.1663482

Johnson et *al*. NCBI Resource Coordinators (2008) Database resources of the National Center for Biotechnology Information, Nucleic Acids Research, 44(D1), D7–D19. doi: 10.1093/nar/gkv1290. URL:blast.ncbi.nlm.nih.gov/Blast.cgi.

Kanehisa, M., Goto, S. (2000) KEGG: kyoto encyclopedia of genes and genomes, *Nucleic Acids Research,* 28(1):- 27-30. doi: 10.1093/nar/28.1.27 URL: http://www.genome.jp/kaas-bin/kaas_main

Kikuchi, G., Motokawa, Y., Yoshida, T., Hiraga, K. (2008) Glycine cleavage system: reaction mechanism, physiological significance, and hyperglycinemia. *Proceedings of the Japan Academy. Series B, Physical and biological sciences,* 84(7), 246–263. doi: 10.2183/pjab.84.246

Krulwich, TA *et al.* (2001) Functions of Tetracycline Efflux Proteins That Do Not Involve Tetracycline. *Journal of Molecular Microbiology and Biotechnology,* 84(7), 246-263. doi: 10.2183/pjab/84.246

Kutmutia, S.K., Drautz-Moses, D.I., Uchida, A., Purbojati, R.W., Wong, A., Kushwaha, K.K., Putra, A., Premkrishnan, B.N.V., Heinle, C.E., Vettath, V.K., Junqueira, A.C.M., Schuster, S.C. (2019) Complete Genome Sequence of *Micrococcus luteus* Strain SGAir0127, Isolated from Indoor Air Samples from Singapore. *Microbiology Resource Announcements*, 8(41): e00646-19 doi: 10.1128/mra.00646-19

Lange, M. (2000) Isoprenoid biosynthesis: The evolution of two ancient and distinct pathways across genomes. *PNAS*, 97(24), 13172-13177. doi: 10.1073/pnas.240454797.

Mitchell, B.A., Brown, M.H., Skurray, R.A. (1998) QacA Multidrug Efflux Pump from *Staphylococcus aureus:* Comparative Analysis of Resistance to Diamidines, Biguanidines, and Guanylhydrazones. *Antimicrobial Agents and Chemotherapy,* 42(2), 475-477 doi: 10.1128/AAC.42.2.475

Mosig, G., Calendar, R. (2002). Horizontal Gene Transfer in Bacteriophages. In Syvanen, M., Kado, C.I. (eds.), *Horizontal Gene Transfer* (2nd. Edn.), Academic Press; Ch. 13, pp. 141-146, VII-VIII. doi: 10.1016/b978-012680126-2/50017-7

NCBI Resource Coordinators (2016) Database resources of the National Center for Biotechnology Information, *Nucleic Acids Research*, 44(D1), D7–D19. doi: 10.1093/nar/gkv1290. URL: blast.ncbi.nlm.nih.gov/Blast.cgi.

Olowomofe, Temitayo, O, Oluyege, JO, Aderiye, BI, Oluwole, OA (2019) Aliphatic Hydrocarbon Profile of Crude-oil Degraded by Bacteria Isolates from Bitumen-polluted Surface Water from Agbabu, Ondo State, *Journal of Pure and Applied Microbiology,* 13(2), 879-888. doi: 10.22207/JPAM.13.2.23

Pandey, N., Cascella, M. (2019) *Beta-Lactam Antibiotics*. Bethesda, MD: National Center for Biotechnology Information Books. www.ncbi.nlm.nih.gov/books/NBK545311/.

Romaniuk, K., Golec, P., Dziewit, L. (2018) Insight Into the Diversity and Possible Role of Plasmids in the Adaptation of Psychrotolerant and Metalotolerant *Arthrobacter* spp. to Extreme Antarctic Environments. *Frontiers in Microbiology, 9,* 3144. doi: 10.3389/fmicb.2018.03144 [including data Table 6]

Shi, L., Liu, J.-F., An, X.-M., Liang, D.-C. (2008) Crystal structure of glycerophospho- diester phosphodiesterase (GDPD) from *Thermoanaerobacter tengcongensis,* a metal ion-dependent enzyme: Insight into the catalytic mechanism. Proteins, 72(1), 280-288. doi: 10.1002/prot.21921.

Souak, D., Boukerb, A.M., Barreau, M., Duclairoir-Poc, C., Feuilloley, M.G.J. (2020) Draft Genome Sequences of *Micrococcus luteus* MFP06 and MFP07, Isolated from the Skin of Healthy Volunteers, *Microbiology Resource Announcements,* 9(25): e00545-20. doi: 10.1128/MRA.00545-20

Sun, D. (2018). Pull in and Push Out: Mechanisms of Horizontal Gene Transfer in Bacteria. *Frontiers in Microbiology*, 9: 2154. doi: 10.3389/fmicb.2018.02154

The UniProt Consortium (2019) UniProt: a worldwide hub of protein knowledge, *Nucleic Acids Research*, 47: D506-515. doi: 10.1093/nar/gky1049. URL: www.uniprot.org/

Timm, C.M., Loomis, K., Stone, W., Mehoke, T., Brensinger, B., Pellicore, M., Staniczenko, P.P.A., Charles, C., Nyak, S., Karig, D.K. (2020) Isolation and characterization of diverse microbial representatives from the human skin microbiome. *Microbiome,* 8: 58. doi: 10.1186/s40168-020-00831-y

Tuleva, B., Christova, N., Cohen, R., Antonova, D., Todorov, T., Stoineva, I. (2009) Isolation and characterization of trehalose tetraester biosurfactants from a soil strain *Micrococcus luteus* BN56, *Process Biochemistry,* 44(2), 135-141. doi: 10.1016/j.procbio.2008.09.016

Wattam, A.R., Davis, J.J., Assaf, R., Boisvert, S., Brettin, T., Bun, C., Conrad, N., Dietrich, E. M., Disz, T., Gabbard, J.L., Gerdes, S., Henry, C.S., Kenyon, R.W., Machi, D., Mao, C., Nordberg, E.K., Olsen, G.J., Murphy-Olson, D.E., Olson, R., Overbeek, R., Parrello, B., Pusch, G.D., Shukla, M., Vonstein, V., Warren, A., Xia, F., Yoo, H., Stevens, R.L. (2017) Improvements to PATRIC, the all-bacterial Bioinformatics Database and Analysis Resource Center, *Nucleic Acids Research*, 45(D1), D535–D542. doi: 10.1093/nar/gkw1017 URL: www.patricbrc.org/.

Yavas, A., Icgen, B. (2018) Aerobic Bacterial Degraders With Their Relative Pathways for Efficient Removal of Individual BTEX Compounds, *Clean-Soil Air Water*, 46(11): 1800068. doi: 10.1002/clen.201800068

Zothanpuia, Passari, A.K., Leo, V.V., Chandra, P., Kumar, B., Nayak, C., Hashem, A., Abd-Allah, E.F., Alqarawi, A.A., Singh, B.P. (2018) Bioprospection of actinobacteria derived from freshwater sediments for their potential to produce antimicrobial compounds. *Microbial Cell Factories*, 17: 68 (2018). doi: 10.1186/s12934-018-0912-0

Benson, D. A., Cavanaugh, M., Clark, K., Karsch-Mizrachi, I., Lipman, D. J., Ostell, J., Sayers, E. W. (2013). GenBank. *Nucleic Acids Research*, 41 (Database issue), D36–D42. doi: 10.1093/nar/gks1195.

Dib, J. R., Liebl, W., Wagenknecht, M., Farías, M. E., Meinhardt, F. (2013) Extrachromosomal genetic elements in *Micrococcus. Applied Microbiology and Biotechnology*, 97(1), 63–75. doi: 10.1007/s00253-012-4539-5

Kooken, Jennifer *et al*. (2012) Characterization of *Micrococcus* strains isolated from indoor air. *Molecular and Cellular Probes,* 26(1), 1-5, doi: 10.1016/j.mcp.2011.09.003.

Madupu, R., Durkin, A.S., Torralba, M., Methe, B., Sutton, G.G., Strausberg, R.L., Nelson, K.E. (2009) Glycerophosphodiester phosphodiesterase family protein from *Micrococcus luteus* SK58, GenBank: EFD50857.1

Pepper, I.L., Gentry, T.J. (2015) Earth Environments. In Pepper, I.L., Gerba C.P., Gentry, T.J. (Eds.), *Environmental Microbiology* (Third Edition), Academic Press; Ch.4, pp. 59-88. doi: 10.1016/B978-0-12-394626-3.00004-1

Rakhashiya, Purvi *et al*. (2015) Whole-genome sequences and annotation of *Micrococcus luteus* SUBG006, a novel phytopathogen of mango. *Genomics Dat*a, 6(C), 10-11. doi: 10.1016/j.gdata.2015.07.022

SnapGene (2004) GSL Biotech LLC: SnapGene. San Diego, CA: Graphpad Software. www.snapgene.com/

A Short Introduction to Particle Physics

Jingxin Crystal Deng

A Short Introduction to Particle Physics is a brief, 50-page bilingual book that covers the basic theories of modern Particle Physics and Quantum Physics. The book includes a variety of images and diagrams that help illustrate and clarify the concepts covered in the book. The book is a medium used to educate younger students by presenting the basics of Particle Physics in a straightforward, comprehendible fashion. It aims to simplify Particle Physics, which is traditionally considered a complex discipline, in order to make it accessible to more readers. Furthermore, it aims to encourage younger audiences to potentially harbour interest in Particle Physics; in light of the current scientific research trends, this may prove to be valuable.

《探索粒子物理》是一本精簡的 50 頁雙語作品，涵蓋了現代粒子物理和量子物理的基本理論。本文利用各種圖像和圖表，以幫助說明和闡明文中涵蓋的概念。傳統上，粒子物理學被視為一門複雜的學科，本文以簡單易懂的方式介紹粒子物理學的基礎知識，使更多的讀者可以理解箇中概念，並成為年輕學子學習粒子物理學的其中一個媒介。此外，本文旨在鼓勵年輕讀者對粒子物理學產生興趣，以協力推進當今科學研究趨勢。

This article was published as a book in fulfillment of the IB's MYP Grade 10 Personal Project.

Introduction

"Particle physics". These simple words might seem very daunting to you. Indeed, there are adults who don't have the slightest idea what quarks are, or how multiple dimensions may exist. But really, particle physics is only confusing because scientists use complicated math equations and abstract ideas to explain it. Why don't we try a different approach? The concepts for particle physics can be easily explained using analogies and colourful graphics. Even young students can dive into the universe at the smallest scale. Still in doubt? Don't worry. Read this very short book about the basics of particle physics, and perhaps you would change your mind.

前言

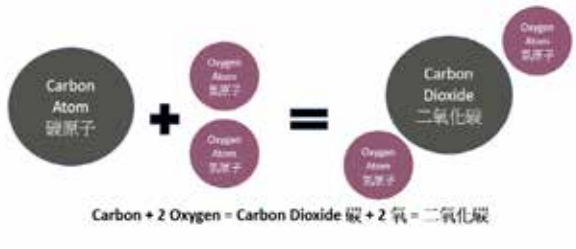
粒子物理。你可能覺得這四個字非常嚇人。的確，有很多成人可能連“夸克”這個詞都沒聽過、完全不能理解額外維度為何可能存在。其實，我們對粒子物理感到困惑只是因為它有很多複雜的數學公式和抽象的概念。為何我們不能走另外一條路，繞過這些可怕的解釋呢 ？我們可以使用類比和圖表來描述粒子物理中的概念，這樣一來，年幼的學生們也能探索粒子物理世界。你仍畏懼嗎 ？別擔心。你可能讀完這本非常簡短的書之後就變成粒子物理的朋友了。

Atoms

If you cut along the dotted white lines, you will realize that the paper quickly becomes very small – but how small can it go? Can you keep cutting it forever?

Imagine a piece of paper being cut in half. Then in half again. And again. If you keep doing this, you will realize that you won't be able to cut the paper into smaller pieces anymore, yet the paper still exists. Like the paper, everything in the world – whether snow, trees, keys, or you and me – can be divided into smaller and smaller pieces, until a point where they can't divide anymore. You might already know this smallest piece to be an "atom". They form larger objects.

Naturally, there are different types of atoms. For example, oxygen is made of oxygen atoms, and carbon is made of carbon atoms. The same type of atoms form pure elements, for example oxygen atoms make up pure oxygen. Different types of atoms can also react with each other to form new things. For example, one carbon atom and two oxygen atoms can bond to form carbon dioxide, a gas that we exhale.



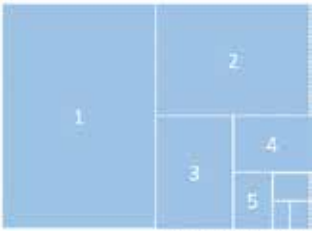
Periodic Table

Over 100 elements have been discovered by the late 19th century, and each element is made of the same type of atoms. Dmitri Mendeleev, a Russian chemist, grouped these elements into a table based on their characteristics. This table is called the "Periodic Table".

Subatomic Particles

Well, the idea that atoms cannot be divided anymore is actually wrong. It was popular belief – until scientists discovered that there are particles smaller than the atom around the early 20th century. We call them subatomic articles. Particle physics – unlike chemistry – is more about those subatomic particles than atoms, elements or the Periodic Table. We will now officially begin our subatomic journey into the world of particle physics.

原子



如果你沿著白線一直剪下去，你很快就會發現這張紙立刻變小。但它到底能被剪到多小呢 ？你能永遠剪下去嗎 ？

想像一張紙，你將它剪成一半，然後再剪成一半之後再剪一次……若你不斷這樣做，那麼紙片馬上就會小得連剪都剪不了了。即便紙片不能再變小了，但是它仍然存在。就像紙一樣，世上萬物（不管是雪花、樹木還是你我）都能被分解成很小很小的碎片，直到再也不能縮小了。這個最小的碎片就是“原子”。原子組成所有的物體。

當然，世上有很多不同類型的原子。我們呼吸的氧氣就是由氧原子構成的，而碳是由碳原子構成的。相同類型的原子能構成元素，就像氧原子構成純氧一樣。不同類型的原子也可以發生反應產生新的物質。比如，一個碳原子和兩個氧原子能構成二氧化碳。

元素週期表

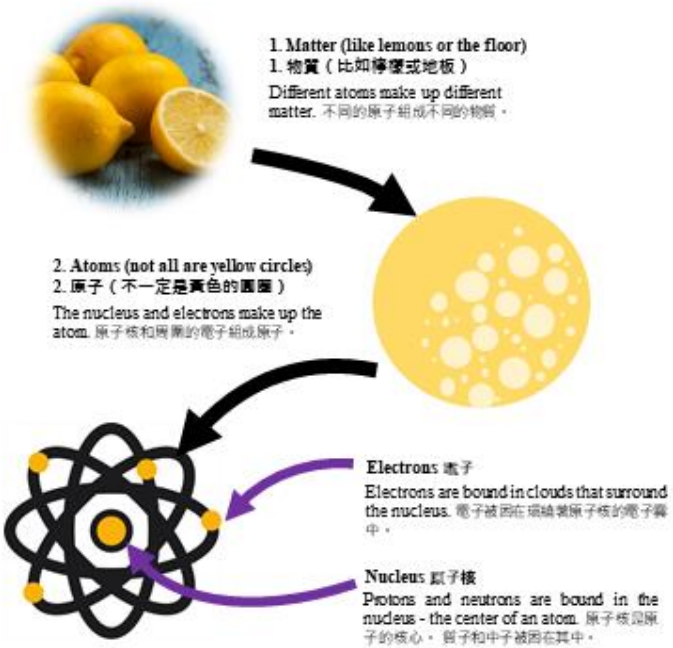
十九世紀末，超過一百個元素已被發現。每一個元素都是由那一類的原子構成的。一位名叫門捷列夫的俄羅斯化學家依照這些元素的性質把它們歸納到一個表中, 這個表叫“元素週期表”。

次原子粒子

其實，原子並不是組成物質最小的單元。科學家們在二十世紀初發現它們是由更小的粒子組成的，我們把這些比原子還小的粒子叫做次原子粒子。粒子物理更加注重這些次原子粒子，而不是更大的原子、元素或元素週期表。因此，這本書也不會多講這方面的知識。現在，我們正式開始探索次原子粒子和粒子物理世界吧。

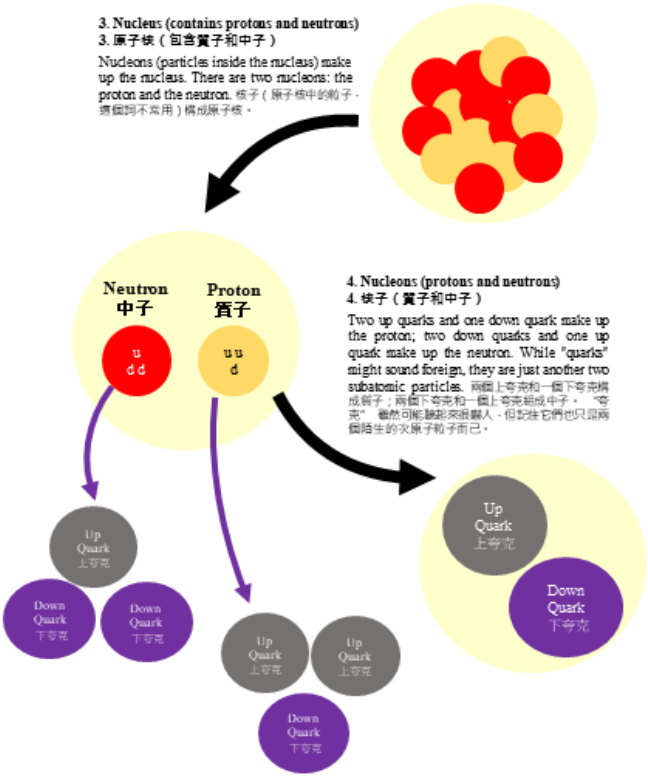
In The Atom

So what exactly are the particles that make up the tiny atom? You may already know that the atom has a nucleus, and particles called electrons surround the nucleus. The nucleus is made of two kinds of particles: protons and neutrons. And yet, even these two are not the smallest. Protons and neutrons are both made of quarks. Do you know what quarks are? We will get to them later.



原子內

那麼，小小的原子到底是由什麼粒子組成的呢？你可能已經知道原子有一個原子核，而且一種名叫電子的粒子包圍原子核。原子核是由兩種粒子構成的：質子和中子。但是，這兩個粒子都不是最小的，他們都由夸克組成。若你不知道夸克是什麼，別擔心，我們馬上就會認識它們了



5. Quarks (not only up and down)

Fundamental Particles

Some of the subatomic particles mentioned above (the electron, the up quark and the down quark) are indivisible. This means that they are the basic building blocks of the world. Particles that cannot be divided anymore are called “fundamental particles”. So the atom is not a fundamental particle, while the electron and the quarks are examples of them. There are many more particles like them.

5. 夸克（除了上和下夸克還有很多）

基本粒子

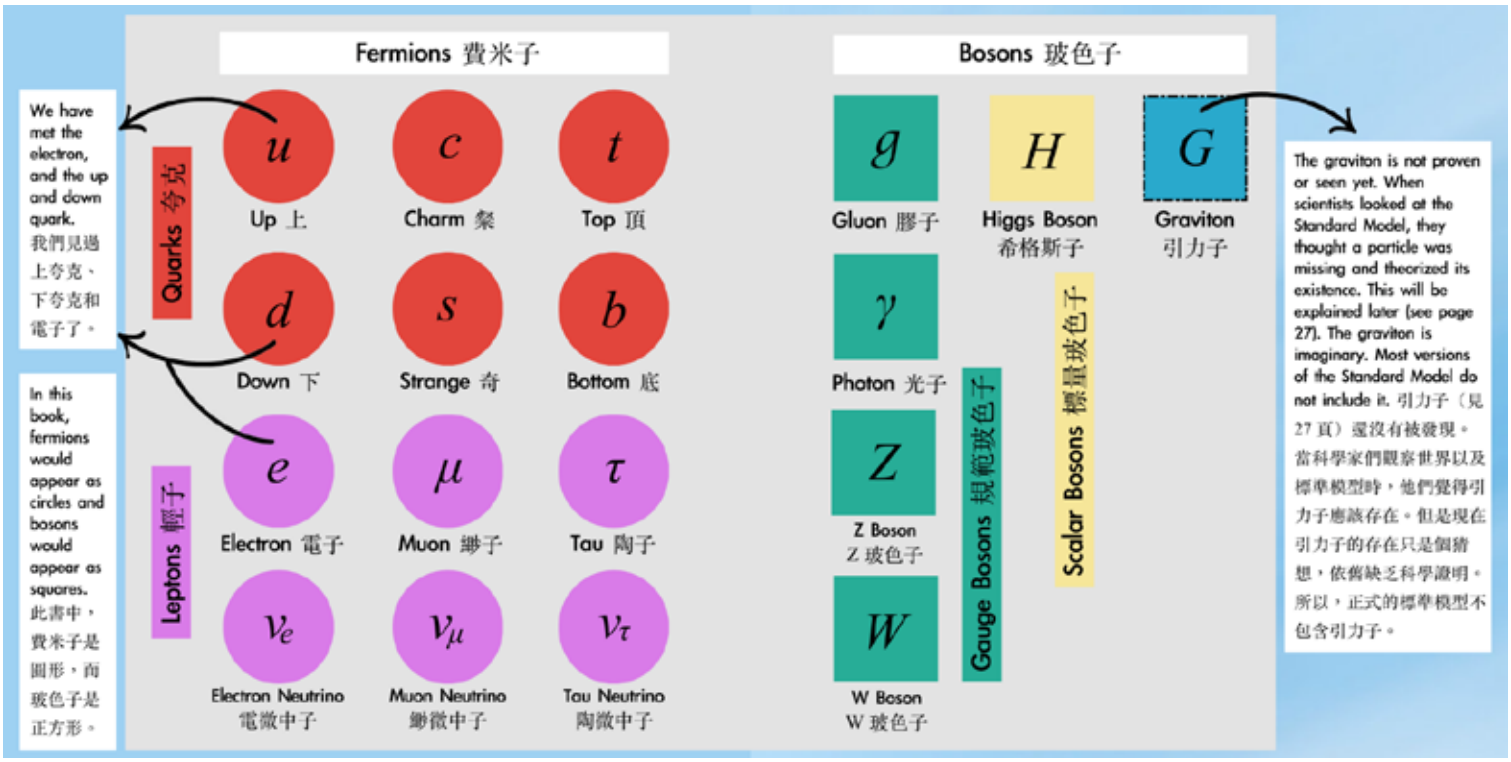
剛剛提到的一些次原子粒子（電子、上夸克和下夸克）都不可以再分解了，這表示它們是最基礎的粒子，構成世上萬物。這種不能再分的粒子叫做“基本粒子”。所以，原子不是基本粒子，而電子和夸克卻是基本粒子之一。

The Standard Model

Just like how all the elements are grouped into a periodic table, all the fundamental particles are grouped into the "Standard Model". The Standard Model of Particle Physics tells us the properties and connections between 17 *known* particles, which are (as far as we know) fundamental. It also tells us how particles that make up matter (fermions) interact with the particles that carry forces (bosons). Right now, it is the most accurate and neat representation we have of the subatomic world. All subatomic particles in the Standard Model will be explained.

標準模型

若元素週期表包含了所有的元素，那麼有沒有一個表格包含所有的基本粒子呢？當然。下面這個表格名叫“標準模型”，涵蓋了所有的基本粒子。“標準模型”告訴我們各個基本粒子的性質和粒子之間的聯繫，它一共有十七個已知的粒子和一個假象的引力子。標準模型也能描述費米子（構成物質的粒子）和玻色子（傳遞力的粒子）之間的聯繫。但是，它並不能告訴我們基本粒子內有什麼東西（我們現在不知道有沒有）、各個粒子的大小、玻色子傳遞的力有多大。但是現在，標準模型能夠最可靠地、最有效地描述粒子物理界的種種聯繫。



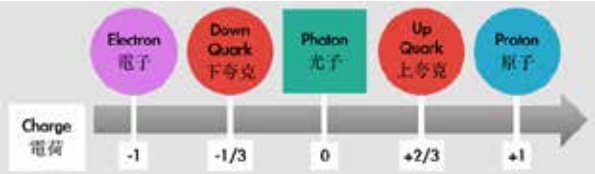
The forces carried by the bosons are not included in the model above, but we will get to them soon. Don't worry if you have no idea what fermions or bosons or leptons are. All will be explained in good time.

上面的標準模型並不包含玻色子傳遞的力，但是你馬上就會學習它們了。若“費米子”、“玻色子”或“輕子”這些詞語對你來說都是完全陌生的，這很正常。大部分的人都沒聽過。這些詞語都會稍後出現。

Electric Charge

Electric charge is a property of a particle that does not change. You've probably all played with magnets before. Hold the North pole and the South pole close enough, and you will feel an invisible force tugging them closer together. And put North and North together, you feel a repelling force. But if you hold the two magnets far enough apart, nothing happens. Electric charge is very similar. Particles with charge have electric fields (an area that influences other particles), and when other particles enter that field, the particles will interact. But once the particle is far enough away from the field, nothing happens.

Not all fundamental particles have whole number electric charges like electrons. The electric charge of quarks are all fractions. Some particles are electrically neutral (their charge is 0).



Spin

Imagine a magnet entering a strong magnetic field. The magnet will change in direction, or "deflect". When scientists sent electrically charged particles (for example electrons) through a magnetic field, they found that the particles deflected, like tiny magnets. They deflected in a way that's very similar to how charged spinning balls would. But this analogy is not accurate, because a fundamental particle is so small that it is more like a point than a ball. And a point can't spin.

The spin of a charged particle is a fixed value. Similar to electric charge, spin is also an inner property of a particle. Different particles have different spins. It turns out spin actually determines if any particle is a fermion or a boson. If the spin is a multiple of 1/2, the particle is a fermion. If the spin is a multiple of 1 (including 0), the particle is a boson. Spin, like charge, can be negative.

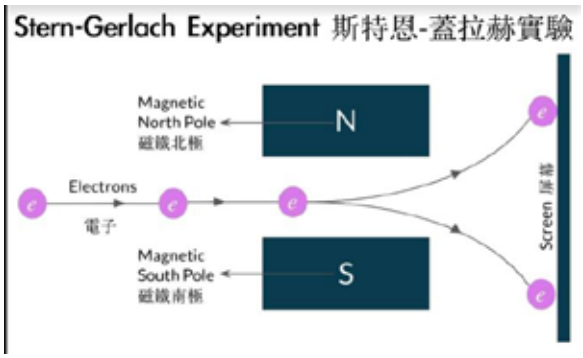
電荷

電荷是粒子的不會改變的性質，是個內在的性質。你玩過磁鐵嗎？當你把一個磁鐵的北極和另一個磁鐵的南極放得比較近時，你會感受到一個隱形的力將磁鐵吸在一起。相反，把兩個北極靠近時，會有一個排斥的力。但若兩個磁鐵相隔太遠，好似沒有力產生。相似地，帶有電荷的粒子有一個電場。當另一個粒子進入這個場，兩個粒子就會互動。但若那個粒子沒有進入電場，什麼事都不會發生。

很多基本粒子的電荷都不是整數。夸克的電荷都是分數。有些粒子是電中性（電荷為零）的。

自旋

想像一個磁鐵穿過一個磁場，那個磁鐵會改變方向，會偏移原來的軌道。當科學家們把帶有電荷的粒子（比如電子）射入磁場時，他們發現那些粒子也偏離了原來的方向，就像小磁鐵一樣。它們表現得像自己旋轉的小球，因此科學家們將這個性質叫做“自旋”。但是旋轉的小球並不是個恰當的比喻，因為基本粒子太小了——它們小到更像一個點，而點是不能旋轉的。

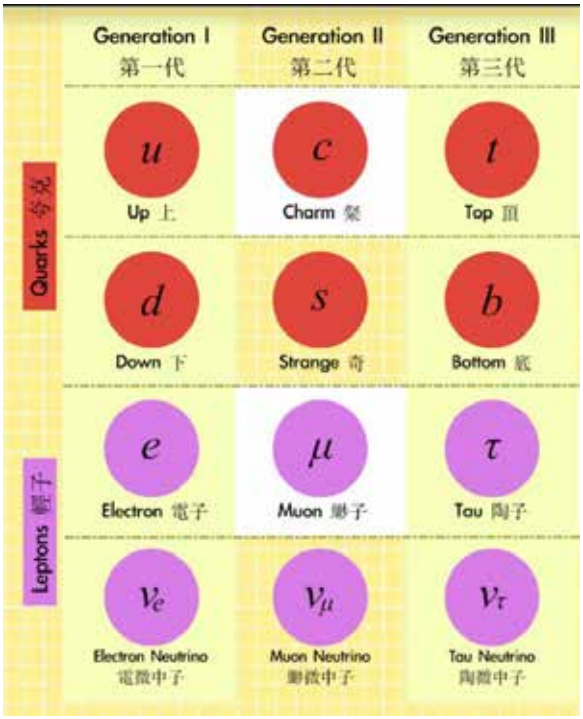


The charged electrons (written as "e" in the diagram) either move up or down after passing through a magnetic field. There are no results between. 帶電荷的粒子要麼偏到上方，要麼偏到下方，沒有中間的結果。

一個粒子的自旋是個固定的數字。自旋與電荷相似，也是一個內在的、不隨環境改變的性質。不同的粒子有不同的自旋，而自旋決定一個粒子是費米子還是玻色子。若自旋是 1/2 的倍數，粒子是費米子；若自旋是 1 的倍數（可以是 0），粒子是玻色子。自旋跟電荷一樣，可以是負數。

Fermions

The left side of the Standard Model shows fermions. Fermions are particles that make up matter. All fundamental fermions have a spin of 1/2. We have seen the up quark and down quark, which make up the proton and neutron. There's also the electron and the electron neutrino. Together, these four fundamental fermions make up generation 1. Strangely, there are also particles of generation 2 and 3, which are just heavier copies of generation 1 and otherwise identical. The table below shows all the fundamental fermions, which is half the Standard Model.



Fermions (can be fundamental where the spin is 1/2, or composite) are defined as particles with spin that's a multiple of 1/2. A particle can be a fermion even if it is made of various fundamental fermions, as long as it has a spin that's a multiple of 1/2. So, fundamental fermions can group together to form composite fermions. A proton is made of two up quarks and one down quark, which all have 1/2 spin. So the proton has 1/2 + 1/2 + 1/2 = 3/2 spin. Since 3/2 is also a multiple of 1/2, protons are also fermions.

The particles of generation 2 and 3 only exist in high energies, such as the early stages of the universe. The heavier versions of the fundamental fermions can also form matter, but they are very unstable. They quickly decay (see page 25) into the lighter fermions: up quarks, down quarks, electrons and electron neutrinos. And that is why the particles of generation 1 form 99.9% of all matter.

費米子

標準模型的左側是費米子。費米子是構成物質的粒子，所有基本費米子的自旋都是 1/2。我們認識上夸克和下夸克，它們組成質子和中子。最左側還有電子和電微中子。這四個基本費米子組成第一代費米子。奇怪的是，自然中還有第二代和第三代粒子，它們只是比第一代費米子重，別無不同。

Up-type quarks 上型夸克
Electric charge:+2/3 電荷：+2/3

The charm quark and the top quark are heavier than the up quark. The three are collectively called "up-type quarks".粲夸克和頂夸克都比上夸克重。這三個夸克被稱為“上型夸克”。

Down-type quarks 下型夸克
Electric charge: -1/3 電荷：-1/3

The strange quark and the bottom quark heavier than the down quark. The three are up quarks, down quarks, called "down-type quarks".奇夸克和底夸克都比下夸克重。這三個夸克被稱為“下型夸克”。

Charged leptons 帶電輕子
Electric charge: -1 電荷： -1

The muon and the tau are simply heavier than the electron. They all have -1 charge, thus called "charged leptons".緲子和陶子比電子重。它們的電荷都是-1，因此被稱為“帶電輕子”

Neutrinos 中微子
Electric charge: 0 電荷： 0

The electron, muon and tau neutrino are electrically neutral and very light in mass, thus called "neutrinos".中微子是微小的電中性粒子。電微、緲微和陶微中子的電荷都是 0，且質量很小。

費米子的定義是“自旋為 1/2 倍數的粒子”。所以，即便一個粒子是由多個基本費米子構成的，它的自旋只要是 1/2 的倍數就是費米子。基本費米子能組合，構成複合費米子。質子是由兩個上夸克和一個下夸克構成的，所以它的自旋是 1/2+1/2+1/2=3/2。而因為 3/2 是 1/2 的倍數，質子也是費米子。

第二代和第三代的夸克和輕子只在高能量的環境中存在，比如宇宙誕生階段。較重的基本費米子（第二和第三代）也能構成物質，但是它們都太不穩定了，很快就會衰（見25頁）變成最輕的費米子，即上下夸克、電子、電微中子。這就是為什麼 99.9%的物質都是由第一代費米子構成的。

The Fundamental Forces of Nature

What use is a set of matter particles when they don't know how to interact and actually build any matter? It's like having a bunch of Lego blocks but no instruction manual to follow. This is where the four fundamental forces of nature comes in. They act as guidelines and tell the fermions how to interact to construct matter. "Fundamental forces" might seem daunting, but you may already know two of the four: gravity and electromagnetism.

The strongest force is the strong force, and that was how it was named. Then it's electromagnetism,which is sometimes called the electromagnetic force. The weak force is weaker than both the strong and the electromagnetic force, but ironically, it is not the weakest. The weakest force is gravity that we experience everyday.

Even though gravity is very weak, its range is the most flexible and is infinite. It can make an apple fall, but it can also hold the solar system together.

Bosons

Bosons are on the right side of the Standard Model. While fermions are matter particles, bosons are force carrying particles that cannot construct matter. Each fundamental force corresponds with one type of boson, as shown on the left. Again, the graviton is not yet observed experimentally, but scientist think that it must exist, as gravity does not have a corresponding boson yet.

In the subatomic world…

Strangely, even though gravity is what keeps us on Earth and affects all objects, it is much weaker than the other three forces. In fact, the strong force is 1,000,000,000,000,000,000,000,000,000,000,000,000,000 times stronger than gravity. That's 10 to the power of 38 in scientific notation. Gravity is so weak that the particles aren't really affected by it. So, gravity is normally not considered when we're talking about interactions between the subatomic particles. That's good for us, because we don't know how gravity fits into everything yet.

四種宇宙基本功

有了一大堆能夠構成物質的費米子有什麼用呢？那些粒子並不知道怎麼互動，更不能單獨構成物質。這就好像你擁有一些樂高積木，但卻沒有說明書，即便有建造飛船的材料也沒辦法建成飛船。而這時，四種宇宙基本力就有用了。它們就像是一本準則，告訴費米子們如何互動、如何構造物質。所以，在我們的世界中，構造物質的粒子沒有基本力的指引，是完全無用的。“宇宙基本力”聽起來可能很困難，但是你可能已經認識其中的兩個了：引力和電磁力。。

最強的力是強力，這就是為什麼科學家們叫它強力。第二強的是電磁力，然後是弱力。弱力比強力和電磁力都弱，但是可笑的是它卻比引力強上千倍。引力是最弱的力。這可能令你吃驚，但是仔細想像生活中的例子你就能發現引力確實很弱。

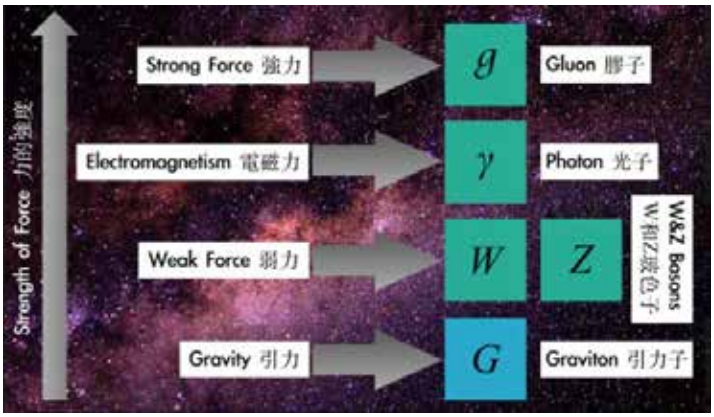
即便引力非常弱，但是它的作用範圍是最有彈性的，也是無限的。引力能夠讓一個蘋果從樹上掉下，但也可以讓整個太陽系穩定地運轉在一起。

玻色子

玻色子在標準模型的右邊。費米子是構成物質的粒子，而玻色子是傳遞力的粒子，不能構成物質。看看左邊的圖，每個基本力都與一個玻色子掛鉤。這裏再次注意：引力子並沒有真正被探測到。但科學家們猜測它應該存在，因為現在還沒有一個玻色子負責引力，而其他力都有負責的玻色子。

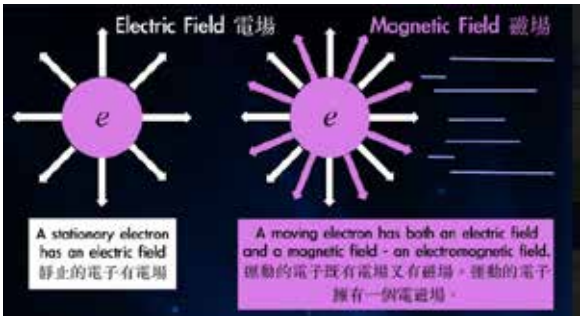
在粒子物理世界裡……

奇怪的是，我們最熟悉的引力和其它宇宙基本力比起來真是弱得可笑。強力比引力強1,000,000,000,000,000,000,000,000,000,000,000,000,000 倍。所以，粒子其實並不怎麼被引力影響，我們在討論它們之間的互動時也把引力忽略不計了。這很方便，因為我們到現在還不知道引力和次原子粒子世界有何聯繫。



Electromagnetism

Electromagnetism, or the electromagnetic force, is very related to our daily lives. You've all had the experience to get shocked by static electricity; that's the electric force. You've certainly bought fridge magnets before; that's the magnetic force. These two forces are more similar than you think. When a charged particle (for example an electron that carries an electric field) moves, it also generates a magnetic field, thus having an electromagnetic field. So, electric and magnetic fields are pretty much the same. They both come from one fundamental force: the electromagnetic force.



Photons

In 1960, scientists came up with the theory that the electromagnetic force is exchanged between fermions through boson particles called photons. Photons are particles with no mass and no electric charge. It has a spin of 1, which makes sense because it is a boson, and bosons have spins that are multiples of 1. Photons are also particles of light, so naturally, they travel at the speed of light. The speed of light is the fastest speed possible in the universe, equal to roughly 300,000,000 meters per second.

Photons exchange the electromagnetic force between particles that have an electric charge. This means that particles without an electric charge (for example the electrically neutral neutrinos) do not interact with the electromagnetic force. Then, how do we represent the exchange of photons? Around the 1950s, a theoretical physicist named Richard Feynman thought of a simple method to represent this: using cartoon-like drawings. These drawings are called Feynman Diagrams. Not only does the Feynman Diagram describe the electromagnetic force, they also tell us what happens when any fundamental particles meet.

Here is an example of a Feynman Diagram. It shows what happens when two electrons with -1 charge meet (remember that like charges repel, see page14). When they reach the first vertex where the exchange occurs. One electron spits out a photon which hits the other electron, and that's why after the exchange, the electrons repel each other. They separate at the second vertex. Of course, the diagram can represent how two particles attract as well.

電磁力

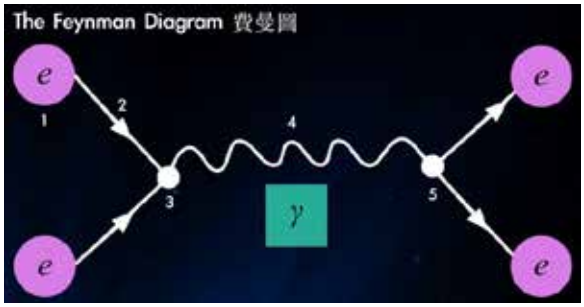
電磁力與我們的生活聯繫緊密。你應該在乾燥的天氣中被朋友“電”過，那就是靜電的電力 ；你也應該買過冰箱貼，那就是磁力。而這兩個力其實非常相似。當一個帶有電荷的粒子（比如一個擁有電場的電子）開始運動時，一個磁場也會被產生。這時，這個運動粒子就有一個電磁場了。所以，電和磁是可以相互轉換的。它們都源於一個基本力：電磁力。

光子

在 1960 年，科學家們發現一種叫光子的玻色子負責傳播電磁力。當費米子以電磁力互動時，它們交換光子。光子是一種沒有質量、電荷為零的粒子，它的自旋是 1。這符合玻色子的自旋特徵：它們的自旋必須是 1 的倍數（見15頁）。光子是光的粒子，所以它們以光速傳播。光速是宇宙中最快的速度，約等於 300,000,000 米每秒。

光子在帶電粒子之間傳播電磁力，這說明不帶電的粒子（比如電中性的中微子）是不會與電磁力互動的。我們如何來表示這種光子的交換呢 ？1950 年左右，一位名叫理查德·费曼的理論物理學家想到了一種簡單的方法：使用卡通式的圖表來表示粒子的互動。這些圖表叫做費曼圖。費曼圖不僅僅能表示電磁力，它還可以表示任何基本粒子互動時的情境。

下圖就是費曼圖的一個例子。它表示兩個帶有-1電荷的電子相遇時的情景（記得同種電荷互相排斥，見14頁）。當它們到達第一個頂點時，一個電子 “吐出 ”一個光子。那光子會撞開另外那個電子。兩電子在第二個頂點分開，因此就互相排斥了。當然，費曼圖也能表示兩個粒子相互吸引的過程。



1. The type of fermion, in this case an electron.
   費米子的類別，這裏指的是電子。
2. Straight lines with arrows indicate fermions and their direction.
   帶有箭頭的直線代表費米子和它們的運動方向。
3. The first vertex where the two fermions meet
   第一個頂點，兩個費米子相遇。
4. Wiggly lines indicate bosons, in this case a photon
   波浪線代表玻色子，這裏是光子。
5. The second vertex where the two fermions separate.
   第二個頂點，兩個費米子分開。

Strong Force

Even though the strong force is the strongest force, its range is limited to a space no larger than a proton. So we don't experience the strong force in daily life, and that might be the reason why you are unfamiliar with it. The strong force plays two major roles. First, between quarks, it "glues" the quarks into groups to form composite particles like protons. And second, in the atom, it holds the nucleus together. Therefore, without the strong force, no atom heavier than hydrogen (which has only 1 proton), no atom can be formed and the world would fall apart.

Between Quarks

The strong force is responsible for "gluing" quarks together so they can form particles like protons and construct all forms of matter. The strong force is exchanged between quarks by bosons called gluons. You probably know where the name "gluon" comes from.

Gluons have no mass and no electric charge. Its spin is 1 because it is a boson. The red, green and blue colors used for the quarks in the diagram above actually has significant meaning in terms of gluon exchange and the strong force.

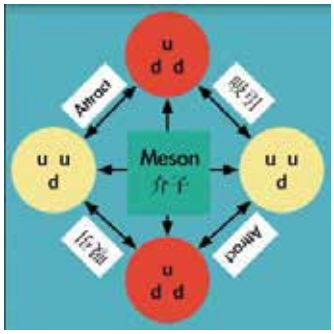
In Atoms

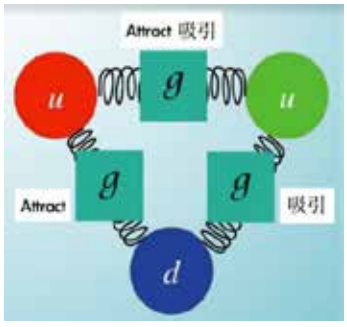
The electromagnetic force causes particles with the same electric charge to repel. That's a problem: if the nucleus is made of protons and neutrons, why don't the protons which all have +1 charge repel each other and escape out of the nucleus? The strong force is the answer. As the strongest force, it can overcome the electromagnetic force which protons repel. In the atom, meson particles relay the strong force and pulls the protons and neutrons together to form the nucleus.

While electromagnetism repels just the protons, the strong force counteracts the electromagnetic force. The strong force, through the exchange of mesons, makes all the protons and neutrons attract so that they can form a stable nucleus.

The Strong Force and Leptons

The strong force has no effect on leptons. This means that quarks are the only fermions that experience the strong force; they can interact with gluons.





強力

雖然強力是四種基本力中最強的一個力，它的範圍被限制在質子大小之內，因此我們在日常生活中不會經歷強力。強力有兩個重要的作用。第一，在夸克之間，它使夸克組合起來構成像質子一樣的複合粒子。第二，在原子內，它負責將質子和中子固定在原子核內。所以，若沒有強力，原子就不會存在，世界也就不會存在了。

在夸克之間

“黏”在一起、讓它們構成像質子這樣的粒子。強力就像膠水一樣。夸克傳遞強力時交換一種名叫“膠子 ”的玻色子。你應該知道膠子這個名字是從哪裡來的了。

膠子沒 有質量，不帶電荷，而且作為玻色子的它自旋為 1。上圖中在夸克上使用的紅、綠和藍並不是完全隨意的。顏色其實在強力和膠子交換方面具有特殊意義。

在原子內

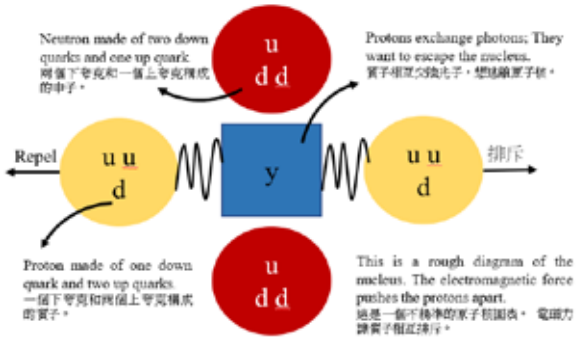
電磁力是第二強的基本力，而它會讓帶有同種電荷的粒子相互排斥。你可能會發現一個問題：原子核是由質子和中子構成的，那麼為什麼電荷為 +1 的質子不會相互排斥、逃出原子核呢 ?強力能解決這個問題。

作為最強的力，它能夠壓倒令質子相互排斥的電磁力。在原子中，強力由介子傳遞。它能夠將質子和中子束縛綑綁在一起來構成原子核。

電磁力讓質子相互排斥，而強力能夠抵銷電磁力的作用。通過介子，強力讓所有的中子和質子互相吸引，構成一個穩定的原子核。

強力和輕子

強力並不影響輕子。夸克是唯一能夠經歷強力的費米子，它們能夠放出或吸收膠子。

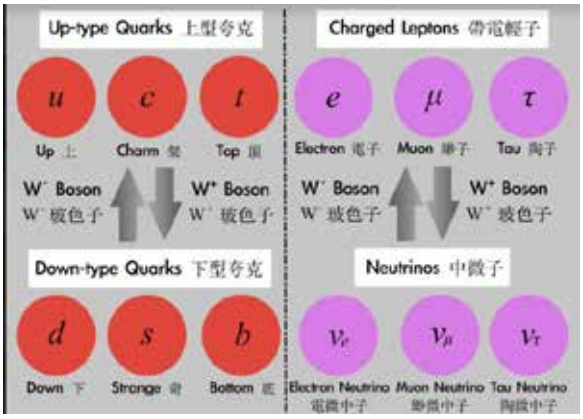


In diagrams on the left (especially Feynman Diagrams), the exchange of gluons is represented using a spring-like line. 在費曼圖中(左圖)，膠子的交換是用圈狀的曲線（與彈簧相似）表示的。

The diagram on the left shows two up quarks and one down quark forming a proton under the effect of gluons. 左邊的圖中，兩個上夸克和一個下夸克。正在交換膠子；強力使它們構成質子。

Weak Force

The third strongest force is the weak force. The current temperature of the universe causes it to be weaker than both the strong and the electromagnetic force, so it's conveniently called the weak force. We don't see its effects in daily life, because the weak force is responsible for strange interactions that change particles entirely. It can change up-type quarks to down-type quarks, and the other way around. It can change charged leptons to neutrinos and the other way around too.



W-Bostons and Z-Boston

The W and Z-Bosons are both responsible for the weak force. The diagram above shows the effects of the electrically charged W-Bosons.The W-plus boson turns up-type quarks into down-type quarks, and charged leptons into neutrinos. The W-minus boson does the opposite. The Z-Boson is not charged, similar to a photon, though it has mass. It exchanges energy, but leaves the particle unchanged.

Decays

The weak force can change particles. This is very useful because it allows heavier particles (generation 2 and 3 particles) to change into lighter particles (generation 1 particles). It even allows heavier elements to turn into lighter ones. This process is called "decay". This is precisely why heavier particles are rare and unstable: because of the weak force, they decay into lighter particles that make up normal matter. Of course, under proper conditions with specific requirements, lighter particles can turn into heavier ones. An example is the process of fusion that occurs at the center of our sun, which will be explained.

弱力

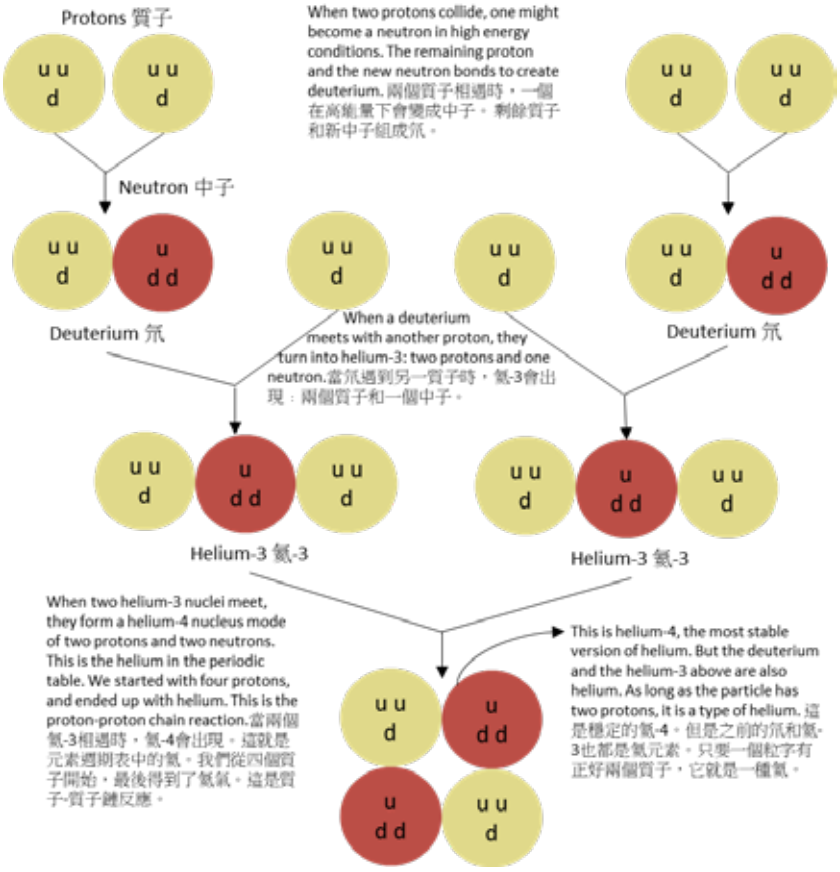
第三強的力是弱力。弱力比強力和電磁力都弱，因此科學家們很沒有創意地把它稱為弱力。弱力對我們來說是非常陌生的。它負責一些奇怪的反應：將一種粒子直接轉成另一種粒子。它能把上型夸克轉成下型夸克，或把下型夸克轉成上型夸克。它也能把帶電輕子轉成中微子，或把中微子轉成帶電輕子。

W 玻色子和 Z 玻色子

W 玻色子和 Z 玻色子都負責傳遞弱力。左圖顯示的是帶電的 W 玻色子的作用。帶正電的 W 玻色子將上型夸克變成下型夸克、將帶電輕子轉成中微子。帶負電的 W 玻色子的作用則相反。而不帶電的 Z 玻色子就像一個擁有質量的光子，它能交換傳遞能量，但是不會改變粒子的種類。

衰變

弱力能夠改變粒子的種類，這非常有用，因為它將較重的粒子（第二代和第三代粒子）變成更輕的粒子（第一代粒子）。弱力甚至可以將重元素變成更輕的元素，這過過程叫做“衰變”。這正是第二代和第三代粒子不穩定、很稀少的原因：弱力的作用使重粒子衰變成構成日常物質的輕粒子。當然，在特別情況下，輕粒子也是能變成更重的粒子的。一個典型的例子就是太陽中心的核聚變過程，下圖會講明。



Gravity

Gravity is the our most familiar force, but also the most unimportant when fundamental particles interact. That's because gravity is the weakest force. It influences anything with mass, but is extremely weak. Just think about it: we can easily jump up and overcome gravity for a while. The electromagnetic force is much stronger than gravity. A small magnet can lift a large piece of metal. This, combined with how the particles are small in mass, causes scientists to ignore gravity completely when discussing how particles interact. But, when we're talking about something with huge mass, like stars or solar systems, gravity's influence is overpowering.

In A Star

Diagram on the right is a star undergoing the nuclear fusion process. The heat and energy produced when protons are turning into helium atoms pushes against the gravity which threatens to collapse the star.

Gravitons

Excluding gravity, all other three fundamental forces have bosons that relay them. Scientists then predicted that gravity would have a responsible boson as well, and they came up with the idea of "gravitons". Gravitons are boson particles that (in theory) exchange the attractive force of gravity between objects with mass. Although the graviton has not seen yet, and also not included in the official Standard Model, physicists actually know quite a lot about it.

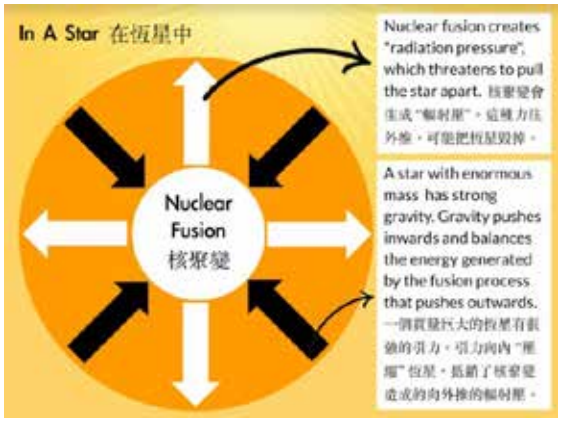
> A graviton should be a massless fundamental boson particle. It should have no electric charge, but should carry energy. It should have a spin of 2, which makes sense as 2 is a multiple of 1, and that's what defines a boson.

Extra Dimensions

Now you may be wondering: if we already know so much about gravitons, why haven't scientists detected them yet? The answer is that gravitons might lie in extra dimensions. It is difficult to understand this because we live in a three-dimensional world. It's possible that we cannot detect extra dimensions because we are too big. But it's possible that something as tiny as a graviton can detect extra dimensions, and even actually exist in them. Plus, since we don't know if these dimensions exist, we have no way of detecting the gravitons. The existence of extra dimensions is actually a very promising idea right now. If you're curious, research String Theory.

引力

我們最熟悉的力就是引力，但是它卻是粒子物理中最不重要的基本力。這是因為引力非常地弱，不太影響基本粒子的互動。想想看我們生活中的例子你就能發現這一點：我們輕輕一跳就能暫時克服引力。電磁力比引力強多了。雖然整個地球的引力將某塊金屬向下拉，但一塊小小的磁鐵就能夠吸起那塊金屬。加上基本粒子的質量如此小，科學家們在討論粒子之間的互動時往往完全忽略引力的影響。但是，我們在討論質量非常大的物體時（比如恆星或太陽系），引力的作用幾乎是壓倒性的。



上圖描述的是一個經歷核聚變過程的恆星。質子變成氦元素過程中，大量的熱和能量被釋放。這種向外力量和恆星向內的引力形成平衡，使恆星發光發熱。

引力子

除了引力之外，另外三個宇宙基本力都有負責它們的玻色子。因此，科學家們猜測引力也被一種玻色子傳遞。他們猜測一種名叫“引力子”的玻色子傳遞引力，負責吸引任何擁有質量的物質。雖然引力子還沒有被科學家們發現，也沒有被列入正式的標準模型中，物理學家通過標準模型已經比較了解它了。

> 引力子應該是一個沒有質量的基本玻色子。它不帶電，但是帶能量。它的自旋是 2，這是合理的，因為 2是1的倍數，而玻色子的自旋應該是 1的倍數。

額外維度

現在你可能在想：若我們已經這麼了解引力子了，為什麼科學家們還沒有發現它們呢？答案是引力子可能藏在額外維度中。這可能非常難以想像，因為我們生活在三維世界中。有可能是因為我們太大了，才導致我們探測不到這些額外維度。但是對於微小的引力子來說，額外維度是可以被探測得到的，他們甚至可以在額外維度中存在。而因為我們至今都不知道額外維度是否真的存在，我們就無法發現引力子了。額外維度的存在其實在粒子物理中是一個很受歡迎的想法。若你好奇的話，看看“弦理論”吧。

Higgs Boson

Now that you've learned this much about the fundamental particles and forces, have you ever thought about why some particles are heavier than others? This inspires a discussion of mass. In the diagram on the right, you can see that the top quark is heavier than the up quark, Top quark which is heavier than the electron neutrino, and the photon is the lightest- it has no mass. How do some particles have mass and others don't?

The Higgs boson is the final piece of the Standard Model. The Higgs boson relates to the Higgs field. It is different from the other fundamental bosons because it doesn't carry force between fermions, rather it lets all particles with mass interact with itself. In other words, the Higgs field interacts with all fundamental particles that have masses (all fermions and the weak force bosons) by giving them the mass they have. The Higgs boson gives particles mass. While the Higgs field might seem abstract, you only need to understand it as an invisible field that stretches throughout the universe, a field that interacts with particles through the exchange of Higgs bosons.

The Higgs boson doesn't interact with all particles in the same way. It interacts with some more than others. The more the Higgs boson interacts with a particle, the more massive the particle is. This is why some particles are more massive than others: they interact with the Higgs field more.

Top Quark 頂夸克

The top quark is the most massive of the fundamental particles. It interacts with the Higgs field the most. 頂夸克是粒子中質量最大的，與希格斯場互動最多。

Electrons 電子

Compared to the top quark, the electron interacts far less with the Higgs field and is much lighter. 和頂夸克相比，電子與希格斯瑒互動得少多了，質量也小多了。

Electron Neutrino 電微中子

Electron neutrinos have tiny, near-zero masses. They interact with the Higgs field least out of all. 中微子的質量幾乎是電微中子 零。它在基本費米子中與希格斯場互動得最少。

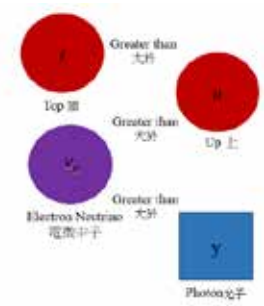
Photons 光子

Photons zoom past the Higgs field without interacting with it. As a result, it has no mass. 光子運動時完全不與希格斯場互動，因此它沒有質量。

HiggsBoson

The Higgs boson has no charge and its spin is 0. But it does have mass, and interacts with itself.

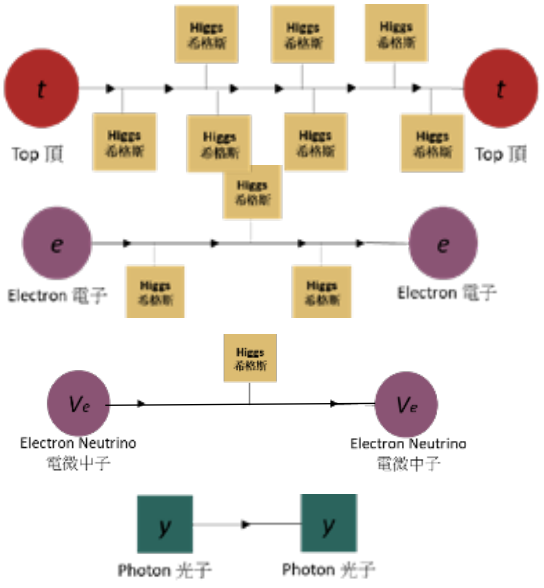
希格斯玻色子



你有沒有思考過為什麼一些基本粒子比另外一些基本粒子的質量更大呢 ?這會引起關於質量的討論。右圖顯示，頂夸克比上夸克大於重，上夸克比電微輕子重，而光子是最輕的，它沒有質量。為什麼有些基本粒子有質電微中子量，而其他粒子卻沒有呢 ？

標準模型中最後的是希格斯玻色子。希格斯玻色子與希格斯場有關。希格斯玻色子與其他玻色子是不同的，因為它並不在費米子之間傳遞力，而是讓所有擁有質量的粒子與自己互動。換一種說法，希格斯場與所有擁有質量的基本粒子（所有費米子和傳遞弱力的玻色子）互動，給予它們應有的質量。雖然希格斯場可能很抽象，但是你只需要把它想成宇宙中一個隱形的場，一個通過交換希格斯粒子來與其他粒子互動的場。

希格斯玻色子並不以同種方式與全部的粒子互動，它與某些粒子互動得較多。希格斯玻色子和某粒子互動得越多，那個粒子的質量就越大。這就是為什麼一些粒子比其它粒子的質量更大：它們與希格場互動得更多。



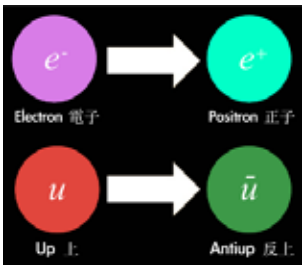
希格斯玻色子

希格斯玻色子是電中性的，自旋為 0。但是它擁有質量，所以也會與希格斯場進行互動，它們與自己互動。

Antimatter

Scientists observed that the twelve fundamental fermion particles each have an "opposite" particle. These mirror particles behave in ways that are completely opposite of the original particle. They are called the "antimatter" of the fundamental particles.

To change ordinary fundamental fermions to anti matter fermions, the sign of the electric charge flips, while other characteristics-like spin-stay the same. Flipping the electric charge would also completely reverse the properties of a particle. The process of flipping electric charge to change a particle to its antiparticle is called "Charge Inversion". For example, an electron of -1 charge will become an antielectron (also commonly called a positron) if its charge changes to +1. We draw a line above the particle to show that it is antimatter. The line is called a bar.



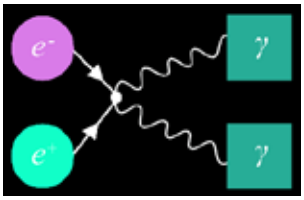
The superscript of "-" and "+" tells us the charge of the particle. The electron has -negative charge, the positron has positive charge. The subscriptis only for charged leptons and their antimatter particles

The bar on top of the up quark ("u" is the symbol) shows that it is an up antiquark.

Annihilation

We don't see antimatter in real life, because when antimatter meets matter, they destroy each other in a process called "annihilation".

Everything in the world is matter,so antimatter cannot exist: it gets annihilated instantly.



The Feynman Diagram shows the annihilation between an electron and a positron. Annihilation releases high energy photons (γ)

In the table on the right, the first row are the up-type antiquarks with charge-2/3, as opposed to the original +2/3 of the up-type quarks. Using the same idea, the down-type antiquarks have charge +1/3, the charged antileptons have charge +1, and the antineutrinos have charge 0. Today, scientists create antimatter in high-energy environments such as in particle accelerators, where particles are smashed together with high speeds. However the antimatter particles still cannot survive for long.

反物質

科學家們發現 12 個基本費米子各有一個“相反”的粒子。這些粒子的特性與原來的粒子完全相反。它們就是基本粒子的“反物質”。

要將普通的基本費米子轉成相對應的反物質，你只需要將反轉一下它的電荷就可以了（其他像自旋性質都保持原樣）。反轉某個粒子的電荷會完全反轉那個粒子的性質。比如，要把電子轉成正電子（也叫正子），只需要將它原來的-1 電荷換成 +1 電荷。為了區分粒子和反粒子，我們在反粒子的上方畫一條橫槓。

“-”和“+”的上標告訴我們粒子的電荷。電子帶有負電荷，而正子帶有正電荷。

“u”上面的橫槓告訴我們它是反上粒子。

湮滅

我們在生活中見不到反物質，因為當反物質觸碰物質時，它們會在一個叫“湮滅”的過程中毀滅彼此。因為世上萬物都是正常物質，所).費曼圖顯以反物質不能存在：它立刻就會被湮滅。 示的是電子和正子的湮滅現象。湮滅過程會釋放高能量光子。

上表顯示的是反費米子。第一行是反上型夸克。它們的電荷是-2/3，是原來上型夸克電荷+2/3 的相反數。同理，反下型夸克的電荷是+1/3、反帶電輕子的電荷是+1、反中微子的電荷是 0。現今，科學家們在高能量環境下創造反物質。比如，反粒子能在粒子加速器（見40頁）中被生成。但是，即便在這種人為環境中，反物質仍不能長期存活。

Color Charge

Color charge is a property of quarks. Its concepts are very similar to those of electric charge. While electric charge arises from the electro magnetic force, color charge arises from the strong force. Just like how every particle that interacts with the electromagnetic force has an electric charge, every quark (quarks are the only fundamental particles that interact with the strong force) carries color charge.

However, while a particle's electric charge can only be two ways – as in either positive or negative -scientists discovered that aquark's color charge can be three ways. Physicists wanted to search for a simile for this three-way charge, so they thought of the primary colors of light. Each color charge corresponds to a primary color: red, green and blue. Unfortunately, a quark's color charge is not related to what type of quark it is. For example, one up quark can have blue charge while the other has green charge. Scientists have to study the behavior of the quarks to determine their color. Antiquarks have anticolor, as you can see in the diagram below.

The anticolor of green is antigreen, also called magenta. Magenta is a mix of red and blue. The anticolor of red is antired, also called cyan. Cyan is a mix of green and blue. The anticolor of blue is antiblue, also called yellow. Yellow is a mix red and green. Remember that the color charge of quarks has nothing to do with actual colors. They are just creative analogies.

Similar to electric charge, like colors repel and opposite colors attract. Thus three quarks with different color charge will attract.

Color-neutral:White

There are two ways to make color neutral white light particles. Once the particle is color neutral, it would not attract any color charged quarks, just like how an electrically neutral particle does not attract any electrons. The first way is to combine all three colors: red, green and blue. This will be a combination of three quarks, and the classic example is making a proton.

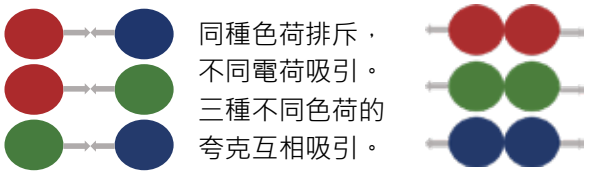
The example above forms a baryon, which will be explained later. Colorless, unstable antibaryons can be made by combining three antiquarks of yellow, magenta and cyan. The second way to create a color-neutral particle is to combine a quark with an antiquark with its anticolor (e.g. blue and yellow). Here, a meson is formed, which will be explained.

色荷

色荷是夸克的一種性質，色荷的概念和電荷的概念大同小異。電荷是由電磁力產生的，而色荷是由強力產生的。就像每個與電磁力互動的基本粒子都帶電荷一樣，每個與強力互動的夸克（夸克是唯一能跟強力互動的基本粒子）都帶色荷。

一個粒子的電荷只能有兩個 “方向”：要麼是正數，要麼是負數。但是科學家們發現一個夸克的色荷可以有三個 “方向”。物理學家們想為這個三方向的色荷找一個類比，於是他們想到了三原色：紅、綠、藍，每個色荷與一個顏色對應。不幸的是，色荷不是由夸克的種類來決定的。兩個相同的夸克可以有不同的色荷。一個上夸克可以是藍色的，而另一個上夸克就可能是綠色的。科學家們必須根據夸克的表現來決定色荷。反夸克有反顏色，見左圖。

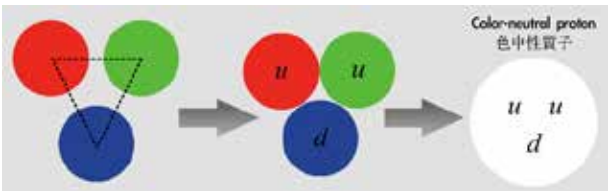
綠的反顏色是反綠，也叫做洋紅。洋紅是紅和藍均勻混合得來的顏色。洋紅的反顏色是反紅，也叫做青。青是綠和藍均勻混合得來的顏色。藍的反顏色是反藍，也叫做黃。黃是紅和綠均勻混合得來的顏色。夸克的色荷與真正的顏色沒有關聯。顏色只是有意思的類比。



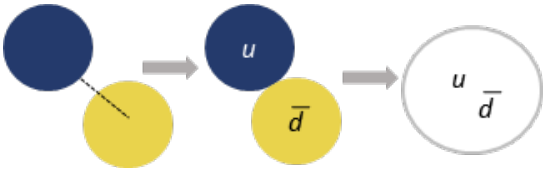
同種色荷排斥，不同電荷吸引。三種不同色荷的夸克互相吸引。

色中性：白色

一個色中性粒子不會吸引任何帶色荷的夸克，就像一個電中性粒子不會吸引任何電子一樣。製造色中性白光粒子有兩種方法。第一種方法是混合三種顏色：紅、綠、藍。這是三個夸克的組合，而最典型的例子就是構建質子。

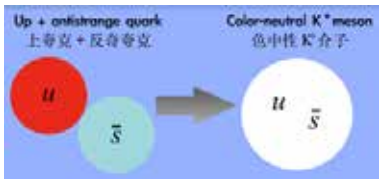


上圖中的例子構成的是一個重子。三個帶有黃、洋紅和青的反夸克也能構成不穩定的色中性反重子。第二種構成色中性粒子的方法是組合一個夸克和帶有對應反顏色的反夸克（比如藍和黃），構成介子。



Mesons

Typically, mesons are subatomic particles made of one quark and one antiquark, bound together by the strong force. An example of a meson was given: one upquark and one antidownquark can bond to make a π meson. As mesons are made of two quarks with spin ½ (antiquarks also have ½ spin, because only the electric charge changes), a meson would have a spin of ½ + ½=1. So, mesons are also composite bosons, because they fit into the boson definition of "particles with spin that is a multiple of 1".



Another example of a mesonis the K^+ meson, or the kaon. It is made of one upquark and one antistrange quark as a color anticolor pair (red and cyan).

The Strong Force

Flip back to the strong force. It was mentioned that mesons are responsible for the strong force that binds the nucleons (protonsand neutrons) together. The meson that does this is the π meson. Mesons can only survive a short time outside the nucleus, because they are made of exotic particles like the antistrange quark.

Exotic Mesons

Mesons are defined as particles made of an even number of quarks. Most mesons are made of 2 quarks, but there are exotic mesons made of 4 quarks (a.k.a. tetra quarks). These mesons are extremely unstable and we don't know a lot about them. Tetraquarks are made of 2 quarks and 2 antiquarks (*q* means "quarks").

Baryons

Typically, baryons are subatomic particles made of three quarks, bound together by the strong force. The most normal examples of baryons are protons and neutrons, as both are made of the common upquarks and down quarks. However, there are many more baryons, such as the Ξ (xi), Λ (lambda), or Σ (sigma). These baryons are made of heavier quarks from generations 2 and 3, and quickly decay into lighter particles. Baryons are made of three quarks of spin ½ + ½ + ½ = $^3/_2$. This means that baryons are composite fermions, as the spin of $^3/_2$. is a multiple of ½.

The "0" as a superscript of "Ξ" tells us the electric charge of the particle, showing that it is electrically neutral. The upquark has $+^2/_3$ charge, while the two strange quarks each have $-^1/_3$ charge. Adding $+^2/_3$ and two $-^1/_3$ together, we get 0.

Hadrons

Both mesons and baryons are hadrons-particles made of quarks. As mesons are bosons and baryons are fermions, we know that hadrons can either be bosons or fermions; it just has to be made of quarks. So there's across over between fermions, bosons, mesons, baryons and hadrons.

介子

通常來說，介子是一種由一個夸克和一個反夸克組成的次原子粒子。有一個例子：一個上夸克和一個反下夸克組成π介子。因為介子是由兩個自旋為½的夸克組成的（反夸克的自旋也是 ½，因為只有電荷變化），所以介子的自旋是½+½=1。因此介子也是複合玻色子，因為它們遵循了玻色子的定義："自旋為 1 的倍數的粒子"。

另一個介子的例子是左邊的介子。它由一個上夸克和一個反奇夸克組成，一個帶紅色荷，一個帶反紅的青色荷。

強力

你已經知道介子負責傳遞把核子（質子和中子）"綁"起來的強力。這個介子就是 π 介子。在原子核之外，介子只能生存很短時間，因為它們是由像反奇夸克這樣不穩定的粒子構成的。

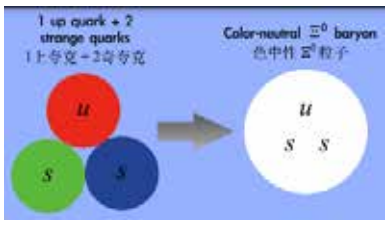
奇異介子

介子的定義是由偶數個夸克組成的粒子。大部分介子都由兩個夸克構成，但是有些奇異介子是由四個夸克組成的（叫做四夸克組合）。四夸克組合由兩個夸克和兩個反夸克組成（q表示夸克）。

重子

通常來說，重子是一種由三個夸克組成的次原子粒子。我們最熟悉的重子是質子和中子，因為這兩個粒子都是由常見的上夸克和下夸克構成的。但是，科學家們還發現了很多其他重子，比如 Ξ 粒子、 Λ 粒子、 Σ 粒子等等。這些重子會快速衰變成更輕的粒子。因為重子是由三個夸克組成的，它的自旋是 ½ + ½ + ½ = $^3/_2$.，$^3/_2$.是½的倍數。因此所有的重子都是費米子。

下圖中的 Ξ 粒子有一個 "0"的上標。這告訴我們 Ξ 粒子的電荷是 "0"，說明它是電中性的。上夸克的電荷是 $+^2/_3$，而兩個奇夸克各帶 $-^1/_3$的電荷。把 $+^2/_3$和兩個 $-^1/_3$加起來，我們就能得到 0。



Another example of a baryon is the Ξ baryon, consisted of an up quark and two strange quarks. The particle will quickly decay into lighter particles in a chain reaction.另一個重子ss 的例子是左邊的Ξ粒子。ss它由一個上夸克和兩個奇夸克組成。這個粒子會迅速衰變成更輕的粒子。

強子

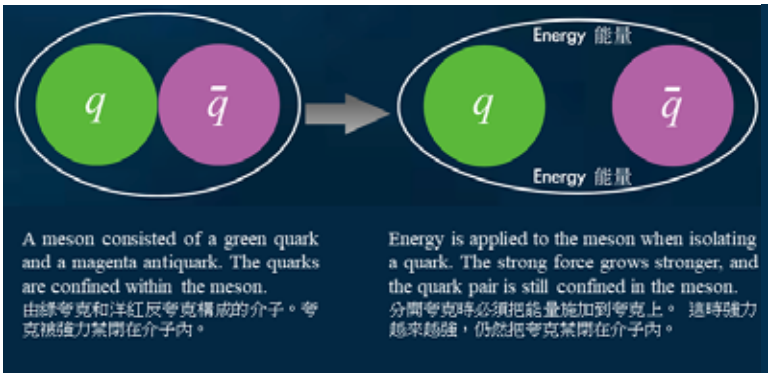
介子和重子都是強子：由夸克組成的粒子。因為介子是玻色子而重子是費米子，我們知道強子既可以是玻色子也可以是費米子，它只要是由夸克組成的就可以基本費米子了。所以費米子、玻色子、介子、重子和強子之間有重疊。

Quark Confinement

It is impossible to study a single quark because of something called "quark confinement". Confinement means being locked or restricted so that something cannot leave a certain area. Quark confinement means that a quark can never be isolated, and so we can never study a quark by itself. Color-charged quarks are always confined in groups within a hadron by the strong force; and a single quark can't move freely outside a hadron. But what happens when we try to separate a quark from a hadron by pulling on it?

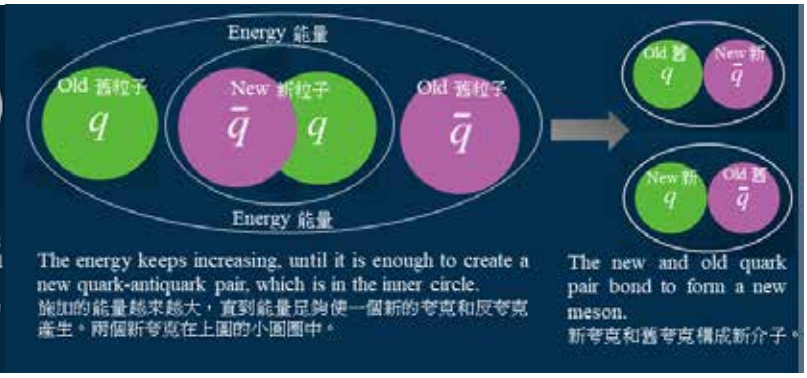
夸克禁閉

夸克永遠不能單獨存在，因為一種叫 “夸克禁閉 ”的現象將夸克鎖在強子之內。 “禁閉 ”的意思是將某個東西控制在某個地方之內，所以那個東西永遠離開不了限定的範圍。 “夸克禁閉 ”的現象讓我們無法將一個夸克孤立出來，所以我們也無法單獨觀察一個夸克。夸克永遠都會被強力兩個兩個地鎖在介子之內，或三個三個地鎖在重子之內。一個單獨的、帶色荷的夸克不可能在強子之外自由地移動，所以能夠自由移動的粒子一定都是色中性的 “白色”強子。但是，若我們真的嘗試將一個夸克從強子中 “拉出來 ”，讓它單獨存在，會發生什麼事呢 ？



A meson consisted of a green quark and a magenta antiquark. The quarks are confined within the meson.
由綠夸克和洋紅反夸克構成的介子。夸克被強力禁閉在介子內。

Energy is applied to the meson when isolating a quark. The strong force grows stronger, and the quark pair is still confined in the meson.
分開夸克時必須把能量施加到夸克上。 這時強力越來越強，仍然把夸克禁閉在介子內。



The energy keeps increasing, until it is enough to create a new quark-antiquark pair, which is in the inner circle.
施加的能量越來越大，直到能量足夠使一個新的夸克和反夸克產生。兩個新夸克在上圖的小圓圈中。

The new and old quark pair bond to form a new meson.
新夸克和舊夸克構成新介子。

We've already talked about how the strong force acting between quarks is responsible for quark confinement (keeping quarks inside a hadron). The interesting thing about the strong force is that it gets stronger and stronger as the quarks get further away. This means that if you want to overcome the strong force and isolate a quark, you need more and more energy. The quarks use that energy to create new quarks that bond together, rather than getting separated. When the new quarks are created, the old quarks are still very close together (the distance between them is smaller than a proton), so you can't observe a quark alone.

But when the quarks are in the hadron (like the diagram between, where the white circle represents the hadron), the quarks can move freely.

Asymptotic Freedom

The strong force acts like an rubberband. When you pull at the rubberband, it gets tight and you need more power to stretch it. But when you let the rubberband slack, there is almost no force in the band. The strong force works exactly like that. When we use force to pull quarks apart, it is strong as the quarks are far away. But it is weak when the quarks are near each other, which allows the quarks to move around freely. This is called "asymptotic freedom", where you can understand "asymptotic" as something you never reach.

你已經知道在夸克之間，強力負責 “夸克禁閉 ”現象，負責將夸克禁閉在強子之內。強力有種非常有趣的性質：當夸克相距越來越遠時，在它們之間作用的強力會越來越強。這等於說若你想克服強力來孤立一個夸克的話，你需要越來越多的能量。而夸克會使用這些能量來創建新的夸克組合、與新的組合結合形成新強子，而不是被能量硬拉開。當新的夸克組合產生時，原來的夸克仍然離得還非常近（它們之間的距離比質子還小），所以我們永遠無法觀察單獨存在的夸克。

但是當夸克在強子內時（如左圖，白色的圈代表強子），強力較弱，因此它們可以自由地活動。

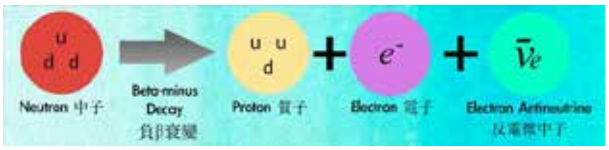
漸進自由

強力就像橡皮筋一樣。當你拉伸橡皮筋的時候，它會變得非常緊，你也需要更多能量來把它拉得更長。但是當橡皮筋是鬆弛的時候，橡皮筋中存在的力是幾乎為零的。強力正是這樣運作的。當夸克相離得很遠的時候，強力非常強，不允許夸克自由運動。而當夸克離得很近的時候，強力非常弱，夸克也就能在強子的範圍內自由運動了。這叫做 “漸進自由 ”。你可以將 “漸進 ”理解為一個永遠達不到的東西。

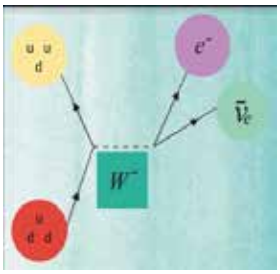
Beta Decays

The weak force makes heavier particles quickly decay into ligher ones. In the process of beta decays (also called β-decays), a protons are made or turns into a neutron or a neutron turns into a porton. As protons are made of two up quarks and one down quark, and neutrons are made of one up quark and two down quarks, beta decay make an up quark change into a down quark (when a proton turns into a neutron), or a down quark change into an up quark (when a neutron turns into a proton). So, there are two types of beta: beta-plus decay, and beta-minus decay.

Beta-minus Decay



The neutron turns into a proton. Because the neutron has charge 0 and the proton has charge +1, an electron with charge -1 must also be released, so the total charge after the decay is still -1 + 1= 0. So, an electron and an electron antineutrino with no charge are released as by products.



Feynman Diagran above, the W Boson's effects are represented using a dashed line, rather than the photon's waves or the gluon's circles. The neutron changes into a proton under the influence of the W Boson. An electron and an electron antineutrino are released.

Beta-plus Decay

The proton turns into a neutron. This means that the overall electric charge decreased from the +1 of the proton to the 0 of the neutron. So, a positron with charge +1 must also be released to balance the electric charge, so that after the decay the charge is still +1 + 0 = +1. A neutrino is also released to balance the extra positron, which is antimatter.

Beta decay is called beta decay because it is a decay that produces beta particle - electrons or positrons with high energy. Beta-minus decays produce electrons with minus (negative) charge, where as beta-plus decay produce positrons with plus (positive) charge. There are many other types of decays, like alphadecay, gammadecay, or simply heavier particles from generations 2 and 3 decaying into lighter ones. The weak force affects all fundamental fermions, unlike the strong force which only affects quarks (not leptons) and the electro magnetic force which only affect charged particles (not neutral neutrinos).

β 衰變

弱力讓較重的粒子衰變成輕的粒子。在 β 衰變

的過程中，一個質子衰變成中子，或一個中子衰變成質子。因為質子是由兩個上夸克和一個下夸克構成的，而中子是由一個上夸克和兩個下夸克構成的，所以 β 衰變只是讓一個上夸克變成一個下夸克(質子變成中子)，或一個下夸克變成一個上夸克(中字變成質子)。因此，β 衰變有兩種：正 β 衰變和負 β 衰變。

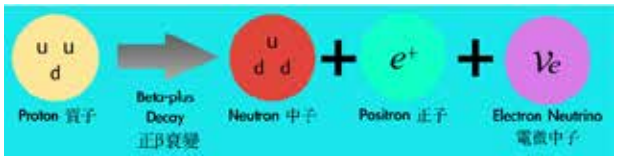
負 β 衰變

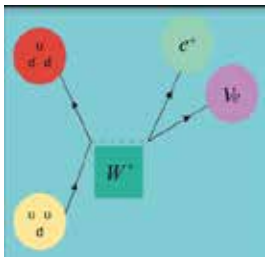
中子變成質子。因為中子不帶電荷而質子帶 +1 的電荷，所以一個帶 -1 電荷的電子必須產生，讓最終的總電荷變為-1+1=0。因此，一個電子和一個不帶電的反電微中子在此過程中被釋放。

左圖是負 β 衰變費曼圖。W 玻色子用虛線表示，而不是用光子的波浪線或膠子的圓圈形的線。一個中子在 W 玻色子的作用下變成一個質子。一個電子和一個反電微中子被釋放。

正 β 衰變

質子變成中子。這說明電荷從質子的 +1 減少到了中子的 0。因此，一個帶 +1 電荷的正子生成，使衰變後的電荷保持在原來的 0。一個中微子也被釋放了，來平衡新生的正子（正子是輕子的反物質）。





左圖是正 β 衰變費曼圖，與之前的費曼圖概念相同。這次是帶正電荷的 W 玻色子，因為帶正電荷的質子變成了電中性的中子，而正 W 玻色子能夠平衡這個負向的電荷轉變。

為何 β 衰變被稱為 β 衰變呢 ?因為它是一種會生成 β 粒子的衰變。 β 粒子是帶有高能量的電子或正子。負 β 衰變會釋放帶有負電荷的電子，而正 β 衰變會釋放帶有正電荷的正子。粒子物理界中還有很多不同的衰變，比如 α 衰變、 γ 衰變，或較重的粒子衰變成更輕的粒子……弱力不像只影響夸克的強力（輕子不受影響）或只影響帶電粒子的電磁力（中微子不受影響），弱力影響全部的基本費米子。

Particle Accelerators

Atoms are tiny, so tiny that we cannot view them under a normal light microscope. Just imagine how tiny quarks must be! It would be impossible to observe any of the subatomic particles without very special equipment, much less study antimatter that do not belong in our world. That's where particle accelerators come in. "Accelerate" means to travel faster and faster. Particle accelerators use electromagnetic fields to send particles into very high speeds and energies, so trackers along the accelerator can detect the electrical signals as the particles zoom past. Currently, there are over 30,000 particle accelerators worldwide.

The Large Hadron Collider (LHC) located in Geneva, Switzerland. It belongs to CERN (European Organization for Nuclear Research), which operates the largest particle physics lab the world. The LHC is the biggest and most powerful particle accelerator. It is an underground circular tunnel 27 kilometers long.

Dark Section

Let's stop talking about subatomic particles, and zoom out to the universe. The Standard Model, though impressive, only describes the visible universe, the part of the universe that we can actually see. Sadly, the visible section is just a tiny fraction o f the stuff out there. The majority of the universe is invisible; we cannot see it using our current scientific equipment. What other things are out there in the universe that are not described by the Standard Model?

We know very little about dark matter. It seems to be what causes stars to form galaxies under the influence of gravity. But, we have no idea what makes up dark matter. It's not antimatter, nor particles in the Standard Model. Particle accelerators are smashing particles together, and we hope that dark matter particles will be produced. However, we have no evidence for these particles yet.

We know even less about dark energy, which is almost 70% of our universe. It could even be made of imaginary particles or empty space! We just know that it is working against gravity to stop the expansion of the universe.

粒子加速器

原子非常小，小到它在顯微鏡下也看不見。這樣你就能想像夸克有多麼小了。我們沒有特別特殊的設備就不可能觀察任何次原子粒子，更不可能研究不在我們世界中存在的反物質。這就是粒子加速器的作用。“加速”的意思就是運動得越來越快。粒子加速器使用電磁場來加速粒子，直到粒子達到非常高的速度和能量。在加速器中的探測器能夠探測到粒子的電場，因為它們的速度和能量都太高了。現在，全球有超過 30000 個粒子加速器。

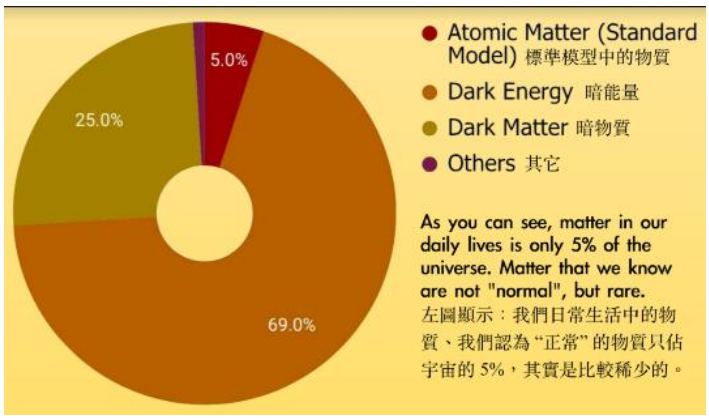
在瑞士日內瓦的大型強子對撞機。它屬於 CERN（欧洲核子研究中心），世上最大的粒子物理實驗室。大型強子對撞機是世界上最大的、最強的粒子加速器。它是個圓形的地下隧道，有 27 千米長。

神秘物質

我們暫時先停止討論次原子粒子世界，來看看浩瀚無邊的宇宙吧。雖然標準模型很厲害，但是它只能形容整個可見宇宙（我們能看見的宇宙）。不幸的是，我們現在能看見的宇宙只是整個宇宙很小的一部分，宇宙很大一部分都是我們用現在的科學技術看不到的。那麼，宇宙中還有什麼奇特的物質是標準模型形容不了的呢 ？

我們對暗物質了解的很少。它好像可以讓恆星在引力的作用下組成銀河系。但是，我們完全不知道暗物質是由什麼粒子構成的。無論如何，它都不是由任何在標準模型中的粒子構成的，也不是由反物質構成的。科學家們正在使用粒子加速器撞擊粒子，希望能找到構成暗物質的粒子。目前，我們還沒有任何證據。

雖然暗能量佔宇宙的 70%，我們對它幾乎是一無所知。它甚至可能是由不存在的虛幻粒子構成的！我們只知道它好像在對付引力，減緩宇宙的膨脹。



References

Butterworth, Jonathan. “What's the Smallest Thing in the Universe?” YouTube,TED-Ed,15Nov.2018,www.youtube.com/watch?v=ehHoOYqAT_U.

Choi, Charles. *“Doomsday and Disembodied Brains? Tiny Particle Rules Universe's Fate.”* NBC News, 13 Sept. 2013, www.nbcnews.com/sciencemain/doomsday-disembodied-brains¬tiny-particle-rules-universes-fate-8C11144331.

Gonick, Larry, and Art Huffman. *The Cartoon Guide to Physics*, New York, Harper Collins Publishers, 2005.

Jackson, Tom. *An Illustrated History of the Foundations of Science*, New York, Shelter Harbor Press,2013.

Johnson, Wyatt. *“The Higgs Boson Simplified Through Animation.”* YouTube, 25 Mar. 2014, www.youtube.com/watch?v=L6AN6UwTTjU.

Schaffner, Paul. “Lepton Decays.” *The Particle Adventure,* Lawrence Berkeley National Laboratory (LBNL), 2013, www.particleadventure.org/lepton_decay.html.

Sci Show. “So What IS the Higgs Boson?” YouTube, 13 Jan. 2012, www.youtube.com/watch?v=WUnDsNL_5nk.

Still, Ben. *Particle Physics Brick by Brick*, Cassell Illustrated, 2017.

蔡志申、楊盛智。“基本粒子。”科學 Online,2011 年 1月，highscope.ch.ntu.edu.tw/wordpress/?p=19087

夸克黑科技。“認識了解夸克。”每日頭條， 2016 年 12月，kknews.cc/science/2azj4kg.html.

吳旭明。“物質的組成 —原子與原子核的組成 —夸克的發現。” YouTube，2016 年 6月， www.youtube.com/watch?v=g4I7Muhlowo.

梁保勝、武杰。“「夸克幽禁」是對物質無限可分的挑戰嗎？”內蒙古農業大學學報，社會科學版， 2009。

“你是費米子 ？還是玻色子 ？”每日頭條，2017 年 4月，kknews.cc/zh-hk/science/en6kzkr.html.

中國科學院。“北京谱仪国际合作组发现四夸克物质 Zc(3900)入选 2013 年物理学重要成果。”中國科學院條件保障與財務局， 2013 年 12月，www.bpf.cas.cn/kyjd/201401/t20140102_4009628.html

張峻輔。“大成功 ？大失敗 ？—漫談超對稱在高能物理學中的妙用。”科學月刊，2013 年 5月， scimonth.blogspot.com/2013/05/blog-post_8948.html.

黃啟新。“夸克，黑洞與上帝。”無可推諉，www.withoutexcuse.net/index.php?option=com_content&view=article&id=93:quark-blackhole-god-tw&catid=25&Itemid=129&lang=zh-tw.

利根、史皮羅普路。“超對稱玩完了 ？”科學人雜誌，2017 年， sa.ylib.com/MagArticle.aspx?Unit=featurearticles&id=2443.

“引力子真的存在嗎 ？”每日頭條，2016 年 12月，kknews.cc/zh¬hk/science/p4zkjk2.html.

Network Analysis of the Equine Gut Microbiome
Following Antibiotic Treatment Reveals Key Species and Dependencies

Yu Han Daisy Wang, Dionne Daiyin Yeung

Executive Summary

The gut microbiome is integral to functions such as maintaining metabolism and the immune system. Disruption, such as from antibiotic usage, potentially leads to deadly diseases such as colorectal cancer and colic. Oral antibiotics are known to disrupt the gut microbiome, but their broader effects and the dynamics of recovery are unknown. Hence, our research aims to answer the question **"What are the long term effects of oral antibiotics?"**

Traditional methods of microbiome analysis are commonly used for microbiome profiling but typically only focus on relatively shallow or compositional based measures of abundance and phylogeny. This means that they may fail to detect the complex relationships and interactions between microbes. Here, we propose the use of co-occurrence analysis to understand the dynamics of recovery in the gut microbiota following antibiotic usage, which allows us to obtain results in two-fold: 1) to infer overall shifts in the hierarchy and structures within the network of the microbiota, and 2) to show key microbes and inter-microbial relationships.

We studied a group of 60 racehorses, 16 of which consumed oral antibiotics. Monthly faecal samples were collected over the span of six months. Metagenomic next generation sequencing was used to determine the species and the number of each species in each sample, which we then used to conduct our analysis.

Traditional microbiome analysis methods, such as alpha and beta diversity, failed to reveal antibiotic-induced changes. But, co-occurrence analysis was able to show how antibiotic usage lead to great disruption in the relationships between microbes, as shown by the networks shattering into smaller, sparser clusters. Furthermore, co-occurrence analysis allowed the identification of *Akkermansia* as a key species, identification supporting relationships between microbes, and also discovery of potentially competitive relationships.

1. Background

1.1 Importance of a Healthy Microbiome

A microbiome defines the collection of microbes that interact with each other within a given space or environment (Berg *et al.,* 2020). Organisms have co-evolved with a plethora of microbes, forming strong symbiotic relationships that benefit the host-microbe system. In humans, the microbiome functions to produce essential resources, mediate interactions, and bioconvert nutrients. The body's microbiomes are also the first line of defence against foreign microbes and viruses (Young, 2017). For instance, changes in the lung microbiome influence one's susceptibility to infection from the influenza virus (Lee *et al.*, 2019).

Microbiomes often have social structure, in which microbes rely on complex interactions to sustain themselves and perform functions. For instance, the metabolic pathways responsible for nourishment rely on the interactions between archaea and cellulolytic bacteria to degrade nutrients. Thus, maintaining eubiosis - a well connected, balanced, and healthy gut microbiome - is of utmost importance (Kauter *et al.*, 2019). Eubiosis in the gut is characterized by temporal and spatial heterogeneity, where each niche is fully colonized with many variants of microbes. The highly diversified gut microbiome is essential for maintaining metabolism, modulating host gene expression (Hooper and Gordon, 2001), and preserving host health (Iebba *et al*., 2016). The gut microbiome has even been

The research for this article was conducted in the *Shuyuan* Molecular Biology Laboratory and presented as a poster at ASM Microbe 2019. Presentations of this work were also entered for the S.T. Yau High School Science Award (Asia) and the Hong Kong Student Science Project Competition (HKSSPC), in which the authors were finalists.

shown to interact with organs outside the gastrointestinal tract and has been linked to the regulation of mood and behavior, strengthening the immune system, and contributing to numerous pathophysiological disorders (Diaz Heijtz *et al*., 2011).

Furthermore, dysbiosis - a disequilibrium in the microbiome with respect to the needs of the individual organism (Hedayat and Lapraz, 2019) - may lead to increased rates of obesity, contracting illnesses (*e.g.* colorectal cancer, Alzheimer's disease), and inhibited organ functionality. Increasing amounts of research point towards the significant role that microbes play in maintaining good health (Iebba *et al.,* 2016).

1.2 Equine Gut Microbiota

The gut microbiota is crucial in equines, where there exists a strong relationship between the composition of the enteral microbiome and its function (Kauter *et al*., 2019). Disturbances can cause issues such as gastrointestinal disease (a leading cause of mortality in horses (Al Jassim and Andrews, 2009)), colic (a highly lethal infection with 63% equine survival rate (Kauter *et al*., 2019)), and laminitis (an extremely painful equine disease (Venable *et al*., 2016)). Thus, maintaining a healthy intestinal microbiome in horses is of paramount importance in averting diseases (McKenney and Pamer, 2015).

While antibiotics are often an effective treatment for infection, their use can also induce and exacerbate several serious conditions. In both humans and equines, antibiotics can induce diarrhea (Bartlett, 1984; Harlow *et al.,* 2013),, *Clostridium difficile* infections (Kauter *et al*., 2019; Båverud *et al*., 2003), colitis (Jacobs, 1994; Larsen, 1997), *etc*. Previous research on antibiotics' effects on animal gut microbiomes have produced conflicting results – some showed no discernible effects (Grønvold *et al*., 2010) while others demonstrated drastic reductions in microbes (Harlow *et al*., 2013) and prolonged disturbances to the gut microbiota (Jernberg *et al.,* 2010; Manichanh *et al.,* 2010; Nobel *et al*., 2015; Korpela *et al*., 2016; Costa *et al*., 2015; Ji *et al*., 2018).

Many of these studies rely on *in vitro* methods, but may result in inaccurate simulations or shallow compositional analysis. Furthermore, they focus only on abundance and phylogenetic data, and are unable to capture antibiotic-induced changes in relationships within the gut microbiome. Hence, we propose the use of co-occurrence analysis to model the impact of antibiotics on relationships in the gut microbiome.

Type	Name	Description
Diversity Analyses	Alpha Diversity	Measures how many kinds of taxa or lineages are in a sample
	Beta Diversity	Measures how the taxa or lineages are shared in between samples
Qualitative and Quantitative Analyses	Qualitative Analyses	Examines presence-absence data over regions
	Quantitative Analyses	Involves relative abundances
Phylogenetic and Taxon Based Analyses	Phylogenetic Analyses	Involve analysis within species through a phylogenetic tree
	Taxon Based Analyses	Make comparisons between species on a taxonomic level

Table 1. An overview of common microbiome profiling methods.

1.3 Profiling the Microbiome

Traditional culture methods (TCMs) culture samples in enriched broth or agar, based medium, then isolate microbes for identification and characterisation (Gupta *et al*., 2019). However, TCMs have a bias towards microbes that grow well under laboratory conditions, and they do not replicate the environment of the microbiome (Dowd *et al*., 2008; Davies *et al*., 2000). Thus, they fail to authentically capture the full diversity of a microbiome (Rhoads *et al*., 2012). Furthermore, TCMs also require an extended amount of time to culture the microbes, making the method very time consuming (Flayhart *et al*., 2007; Wolcott and Dowd, 2008). By relying on the presence of DNA, RNA, and proteins, culture-independent methods bypass the limitations of culturing of microbes, while speeding up the identification of microbes (Deurenberg *et al*., 2017; Poretsky *et al*., 2014), addressing the shortcomings listed above (Riesenfeld *et al*., 2004). In particular, 16s rRNA sequencing has become particularly popular (Zapka *et al*., 2017).

There are three major components to common microbiome profiling methods: diversity analyses, qualitative and quantitative analyses, and phylogenetic and taxon-based analyses (Hamady and Knight, 2009), as detailed in Table 1. These methods consider the abundances or phylogeny of microbes, but fail to account for relationships that are crucial in understanding the microbiome. This highlights the need for network-based analysis to provide a more in-depth understanding of the gut microbiome's dynamic changes in response to antibiotic treatment.

Co-occurrence analysis, which uses the repeated measurements of objects' abundance to model relationships through significant co-occurrence or mutual exclusion patterns, has shown recent success in aiding understanding of the complex relationships between microorganisms (Gilbert *et al*., 2014). Co-occurrence networks have also shown success in revealing site specializations of microbes, coexistence and clustering of microbes (Faust *et al*., 2012), as well as environmental-based microbial community structure shifts (Faust and Raes, 2016).

While co-occurrence analyses have been applied successfully in better understanding static systems, very few have applied this method of analysis to study changes in the microbiome longitudinally, and certainly not to study the impact of a large change such as antibiotic treatment.

1.4 Investigation

In this longitudinal study, we aim to answer the following research question: **"What is the impact of oral antibiotics on the equine gut microbiome?"**

We chose to investigate the equine gut microbiome because it can shed valuable insight into the functionalities and patterns of the gut microbiome in humans. Furthermore, uniquely the Hong Kong Jockey Club racehorses provide a homogeneous group for study with well documented physical performance and veterinary treatments.

2. Methodology

2.1 Subject

Sixty thoroughbred racehorses (median age 6 years) under a single trainer, with similar feed and work regimes, were recruited from the Hong Kong Jockey Club. Faecal samples were collected from each horse monthly over six months, with a total of 360 samples. Medical records of the horses before and during this period detailed antibiotic usage.

2.2 Sequencing

Genomic DNA from fresh faecal samples was extracted using the PureLink Genomic DNA mini kit. PCR amplification of the variable region 3-5 (V3-V5) of the 16S rRNA gene was performed before using the Illumina MiSeq sequencing platform to generate forward and reverse reads.

2.3 Data Processing

Sequence analysis was performed with Quantitative Insight Into Microbial Ecology 2 (QIIME2) (Bolyen *et al*., 2019). Reference data includes the raw sequences of each faecal sample and clinical metadata for each subject. The q2-demux plugin was used to demultiplex and filter the quality of the raw sequences. Raw sequences were then de-noised using DADA2 (q2-dada2) (Callahan *et al.*, 2016). A phylogenetic tree was constructed using fasttree2 (Price *et al*., 2010) from amplicon sequence variants (ASVs) aligned with mafft (Katoh *et al*., 2002). Samples were rarefied, then used to estimate alpha diversity metrics (observed OTUs, Shannon Index (Shannon and Weaver, 1963)), beta diversity metrics (weighted UniFrac (Lozupone *et al*., 2007)), and Principle Coordinate Analysis (PCoA). Taxonomic association was carried out using the q2-feature-classifier (Bokulich *et al*., 2018) classify-sklearn naive Bayes taxonomy classifier against Greengenes 16S rRNA reference sequences at 97% similarity (DeSantis *et al*., 2006). A flowchart of our data analysis is shown in Figure 1.

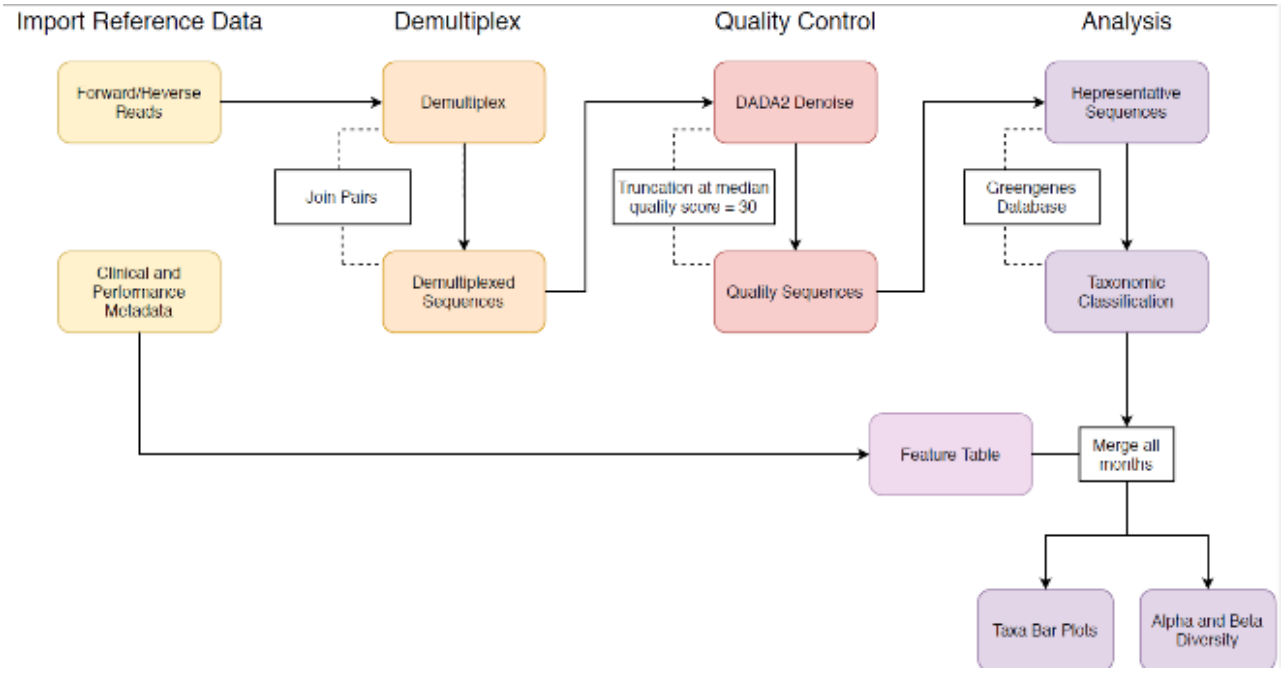


Figure 1. A flowchart of our data analysis workflow

2.4 Alpha Diversity Analysis

QIIME2 was used to calculate Shannon Index (H) (Shannon and Weaver, 1963) of the equine gut microbiome. H, as calculated by equation 1 below, quantifies the diversity of samples by considering both abundance and evenness of the species present in and between the samples. The variable 'p' is the number of individuals of a single species divided by the total number of individuals of all species. In this work, the number of 16S DNA reads of a particular species is a proxy for the number of individuals. A higher H reflects a higher diversity, and *vice versa*. H provides a useful overview, even if species in small proportion may be over-weighted.

2.5 Principle Coordinate Analysis (PCoA)

$$H = -\sum_{i=1}^{s} p \times ln(p) \quad (1)$$

PCoA plots were generated based on the beta diversity metric weighted UniFrac distance, calculated from OTU tables. Weighted UniFrac was chosen as our distance measure as it accounts for both phylogeny and sample abundance, and is thus more suited to our longitudinal study (Lozupone *et al*., 2007).

2.6 CoNet

CoNet [37] was used to create and analyse co-occurrence networks from the OTU tables. CoNet takes in an abundance matrix as input – each row represents a different microbe, and each column represents observations. Co-occurrence analysis looks for potential relationships between two nodes based on how often they "co-occur", or are present in the same measurement. The generated graphs (Figure 5-12) consists of species and edges, where each node denotes a different microbe, and correlated nodes are connected by edges.

Networks were generated using CoNet for the following months of antibiotic intake: control (horses without antibiotic use in the previous 12 months), month 0 (month of antibiotic use), month 1 (1 month following antibiotic use), month 2 (2 months following antibiotic use), month 3 (3 months following antibiotic use), month 4 (4 months following antibiotic use), month 5 and 6 (5 and 6 months following antibiotic use), month 7 and 8 (7 amd 8 months following antibiotic use). Six horses were randomly chosen from the control group to create the control network in order to ensure comparable values across the networks. Months 5 and 6, as well as months 7 and 8, were combined in order to ensure comparable sample sizes between networks. Pearson correlation, Spearman correlation, Kendall correlation, Distance correlation, mutual similarity, and Bray Curtis Dissimilarity were each set to a threshold of 0.5. A p-value threshold of <0.05 was set to ensure all edges were statistically significant. Mutual similarity,

a measure of dependency between different elements in the network, was chosen for this project, hence the edges did not have a positive or negative measure.

3. Results

3.1 Alpha Diversity

As demonstrated in Figure 2, the Shannon index of the equine gut gut microbiome seems to fluctuate naturally through time with a range of roughly 6.5 to 8.

Shannon diversity of the equine gut microbiome does not seem to be greatly affected by the consumption of antibiotics, as seen in figure 3. Changes present are confined to a similar range to the natural fluctuations found in Figure 2.

3.2 Principle Coordinate Analysis (PCoA)

A PCoA plot generated based on Weighted UniFrac [43] is shown in figure 4. Each colour corresponds to a different month after antibiotic usage. Most horses were quite similar, clustering in the middle left of the graph. However, some deviation was present, with some points deviating to the lower right of the graph.

3.3 Networks

Networks were colored by degree, which is the amount of edges that a node is connected to. Warmer colours and smaller nodes indicate lower degrees, and vise versa. For a description of what each graph represents, please refer to section 2.6.

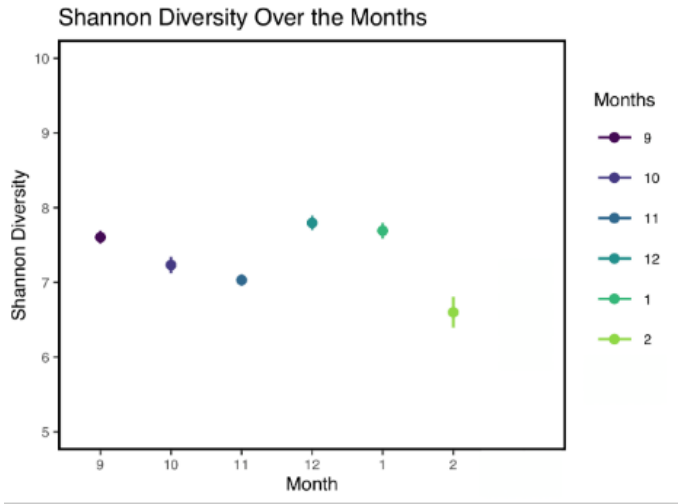


Figure 2. Shannon diversity of the entire cohort of horses in each month. (Error bars are shown as ± 1 s.d.)

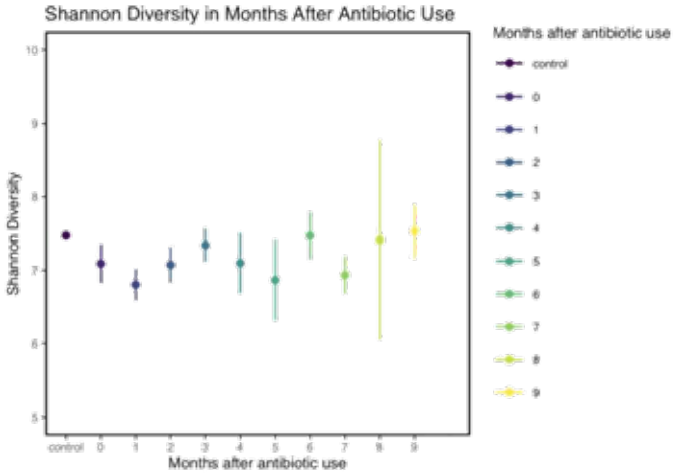


Figure 3. Shannon index in horses in months before and after antibiotic use. (Error bars are shown as 1 s.d.).

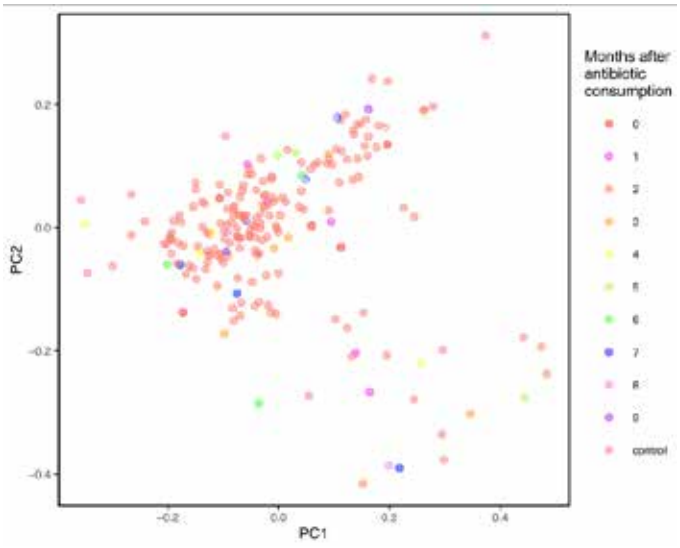


Figure 4. PCoA based on weighted UniFrac distance. Colour usage reflects the month after antibiotic consumption.

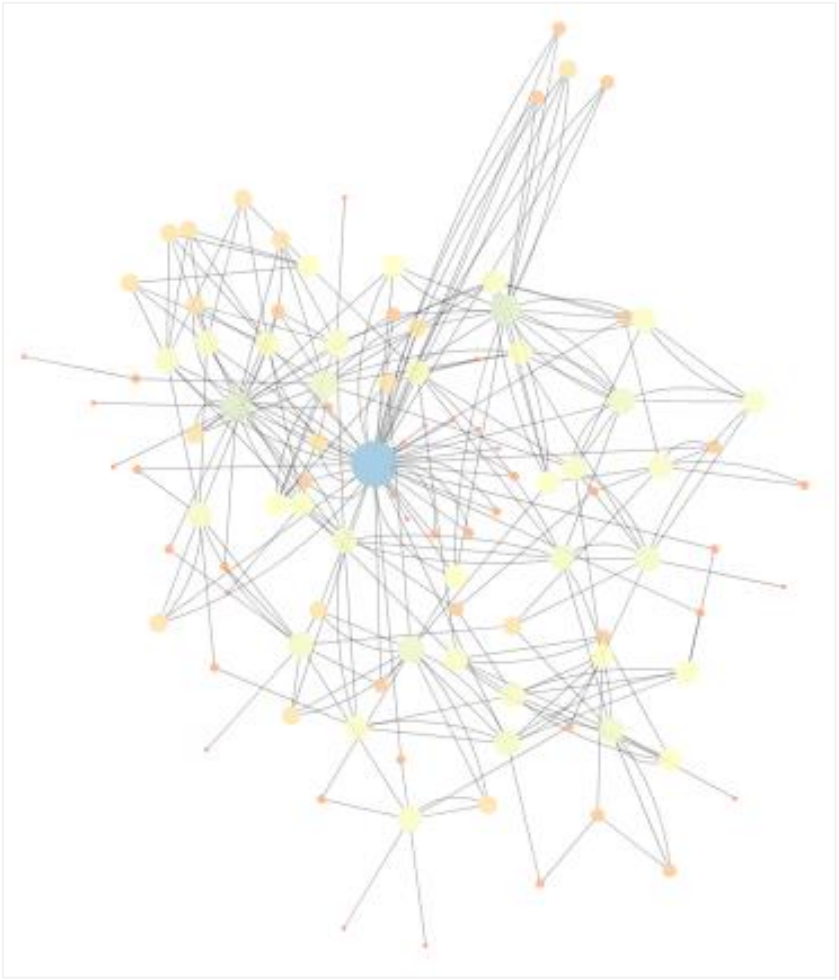


Figure 5. Control - The control network is relatively well connected and dense, with a central node present in blue.

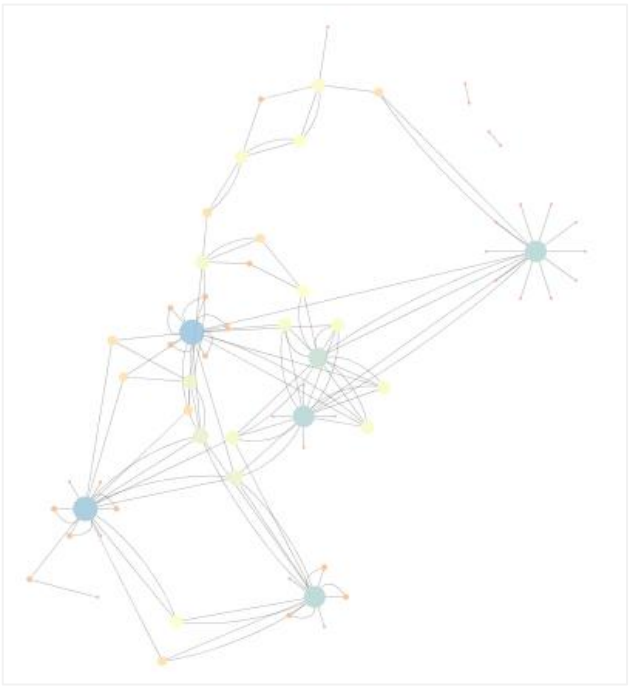


Figure 6. Month 0 - Immediately after antibiotic consumption, the network becomes much more sparse and disconnected

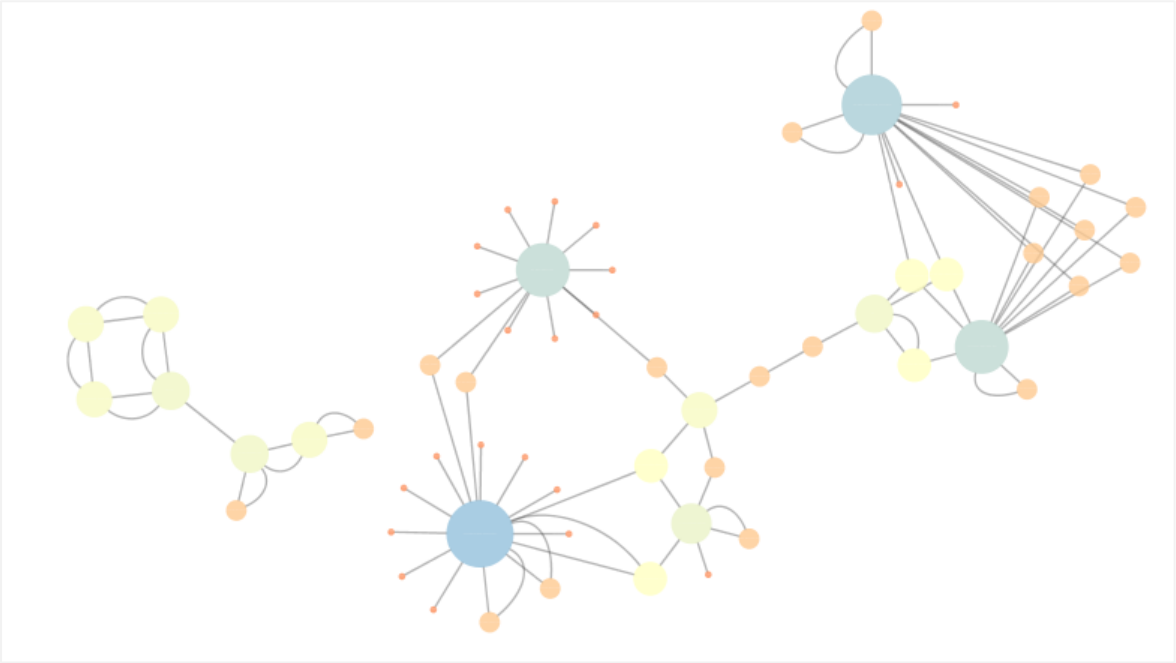


Figure 7. Month 1 - Some recovery seems apparent one month after antibiotic consumption, with the network becoming more connected.

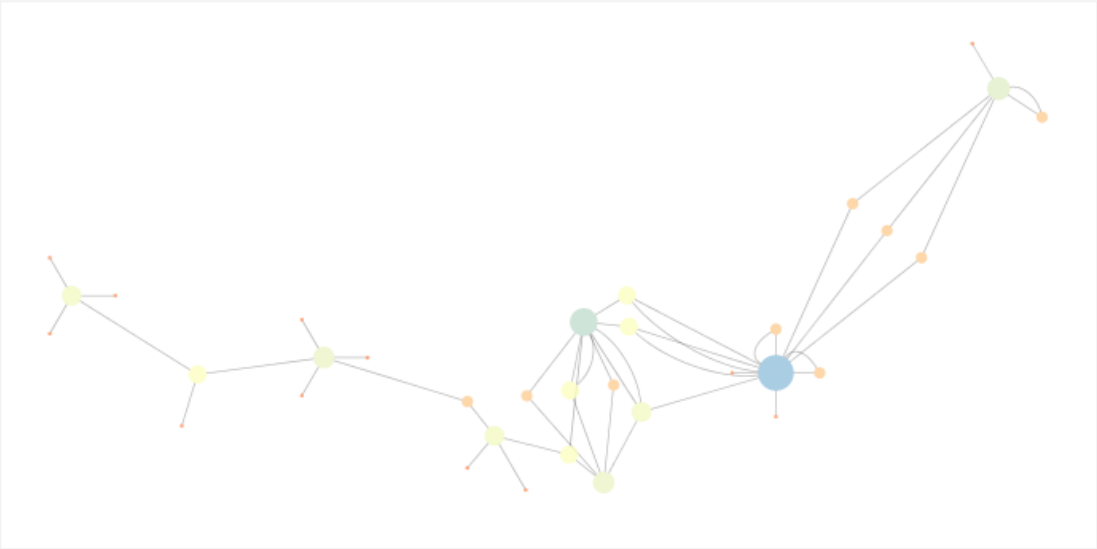


Figure 8. Month 2 - The network appears more sparsely connected once again, showing fluctuations in recovery.

Figure 9. Month 3 - There seems to be clusters forming and an increase in density.

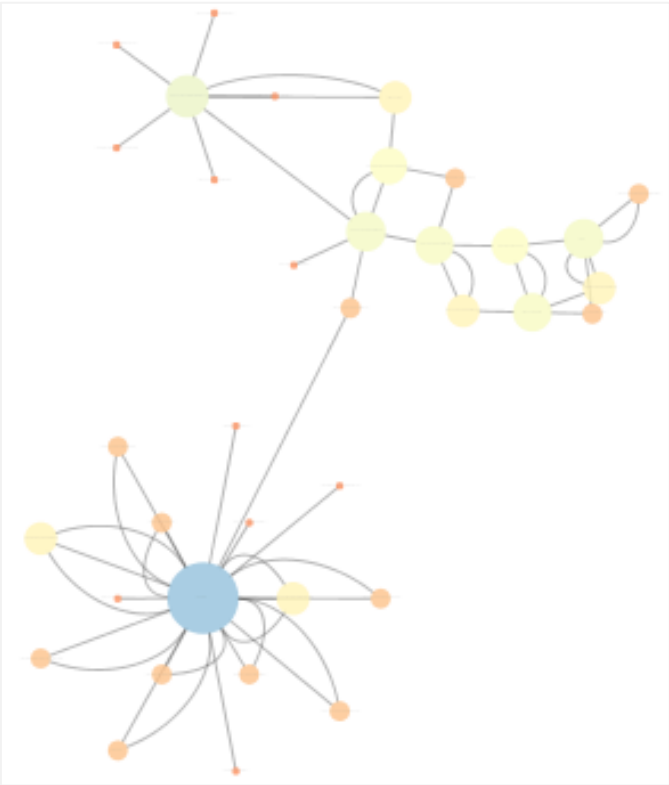


Figure 10. Month 4 - The network regains connectivity, as there are no longer clear fragmentations.

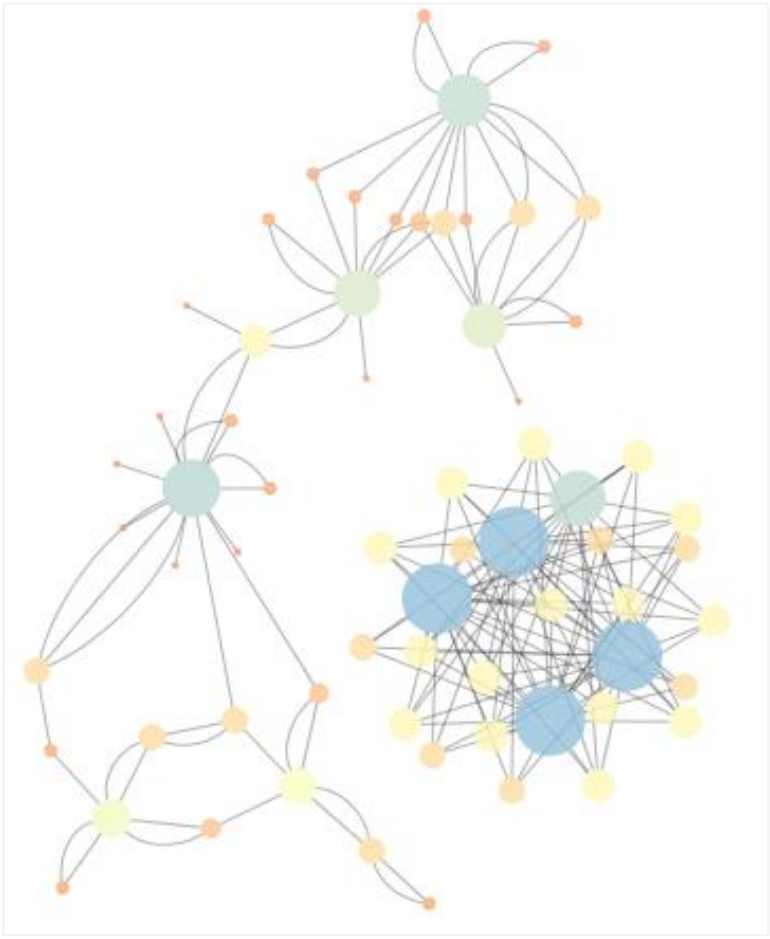


Figure 11. Months 5 and 6 - The network is once again quite dense, and clustering is prominent.

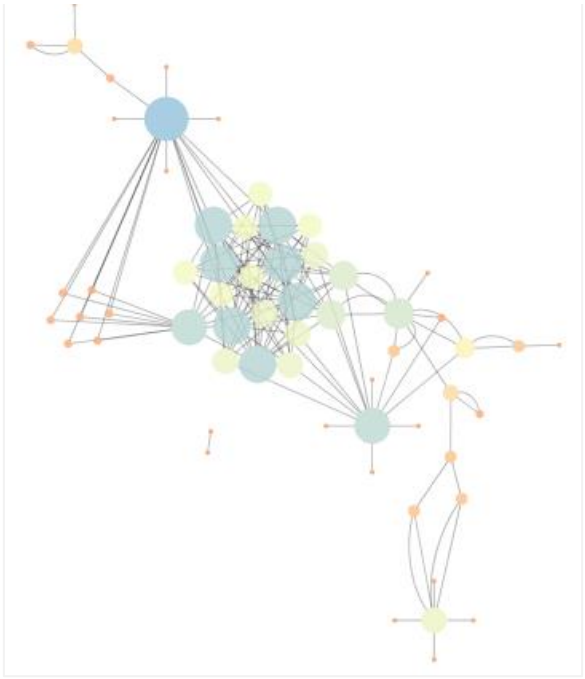


Figure 12. Months 7 and 8 - The network is dense and centralized, though it still does not seem to be as diverse and well connected as the control (figure 5).

Parameters	Control	Month 0	Month 1	Month 2	Month 3	Month 4	Month 5 & 6	Month 7 & 8
# Nodes	103	66	60	34	83	36	65	61
# Edges	258	129	88	48	163	58	173	170
Average # of Neighbours	3.728	2.697	2.433	2.412	2.795	2.222	4.646	5.279
Connected Components	1	3	2	1	3	1	2	2
Network Centralisation	0.343	0.195	0.220	0.244	0.302	0.417	0.280	0.236

Table 2: A summary of various metrics taken from the generated networks

3.4 Network Metrics

In order to compare the graphs, a few key metrics were utilized. Average number of neighbours is the average number of connections a node has. The number of connected components describes the number of components in the graph where each node can be reached from another node in the component. Network centralization describes the extent to which a graph is organised around a few focal points. A summary of the results can be seen in Table 2.

The control network (Figure 5) has a high amount of average number of neighbours (3.728), one connected component, and a centralisation of 0.343, reflecting a highly connected network.

However, antibiotic intake (Figure 6) causes fragmentation in the immediately following month of consumption. This month is reflected by the number of distinct connected components increasing to 3, the average number of neighbours decreasing to 2.697, and centralisation decreases to 0.195. Network metrics generally show indications of returning to baseline values, but fluctuate over time. The number of connected components fluctuates between 1 and 3, ending in 2 in months 7 and 8 (Figure 12). Centralisation shows a weak upwards trend, reaching a peak of 0.417 in month 4 (Figure 10), but then decreasing to 0.236 in months 7-8 (Figure 12). The average number of neighbours shows a weak upwards trend, eventually increasing to 5.279 by months 7-8 (Figure 12), significantly greater than the baseline value.

3.5 Key Species and Relationships

3.5.1 Akkermansia

We further used co-occurrence analysis to determine potential key species. In particular, we considered *Akkermansia*, with the most connections (58) in the control network (Figure 5), but only a degree of 2 in month 0 (Figure 6). A graph of *Akkermansia* and its connections in the control network is shown in Figure 13, and a similar graph of *Akkermansia* in month 0 is shown in Figure 14.

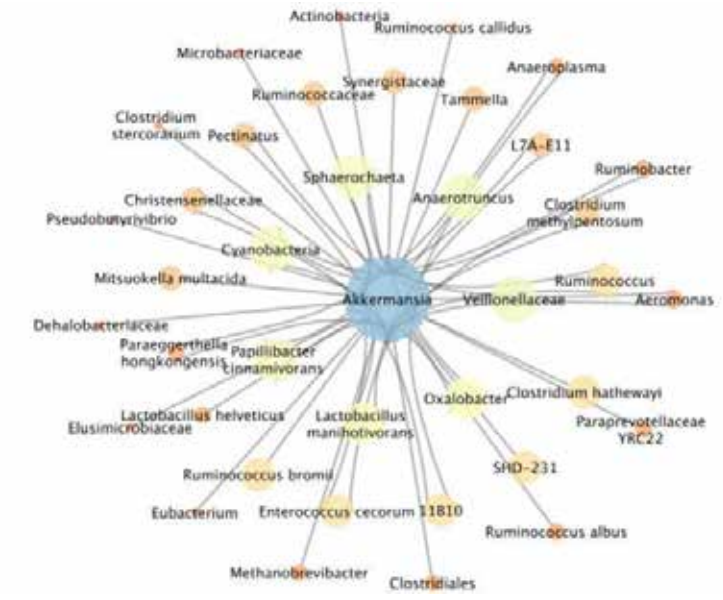


Figure 13. The degree of *Akkermansi*a in the control network (Figure 5)

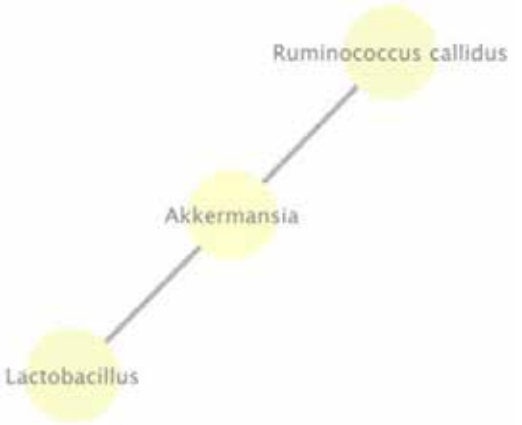


Figure 14. The degree of *Akkermansia* in the Month 0 network (Figure 6).

3.5.2 Methanogens

A competitive relationship between *Methanobrevibacter* and *Methanocorpusculum* was identified from the OTU tables. The normalized abundance, calculated by dividing the number of reads of each methanogen by the total number of reads in each sample, was used in Figure 15.

Co-occurrence analysis were also used to explore how *Methanobrevibacter* and *Methanocorpusculum* may compete for relationships, as shown in Figure 16.

Normalized Abundance of *Methanocorpusculum* vs *Methanobrevibacter*

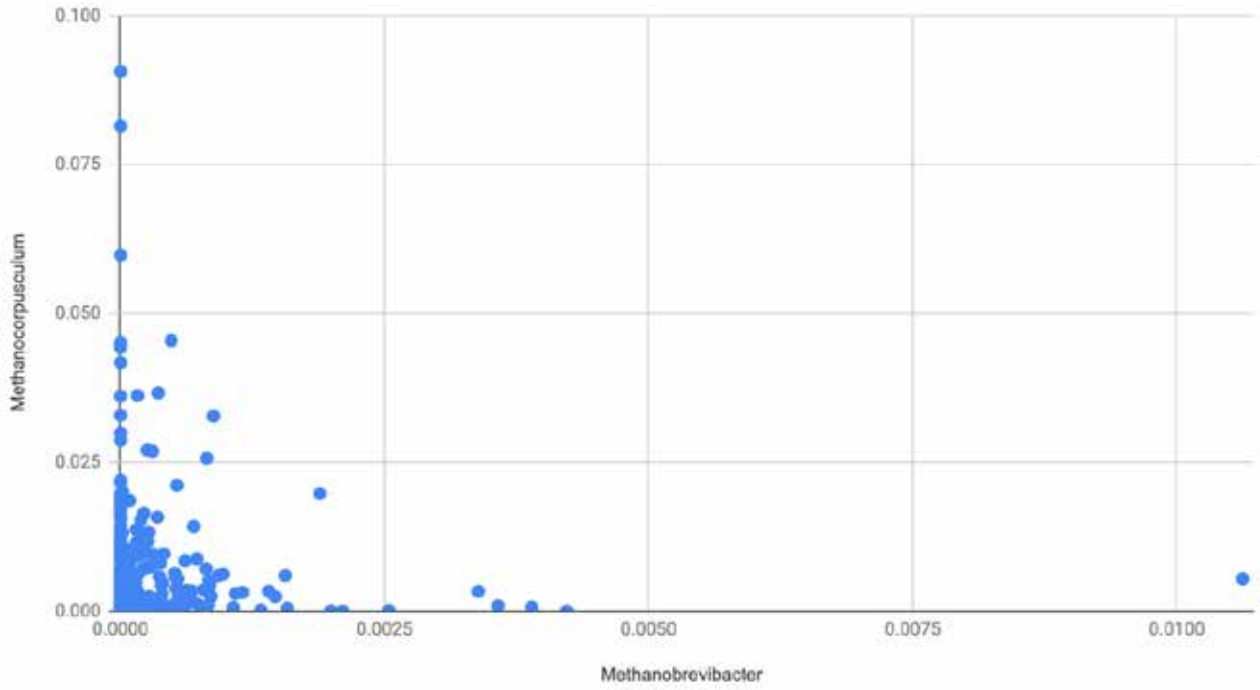


Figure 15. Normalized abundance of *Methanocorpusculum* versus the normalized abundance of *Methanobrevibacter*.

Degree of *Methanobrevibacter* versus degree of *Methanocorpusculum*

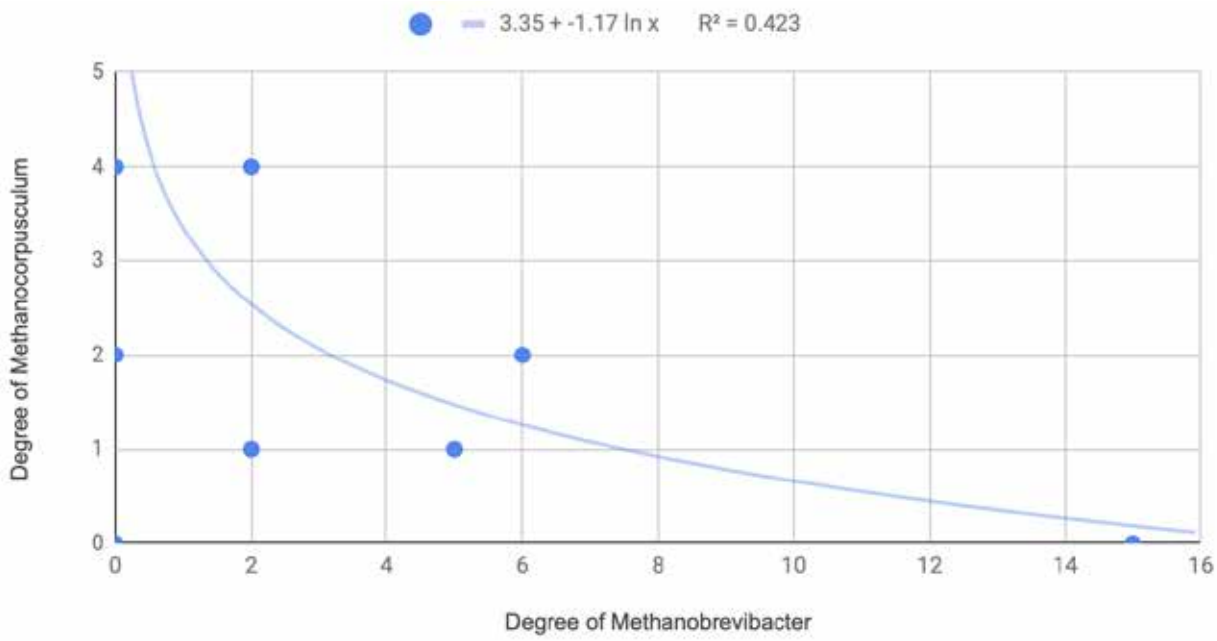


Figure 16. Degree (the number of edges connected to each node) of the Methanobrevibacter node versus the degree of the *Methanocorpusculum* node for each graph.

4. Discussion

4.1 Traditional Methods

4.1.1 Alpha Diversity

Referring to figure 2, we see that alpha diversity of the gut microbiome seems to fluctuate naturally over time, with no real apparent trend but within a narrow range.

Referring to figure 3, though fluctuations were present, antibiotic use also seems to have no visible impact on the mean Shannon index. With regards to the Shannon index of the equine gut microbiome after antibiotic treatment (Figure 3), H fluctuates within the range of 6-8, mirroring the behaviour seen in Figure 2, suggesting that any fluctuations were just natural fluctuations over time but with a slightly broader range.

We acknowledge that our Shannon index values are abnormally high, with the typical range being 1.5-3.5 (Magurran, 2013) and our range being 6-8. We attribute this to the Shannon index's tendency to overweight species present in small proportions as it tends to produce comparatively high numbers for rare species. Given that our samples contained a large number of species, many of which were only present in small amounts in a few samples, it is to be expected that our Shannon index values are higher than normal. However, despite its shortcomings, we still believe that the Shannon index provides a useful overview of the diversity of our samples, and still provides an insightful comparison across cohorts.

4.1.2 PCoA

As seen in figure 4, although deviation of data points was present, the deviation was seemingly randomly distributed throughout the months before and after antibiotic intake. As such, this deviation cannot be attributed towards antibiotic usage, and requires further investigation.

4.1.3 Traditional Methods - A Conclusion

Referring to sections 4.1.1 and 4.1.2, we see that traditional methods, for various reasons, fail to capture changes in the equine gut microbiome following antibiotic usage.

4.2 Network Connectivity

As shown in Section 3.4, antibiotic usage causes fragmentation in the networks, as reflected by changes in various network metrics. We hypothesize that this fragmentation is due to the disruptive impacts of antibiotics, which may have resulted in the diminishing or defauning of key species in the equine gut microbiota.

Network connectivity can be indicative of a healthy and resilient gut microbiome (Naqvi *et al*., 2010), especially because microbes rely on functions and substrates provided by other microbes to sustain themselves and propel their metabolic pathways [4]. Hence, a loss of connections may lead to the collapse of dependencies in a microbial network, resulting in a less healthy and less functional microbiome.

However, we must acknowledge that we are uncertain of what constitutes a "healthy" level. Although studies have shown that increased connectivity in a network can be used as an indicator of gut health (Naqvi *et al*., 2010), there is no consensus upon what this level may be. Here, we believe that what really constitutes as a healthy gut microbiome is an optimum range of values rather than a singular value. Further work, however, is needed to determine this range.

4.3 Key Species and Relationships

4.3.1 Key Species

Co-occurrence networks were used to observe noteworthy relationships within the gut microbiota. On a large scale, key microbes were identified with overall degree across the networks. Microbes from the family *Succinivibrionaceae* were consistently among the most connected nodes in the network, while the genera *Akkermansia* was the most connected node in the control network, as depicted in Figure 13. Interestingly, these microbes are not from either of the most abundant phyla in the gut microbiota, Firmicutes and Bacteroidetes, rather *Proteobacteria* and *Verrucomicrobia* (Rinninella *et al*., 2019). Hence we see abundance alone is not indicative of key relationships or dependencies.

The high degree of *Akkermansia* (58), *particularly A. muciniphila*, as shown in 13, can be attributed to its role as a key species in the gut microbiota, aiding the rest of the microbiome as one of the many mucin degrading specialists in the glycobiome. They produce monosaccharides and oligosaccharides, short-chain fatty acids (SCFA), and butyrate, while detoxifying hydrogen sulfide and preventing pathogen colonization (Ottman *et al*., 2017). *Akkermansia muciniphila* thereby fulfills its role as a supposed "feeder" and key microbe, while facilitating gut health. Furthermore, there is a link between the gut microbiome and the immune system of remote areas, such as the lungs, by producing molecules that transport via the bloodstream (Zhang *et al*., 2020). In particular, the production of SCFAs shows promise in treating viral infections and lung disease, such as Respiratory Syncytial Virus (Antunes *et al*., 2019), and is noted as a potential therapeutic agent to COVID-19 (Sarkar *et al*., 2020). *Akkermansia muciniphila*, however, is highly sensitive to changes in the environment, which may explain why it is no longer the most connected node following antibiotic treatment (Ottman *et al*., 2017), as shown by the sharp drop in connections depicted in figure 14. This further demonstrates the damage of antibiotics, in which, the loss of a key microbe significantly impacts the connectivity and health of the overall network.

4.3.2 Supporting Relationships

Co-occurrence networks were also utilized to uncover more specific relationships between microbes. In particular, relationships between methanogens – archaea that produce methane under hypoxic conditions – were investigated. The methanogens present in the networks were *Methanobrevibacter*, *Methanocorpusculum*, and *vadinCA11*. These methanogens were most often connected with two families: *Lachnospiraceae* and *Ruminococcaceae*. This is likely attributed to how *Lachnospiraceae* are butyrate producers, wherein butyrate is metabolized by colonocytes in the epithelium that produce CO_2 as a byproduct (Litvak *et al*., 2018). *Ruminococcaceae* is a fibrolytic microbe that produces CO_2 and H_2 (Piao *et al*., 2014). Methanogens subsist off of CO_2 and H_2 (Sowers, 2009), hence *Lachnospiraceae* and *Ruminococcaceae* potentially both act as "feeders" to the methanogens.

Further, methanogens tended to appear to be smaller, poorly connected nodes attached to larger "feeders" that also sustain other small microbes. In the "control" network (Figure 5), *Methanobrevibacter* was connected to *Akkermansia*, a key microbe. Hence, we see how co-occurrence networks are able to elucidate relationships between microbes, and suggest which microbes may be considered "key".

4.3.3 Competing Relationships

Potentially competitive relationships were explored with relation to both traditional and network analyses. In figure 15, the distribution of data points is mainly centred around the origin and spread along the two axes, similar to results in vitro (Guo *et al*., 2019). It can thus be inferred that *Methanobrevibacter* and *Methanocorpusculum* may be competing microbes, likely competing for the same resources (H_2 and CO_2, acetate, methanol, and formate) to survive. Further network-centric analysis has been used to determine the nature of their competition.

Referring to Figure 16, the degree of the two methanogens reveals a decaying logarithmic relationship, suggesting that the methanogens were also competing for connections within the community. However, the R2 value is only 0.423, suggesting that other factors that are currently uncounted for may also be affecting this relationship. However, while the methanogens shared some microbes, including *Lachnospiriceae* and *Ruminococcaceae*, they each had mostly unique sets of connections.

Co-occurrence network analysis is able to identify potential key microbes, detect valid connections between microbes, and demonstrate how microbes may compete with each other, showing its promise as a tool to better understand the nature and relationships of the gut microbiome. We also show how traditional analyses can be used in conjunction with network analyses to elucidate complex relationships.

However, since co-occurrence analysis is based on correlations inferred from relative abundance, it cannot provide reasons behind relationships and, on its own, cannot determine directionality. Nevertheless, the relationships it infers can be further investigated by tracer studies (using labelled substrates), such as in a study by Sung *et al*., 2017, and by metabolomic studies using imaging mass spectrometry (MALDI-TOF) (Singhal *et al*., 2015).

5. Conclusion

5.1 Main Findings

- Traditional methods (alpha diversity, beta diversity) failed to capture oral antibiotic-induced changes in the equine gut microbiota.
- Co-occurrence networks, however, demonstrate that oral antibiotics cause network fragmentation, with effects still visible in networks 7-8 months after antibiotic consumption. Fluctuations of network metrics reveal uncertain indications of recovery.
- Longitudinal sampling and network analysis helps to identify relationships between species: *e.g.* a possible competitive relationship between *Methanobrevibacter* and *Methanocorpusculum* is demonstrated, as well as competition for "feeder" species, such as *Lachnospiraceae* and *Ruminococcaceae*
- Influential microbes, such as *Akkermansia*, can be clearly identified with co-occurrence networks.

5.2 Further Work

We anticipate closer analysis of the networks with a larger sample size and smaller time frames (*i.e.* days or weeks). Given that literature suggests that the most prominent changes within the gut microbiome following antibiotic consumption are immediately after antibiotic intake (Costa *et al*., 2015), smaller time frames will help us better understand the connections and relationships, as well as directionality, within the network. Larger sample sizes and increased time frames would allow for investigation of potential hidden Markov networks, which might also reveal directionality within the networks (Harris, 2016).

5.3 Evaluations

In this paper, we have demonstrated how traditional approaches of profiling and understanding microbiomes through abundance or phylogenetic based approaches fail to capture the nuances of changes in the microbial relationships. Increasing amounts of research indicate that the gut microbiome plays crucial roles in health, with its impacts ranging from obesity (Le Chatelier *et al*., 2013) to modulating brain behaviour (Diaz Heijtz *et al*., 2011), to even possibly influencing one's immunity to COVID-19 (Dhar and Mohanty, 2020). Should we continue to only conduct simple compositional analyses, we may overlook the complex interactions between microbes, which has wide reaching impacts upon both human and non-human health.

Although co-occurrence analysis has obtained more prominence in other fields, especially in the social sciences (Chen *et al*., 2016), it is currently a relatively under-utilised method within the field of microbial ecology. In this paper, we have shown that co-occurrence networks have the potential to capture changes in microbiomes that traditional methods may ignore by analysing the changes in the equine gut microbiome following antibiotic usage. The method that we demonstrate, where we generate co-occurrence networks based on abundance, is one that is easy to implement, yet reveals great nuance about a microbiome. Thus, we see that the co-occurrence analysis approach demonstrated in this paper is one that has broad application and great potential in furthering our understanding of the complex realities of host-microbe and microbe-microbe interactions.

We recognize that, without a perfectly homogeneous sample, there will always be confounding factors, especially considering the range of parameters that might influence the behavior of each microbe.

Combined with the chaotic nature of ecology, it may seem like the conclusions we draw appear uncertain. In ecology, however, it is especially important and valid to focus on the big picture and the overarching patterns (Sanderson and Pimm, 2015). While network analysis cannot be used to accurately predict or model changes down to an individual level, we can infer and extrapolate overall changes in connections and relationships, as done in this study.

We also acknowledge that our approach is relatively simplistic as it only considers correlation, and thus does not provide reasons behind the relationships. However, we believe that it still provides a good overview of potential relationships within the microbiome, helping us better understand potential interactions, as well as where to focus future efforts on.

Nevertheless, with a growing trend towards "Big Data", it is imperative that we find new ways of efficiently finding relationships within large datasets.

In the future, we envision that the widespread applicability of network-based approaches, as well as the ease of implementation of such methods, will allow us to use them to further our understanding of host-microbe interactions, provide important insights to treatments on microbiome health, and push the frontiers of research in microbial ecology.

Acknowledgements

Special thanks to the Hong Kong Jockey Club for support and access to the data and medical records; Professor Frederick Leung for help with sequencing; and the *Shuyuan* Science team at The Independent Schools Foundation Academy, in particular Ms. Grace Lai and Dr. Simon Griffin, for their indispensable guidance throughout the entire project.

References

Al Jassim, R. A., Andrews, F. M. (2009). The bacterial community of the horse gastrointestinal tract and its relation to fermentative acidosis, laminitis, colic, and stomach ulcers. *The Veterinary Clinics of North America. Equine Practice*, 25(2), 199–215. doi: 10.1016/j.cveq.2009.04.005

Antunes, K. H., Fachi, J. L., de Paula, R., da Silva, E. F., Pral, L. P., Dos Santos, A. Á., Dias, G., Vargas, J. E., Puga, R., Mayer, F. Q., Maito, F., Zárate-Bladés, C. R., Ajami, N. J., Sant'Ana, M. R., Candreva, T., Rodrigues, H. G., Schmiele, M., Silva Clerici, M., Proença-Modena, J. L., Vieira, A. T., … de Souza, A. (2019). Microbiota-derived acetate protects against respiratory syncytial virus infection through a GPR43-type 1 interferon response. *Nature Communications*, 10(1): 3273. doi: 10.1038/s41467-019-11152-6

Bartlett, J. G. (1984). Antibiotic-associated colitis. *Disease-a-Month : DM*, 30(15), 1–54. PMID: 6391879

Båverud, V., Gustafsson, A., Franklin, A., Aspán, A., Gunnarsson, A. (2003). Clostridium difficile: prevalence in horses and environment, and antimicrobial susceptibility. *Equine Veterinary Journal*, 35(5), 465–471. doi: 10.2746/042516403775600505

Berg, G., Rybakova, D., Fischer, D., Cernava, T., Vergès, M. C., Charles, T., Chen, X., Cocolin, L., Eversole, K., Corral, G. H., Kazou, M., Kinkel, L., Lange, L., Lima, N., Loy, A., Macklin, J. A., Maguin, E., Mauchline, T., McClure, R., Mitter, B., … Schloter, M. (2020) Microbiome definition re-visited: old concepts and new challenges. *Microbiome*, 8(1): 103. doi: 10.1186/s40168-020-00875-0

Bokulich, N. A., Kaehler, B. D., Rideout, J. R., Dillon, M., Bolyen, E., Knight, R., Huttley, G. A., Gregory Caporaso, J. (2018). Optimizing taxonomic classification of marker-gene amplicon sequences with QIIME 2's q2-feature-classifier plugin. *Microbiome*, 6(1): 90. doi: 10.1186/s40168-018-0470-z

Bolyen, E., Rideout, J. R., Dillon, M. R., Bokulich, N. A., Abnet, C. C., Al-Ghalith, G. A., Alexander, H., Alm, E. J., Arumugam, M., Asnicar, F., Bai, Y., Bisanz, J. E., Bittinger, K., Brejnrod, A., Brislawn, C. J., Brown, C. T., Callahan, B. J., Caraballo-Rodríguez, A. M., Chase, J., Cope, E. K., … Caporaso, J. G. (2019). Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. *Nature Biotechnology*, 37(8), 852-857. doi: 10.1038/s41587-019-0209-9

Callahan, B. J., McMurdie, P. J., Rosen, M. J., Han, A. W., Johnson, A. J., Holmes, S. P. (2016). DADA2: High-resolution sample inference from Illumina amplicon data. *Nature Methods*, 13(7), 581–583. doi: 10.1038/nmeth.3869

Chen, X., Chen, J., Wu, D., Xie, Y., Li, J. (2016) Mapping the Research Trends by Co-word Analysis Based on Keywords from Funded Project, *Procedia Computer Science*, 91, 547-555. doi: 10.1016/j.procs.2016.07.140.

Costa, M. C., Silva, G., Ramos, R. V., Staempfli, H. R., Arroyo, L. G., Kim, P., Weese, J. S. (2015). Characterization and comparison of the bacterial microbiota in different gastrointestinal tract compartments in horses. *Veterinary Journal*, 205(1), 74-80. doi: 10.1016/j.tvjl.2015.03.018

Davies, S., Zadik, P. M., Mason, C. M., Whittaker, S. J. (2000). Methicillin-resistant Staphylococcus aureus: evaluation of five selective media. *British Journal of Biomedical Science*, 57(4), 269-272. PMID: 11204854

DeSantis, T. Z., Hugenholtz, P., Larsen, N., Rojas, M., Brodie, E. L., Keller, K., Huber, T., Dalevi, D., Hu, P., Andersen, G. L. (2006). Greengenes, a chimera-checked 16S rRNA gene database and workbench compatible with ARB. *Applied and Environmental Microbiology*, 72(7), 5069-5072. doi: 10.1128/AEM.03006-05

Deurenberg, R. H., Bathoorn, E., Chlebowicz, M. A., Couto, N., Ferdous, M., García-Cobos, S., Kooistra-Smid, A. M., Raangs, E. C., Rosema, S., Veloo, A. C., Zhou, K., Friedrich, A. W., Rossen, J. W. (2017). Application of next generation sequencing in clinical microbiology and infection prevention. *Journal of Biotechnology*, 243, 16-24. doi: 10.1016/j.jbiotec.2016.12.022

Dhar, D., Mohanty, A. (2020). Gut microbiota and Covid-19 – possible link and implications. *Virus Research*, 285, 198018. doi: 10.1016/j.virusres.2020.198018

Diaz Heijtz, R., Wang, S., Anuar, F., Qian, Y., Björkholm, B., Samuelsson, A., Hibberd, M. L., Forssberg, H., Pettersson, S. (2011). Normal gut microbiota modulates brain development and behavior. *PNAS*, 108(7), 3047–3052. doi:10.1073/pnas.1010529108

Dowd, S. E., Sun, Y., Secor, P. R., Rhoads, D. D., Wolcott, B. M., James, G. A., Wolcott, R. D. (2008). Survey of bacterial diversity in chronic wounds using pyrosequencing, DGGE, and full ribosome shotgun sequencing. *BMC Microbiology*, 8: 43. doi: 10.1186/1471-2180-8-43

Faust, K., Raes, J. (2016). CoNet app: inference of biological association networks using Cytoscape. *F1000Research*, 5: 1519. doi: 10.12688/f1000research.9050.2

Faust, K., Sathirapongsasuti, J. F., Izard, J., Segata, N., Gevers, D., Raes, J., Huttenhower, C. (2012). Microbial co-occurrence relationships in the human microbiome. *PLoS Computational Biology*, 8(7): e1002606. doi: 10.1371/journal.pcbi.1002606

Flayhart, D., Borek, A. P., Wakefield, T., Dick, J., Carroll, K. C. (2007). Comparison of BACTEC PLUS blood culture media to BacT/Alert FA blood culture media for detection of bacterial pathogens in samples containing therapeutic levels of antibiotics. *Journal of Clinical Microbiology*, 45(3), 816-821. doi: 10.1128/JCM.02064-06

Gilbert, J. A., Jansson, J. K., Knight, R. (2014). The Earth Microbiome project: successes and aspirations. *BMC Biology*, 12: 69. doi: 10.1186/s12915-014-0069-1

Grønvold, A. M., L'Abée-Lund, T. M., Strand, E., Sørum, H., Yannarell, A. C., Mackie, R. I. (2010). Fecal microbiota of horses in the clinical setting: potential effects of penicillin and general anesthesia. *Veterinary Microbiology*, 145(3-4), 366-372. doi: 10.1016/j.vetmic.2010.03.023

Guo, X., Silva, K., Boedicker, J. Q. (2019). Single-cell variability of growth interactions within a two-species bacterial community. *Physical Biology*, 16(3): 036001. doi: 10.1088/1478-3975/ab005f

Gupta, S., Mortensen, M. S., Schjørring, S., Trivedi, U., Vestergaard, G., Stokholm, J., Bisgaard, H., Krogfelt, K. A., Sørensen, S. J. (2019). Amplicon sequencing provides more accurate microbiome information in healthy children compared to culturing. *Communications Biology*, 2: 291. doi: 10.1038/s42003-019-0540-1

Hamady, M., Knight, R. (2009). Microbial community profiling for human microbiome projects: Tools, techniques, and challenges. *Genome Research*, 19(7), 1141-1152. doi: 10.1101/gr.085464.108

Harlow, B. E., Lawrence, L. M., Flythe, M. D. (2013). Diarrhea-associated pathogens, lactobacilli and cellulolytic bacteria in equine feces: responses to antibiotic challenge. *Veterinary Microbiology*, 166(1-2), 225–232. doi: 10.1016/j.vetmic.2013.05.003

Harris, D. J. (2016). Inferring species interactions from co-occurrence data with Markov networks. *Ecology*, 97(12), 3308-3314. doi: 10.1002/ecy.1605

Hedayat, K. M., Lapraz, J.-C. (2019) The Theory of Endobiogeny: Volume 2: Advanced Concepts for the Treatment of Complex *Clinical Conditions.* Academic Press.

Hooper, L. V., Gordon, J. I. (2001). Commensal host-bacterial relationships in the gut. *Science*, 292(5519), 1115–1118. doi: 10.1126/science.1058709.

Iebba, V., Totino, V., Gagliardi, A., Santangelo, F., Cacciotti, F., Trancassini, M., Mancini, C., Cicerone, C., Corazziari, E., Pantanella, F., Schippa, S. (2016). Eubiosis and dysbiosis: the two sides of the microbiota. *The New Microbiologica*, 39(1), 1-12. PMID: 269 22981

Jacobs, N. F., Jr (1994). Antibiotic-induced diarrhea and pseudomembranous colitis. Postgraduate Medicine, 95(8), 111-120. doi: 10.1080/00325481.1994.11945865

Jernberg, C., Löfmark, S., Edlund, C., Jansson, J. K. (2010). Long-term impacts of antibiotic exposure on the human intestinal microbiota. *Microbiology*, 156(Pt 11), 3216-3223. doi: 10.1099/mic.0.040618-0

Ji, S., Jiang, T., Yan, H., Guo, C., Liu, J., Su, H., Alugongo, G. M., Shi, H., Wang, Y., Cao, Z., Li, S. (2018). Ecological Restoration of Antibiotic-Disturbed Gastrointestinal Microbiota in Foregut and Hindgut of Cows. Frontiers in Cellular and Infection *Microbiology*, 8: 79. doi: 10.3389/fcimb.2018.00079

Katoh, K., Misawa, K., Kuma, K., Miyata, T. (2002). MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic Acids Research*, 30(14), 3059-3066. doi: 10.1093/nar/gkf436

Kauter, A., Epping, L., Semmler, T., Antao, E. M., Kannapin, D., Stoeckle, S. D., Gehlen, H., Lübke-Becker, A., Günther, S., Wieler, L. H., Walther, B. (2019). The gut microbiome of horses: current research on equine enteral microbiota and future perspectives. *Animal Microbiome*, 1(1): 14. doi: 10.1186/s42523-019-0013-3

Korpela, K., Salonen, A., Virta, L. J., Kekkonen, R. A., Forslund, K., Bork, P., de Vos, W. M. (2016). Intestinal microbiome is related to lifetime antibiotic use in Finnish pre-school children. *Nature Communications*, 7: 10410. doi: 10.1038/ncomms10410

Larsen, J. (1997). Acute colitis in adult horses. A review with emphasis on aetiology and pathogenesis. *The Veterinary Quarterly*, 19(2), 72-80. doi: 10.1080/01652176.1997.9694745

Le Chatelier, E., Nielsen, T., Qin, J., Prifti, E., Hildebrand, F., Falony, G., Almeida, M., Arumugam, M., Batto, J. M., Kennedy, S., Leonard, P., Li, J., Burgdorf, K., Grarup, N., Jørgensen, T., Brandslund, I., Nielsen, H. B., Juncker, A. S., Bertalan, M., Levenez, F., … Pedersen, O. (2013). Richness of human gut microbiome correlates with metabolic markers. *Nature*, 500(7464), 541–546. doi: 10.1038/nature12506

Lee, K. H., Gordon, A., Shedden, K., Kuan, G., Ng, S., Balmaseda, A., Foxman, B. (2019). The respiratory microbiome and susceptibility to influenza virus infection. *PLoS ONE*, 14(1): e0207898. doi: 10.1371/journal.pone.0207898

Litvak, Y., Byndloss, M. X., Bäumler, A. J. (2018). Colonocyte metabolism shapes the gut microbiota. *Science*, 362(6418): eaat9076. doi: 10.1126/science.aat9076

Lozupone, C. A., Hamady, M., Kelley, S. T., Knight, R. (2007). Quantitative and qualitative beta diversity measures lead to different insights into factors that structure microbial communities. *Applied and Environmental Microbiology*, 73(5), 1576-1585. doi: 10.1128/AEM.01996-06

Magurran, A. E. (2013) Measuring Biological Diversity. Wiley-Blackwell.

Manichanh, C., Reeder, J., Gibert, P., Varela, E., Llopis, M., Antolin, M., Guigo, R., Knight, R., Guarner, F. (2010). Reshaping the gut microbiome with bacterial transplantation and antibiotic intake. *Genome Research*, 20(10), 1411-1419. doi: 10.1101/gr.107987.110

McKenney, P. T., Pamer, E. G. (2015). From Hype to Hope: The Gut Microbiota in Enteric Infectious Disease. *Cell*, 163(6), 1326-1332. doi: 10.1016/j.cell.2015.11.032

Naqvi, A., Rangwala, H., Keshavarzian, A., Gillevet, P. (2010). Network-based modeling of the human gut microbiome. *Chemistry & Biodiversity*, 7(5), 1040-1050. doi: 10.1002/cbdv.200900324

Nobel, Y. R., Cox, L. M., Kirigin, F. F., Bokulich, N. A., Yamanishi, S., Teitler, I., Chung, J., Sohn, J., Barber, C. M., Goldfarb, D. S., Raju, K., Abubucker, S., Zhou, Y., Ruiz, V. E., Li, H., Mitreva, M., Alekseyenko, A. V., Weinstock, G. M., Sodergren, E., Blaser, M. J. (2015). Metabolic and metagenomic outcomes from early-life pulsed antibiotic treatment. *Nature Communications*, 6: 7486. doi: 10.1038/ncomms8486

Ottman, N., Geerlings, S. Y., Aalvink, S., de Vos, W. M., Belzer, C. (2017). Action and function of Akkermansia muciniphila in microbiome ecology, health and disease. Best practice & research. *Clinical Gastroenterology*, 31(6), 637-642. doi: 10.1016/j.bpg.2017.10.001

Piao, H., Lachman, M., Malfatti, S., Sczyrba, A., Knierim, B., Auer, M., Tringe, S. G., Mackie, R. I., Yeoman, C. J., Hess, M. (2014). Temporal dynamics of fibrolytic and methanogenic rumen microorganisms during in situ incubation of switchgrass determined by 16S rRNA gene profiling. *Frontiers in Microbiology*, 5: 307. doi: 10.3389/fmicb.2014.00307

Poretsky, R., Rodriguez-R, L. M., Luo, C., Tsementzi, D., Konstantinidis, K. T. (2014). Strengths and limitations of 16S rRNA gene amplicon sequencing in revealing temporal microbial community dynamics. *PLoS ONE*, 9(4): e93827. doi: 10.1371/journal.pone.0093827

Price, M. N., Dehal, P. S., Arkin, A. P. (2010). FastTree 2 – approximately maximum-likelihood trees for large alignments. *PloS ONE*, 5(3): e9490. doi: 10.1371/journal.pone.0009490

Rhoads, D. D., Cox, S. B., Rees, E. J., Sun, Y., Wolcott, R. D. (2012). Clinical identification of bacteria in human chronic wound infections: culturing vs. 16S ribosomal DNA sequencing. *BMC Infectious Diseases*, 12: 321. doi: 10.1186/1471-2334-12-321

Riesenfeld, C. S., Schloss, P. D., Handelsman, J. (2004). Metagenomics: genomic analysis of microbial communities. *Annual Review of Genetics*, 38, 525-552. doi: 10.1146/annurev.genet.38.072902.091216

Rinninella, E., Raoul, P., Cintoni, M., Franceschi, F., Miggiano, G., Gasbarrini, A., Mele, M. C. (2019). What is the Healthy Gut Microbiota Composition? A Changing Ecosystem across Age, Environment, Diet, and Diseases. *Microorganisms*, 7(1): 14. doi: 10.3390/microorganisms7010014

Sanderson, J. G. and Pimm, S. L. (2015) Patterns in Nature: the analysis of species co-occurrences, The University of Chicago Press.

Sarkar, P., Borah, S., and Sharma, H. K. (2020) Can microbial SCFA, butyrate be the alternate savior against COVID-19?, *Current Trends in Pharmaceutical Research*, 7(1): 07.

Shannon, C. E. and W. Weaver, W. (1963) A Mathematical Theory of Communication. University Of Illinois Press.

Singhal, N., Kumar, M., Kanaujia, P. K., Virdi, J. S. (2015). MALDI-TOF mass spectrometry: an emerging technology for microbial identification and diagnosis. *Frontiers in Microbiology*, 6: 791. doi: 10.3389/fmicb.2015.00791

Sowers, K. (2009) Methanogenesis. In Schmidt, T. M. (Ed.) Encyclopedia of Microbiology (Fourth Edition), *Academic Press*; 53-73. doi: 10.1016/B978-0-12-801238-3.02447-8

Sung, J., Kim, S., Cabatbat, J., Jang, S., Jin, Y. S., Jung, G. Y., Chia, N., Kim, P. J. (2017). Global metabolic interaction network of the human gut microbiota for context-specific community-scale analysis. *Nature Communications*, 8: 15393. doi: 10.1038/ncomms15393

Venable, E. B., Bland, S.D., McPherson, J.L., Francis, J. (2016). Role of the gut microbiota in equine health and disease. *Animal Frontiers*, 6(3), 43–49. doi: 10.2527/af.2016-0033

Wolcott, R. D., Dowd, S. E. (2008). A rapid molecular method for characterising bacterial bioburden in chronic wounds. *Journal of Wound Care*, 17(12), 513-516. doi: 10.12968/jowc.2008.17.12.31769

Young, V. B. (2017). The role of the microbiome in human health and disease: an introduction for clinicians. *BMJ (Clinical Research ed.),* 356: j831. doi: 10.1136/bmj.j831.

Zapka, C., Leff, J., Henley, J., Tittl, J., De Nardo, E., Butler, M., Griggs, R., Fierer, N., Edmonds-Wilson, S. (2017). Comparison of Standard Culture-Based Method to Culture-Independent Method for Evaluation of Hygiene Effects on the Hand Microbiome. *mBio*, 8(2): e00093-17. doi: 10.1128/mBio.00093-17

Zhang, D., Li, S., Wang, N., Tan, H. Y., Zhang, Z., Feng, Y. (2020). The Cross-Talk Between Gut Microbiota and Lungs in Common Lung Diseases. *Frontiers in Microbiology*, 11: 301. doi: 10.3389/fmicb.2020.00301

Horizontal Gene Transfer and Genomic Inversion shape an *E. coli* Evolution in the Presence of *Kosakonia Cowanii*

Wilhelmina Evelyn Moore

Abstract

E. coli typically thrives only in enteric environments. However, horizontal gene transfer can allow normally enteric bacteria to acquire halotolerant genes and characteristics, thus being able to survive in seawater. This study of seawater microbes had identified three halotolerant strains, including one *E. coli* containing halotolerance genes homologous to those in *Kosakonia cowanii.* Full genome sequencing enabled an exploration of aspects of this adaptation. In this study, we show that horizontal gene transfer and genomic inversions can cause normally enteric bacteria such as *E. coli* to develop characteristics that allow them to survive in new environments.

1. Background

Horizontal gene transfer is the process in which genetic information, in the form of plasmids, is transferred between bacteria. Horizontal gene transfer can often be a mechanism for bacteria to adapt to previously uninhabitable environments, which can also be a route to pathogenicity, as found by Maruyama *et al*. (2009), Lara-Ramírez *et al*. (2011), and Diene *et al*. (2012). Horizontal gene transfer as a mechanism for environmental adaptation has been previously examined by Cui *et al*. (2012) and also observed in *Streptococcus mutans* (Maruyama *et al*., 2009) and in *Helicobacter pylori* (Lara-Ramírez *et al*., 2011). This work specifically focuses on how horizontal gene transfer allows bacteria to become halotolerant. Halotolerance is the ability for bacteria to survive in environments with high concentrations of salt (such as seawater), which are usually uninhabitable. Halotolerance has previously been linked to several genes that control cell membrane proteins, and it is possible that halotolerance can be acquired by the transferring of these genes between bacteria. This could happen through the route of phage-mediated horizontal gene transfer. The existence of prophage sequences (a genome sequence inserted by a bacteriophage) could indicate that phage-mediated gene transfer has occurred.

2. Methods

2.1 Sample Isolation and Identification

Samples of seawater were collected from Repulse Bay, Hong Kong and screened for *E. coli* and coliforms using 3M Petrifilm® plates. Chosen isolates were passaged 8 times on LB agar and following gDNA extraction, shotgun sequencing was performed via the Illumina MiSeq NGS platform and species characterized using NCBI BLAST. Full genome sequences were assembled by combining Nanopore MinIon long-reads with the Illumina short-reads using Unicycler (Wick *et al.*, 2017).

2.2 Genomic Analysis

KAAS (Moriya *et al*., 2007) was used to map KEGG pathways. Likely prophage sequences were identified by PHASTER. NCBI BLAST and AutoMLST (Alanjary *et al*., 2019) were used for alignment and to map phylogeny. Synmap (Haug-Baltzell *et al*., 2017) was used to create synteny maps. GeVo (Lyons & Freeling, 2008), Clustal Omega (Madeira *et al*., 2019), Blast ATLAS (Buckingham *et al*., 2010), and PATRIC (Wattam *et al*., 2016) were used for multiple sequence alignment comparisons. Ori-Finder 2 (Luo *et al*., 2014) was used to find the origin of replication.

The research for this article was conducted in the *Shuyuan* Molecular Biology Laboratory and a poster about this work was prepared for the ASM Microbe conference in 2020.

2.3 Halotolerance Testing

Halotolerance was investigated by measuring growth curves (OD_{600} via BMG CLARIOstar Plus) at a range of NaCl concentrations: isolates were grown in LB medium containing 1%, 3.5%, 5%, 7.5% or 10% w/v NaCl.

3. Result

3.1 Identification of Strains

NCBI BLAST identified WEM22 and WEM41 as *Kosakonia cowanii* with a homology of 86% and 96% respectively. WEM25 was shown to be *E. coli* with a homology of 100%. The DNA sequence of WEM22 has a plasmid located which was identified by BLAST as related to the *Kosakonia cowanii* Esp_Z (*Kco_Z*) genome with a homology of 51%.

3.2 Identification of Prophages

All three isolates contain multiple intact prophages and WEM22, which is more halotolerant, contains more prophages than the less halotolerant WEM41. WEM25, the *E. coli*, has the most prophages.

Isolate	Status	Prophage
22	Intact	PHAGE_Escher_HK75_NC_016160(13)
22	Intact	PHAGE_Escher_HK639_NC_016158(14)
22	Intact	PHAGE_Salmon_SEN1_NC_029003(11)
25	Intact	PHAGE_Salmon_SEN34_NC_028699(33)
25	Intact	PHAGE_Entero_SfI_NC_027339(24)
25	Intact	PHAGE_Salmon_SSU5_NC_018843(92)
25	Intact	PHAGE_Escher_pro483_NC_028943(27)
25	Intact	PHAGE_Escher_RCS47_NC_042128(6)
41	Intact	PHAGE_Escher_HK75_NC_016160(13)
41	Intact	PHAGE_Salmon_SEN1_NC_029003(11)

Table 1. Intact prophages identified in each isolate.

Figure 1. Osmotic upshift pathway responsible for halotolerance (captured from KEGG)

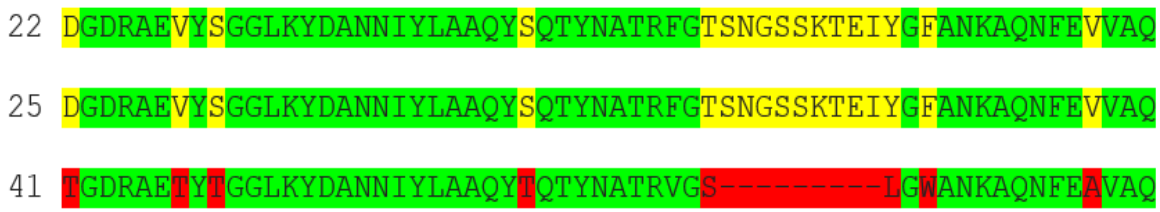


Figure 2. Comparison between the OmpC protein sequence of three isolates

3.3 KEGG Pathway Analysis

For all three isolates, KEGG analysis shows the presence of the osmotic upshift pathway via *EnvZ* and *OmpR*, which regulates the activity of the membrane protein gene products of *OmpC* and *OmpF*.

3.4 Protein Sequence Alignment

Comparisons of *OmpC* in the three isolates show a greater degree of homology between WEM22 (*K. cowanii*) and WEM25 (*E. coli*), while WEM41 (*K. cowanii*) has less in common than either two sequences.

3.5 Halotolerance

All three strains have the ability to grow in NaCl concentrations up to 7.5% (around twice the salinity of seawater), but exhibit different growth characteristics over this range. Although both *Kosakonia* strains have a strong rate of growth in NaCl concentrations up to 5% w/v, WEM22 grows moderately in 7.5% w/v whilst WEM41 grows more slowly. *E. coli* WEM25, on the other hand, grows moderately at NaCl concentrations of 3.5% and 5% w/v while at 7.5% it grows at a similar rate to *Kosakonia* WEM41.

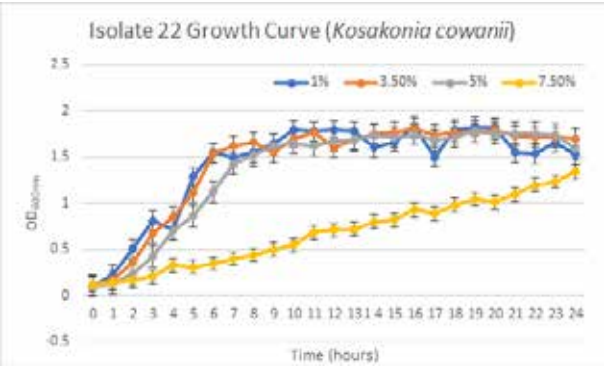


Figure 3. Growth curve of WEM22 in different concentrations of saltwater

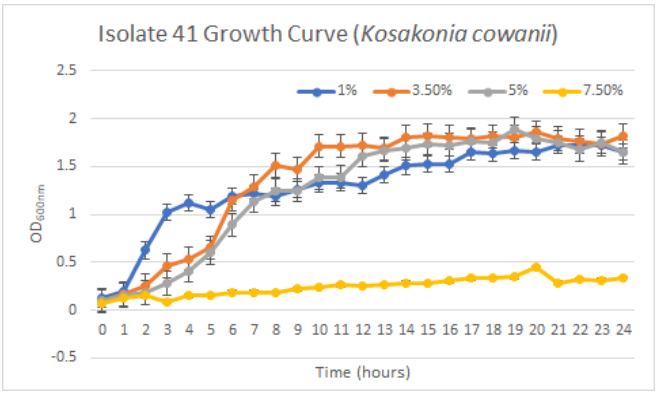


Figure 4. Growth curve of WEM25 in different concentrations of saltwater

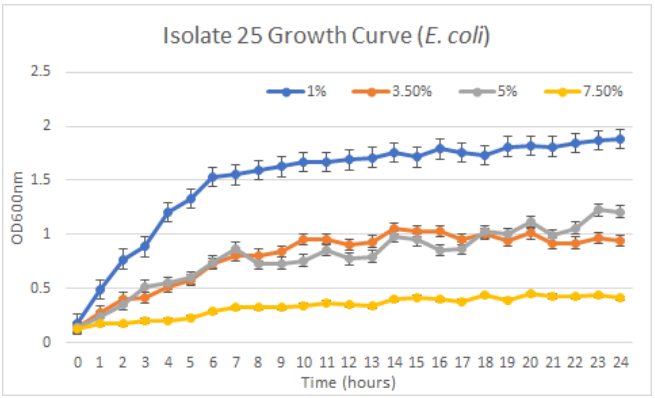


Figure 5. Growth curve of WEM41 in different concentrations of saltwater

3.6 Phylogeny

Phylogenetic trees show both *K. cowanii* 22 and *K. cowanii* 41 to be closely related and, surprisingly, *E. coli* 25 to be closer to these than to generic K12 strain *E. coli*.

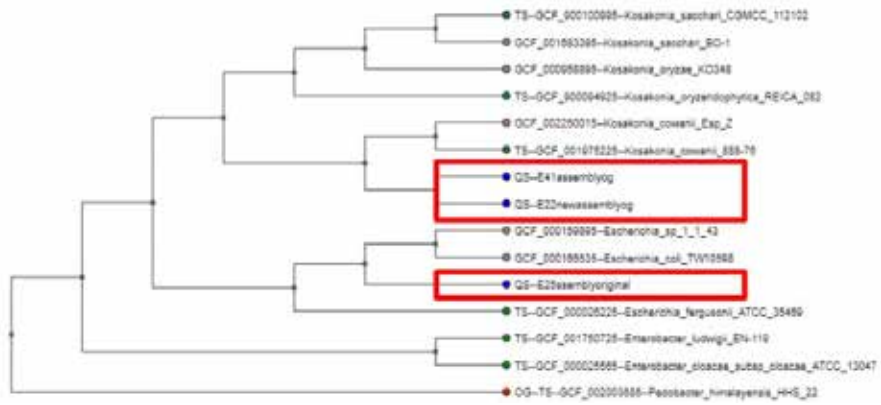


Figure 6. AutoMLST phylogenetic tree between WEM22, WEM25, and WEM41

3.7 Synteny Mapping

Synteny mapping not only reveals that the *E. coli* isolate is closer to *Kosakonia* than to the generic K12 *E. coli* but also shows an 'X' pattern characteristic of a large-scale genomic inversion.

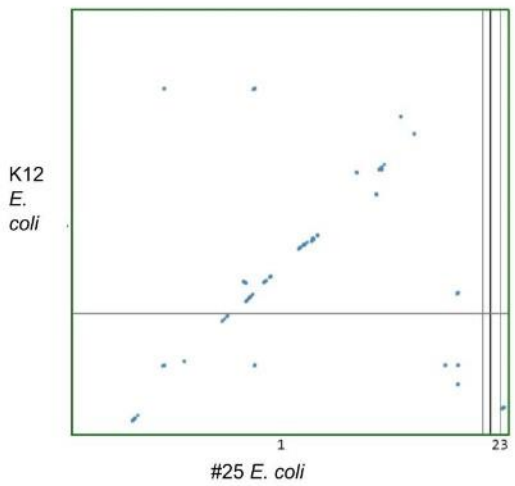


Figure 7. WEM25 and K12 E. coli synteny map

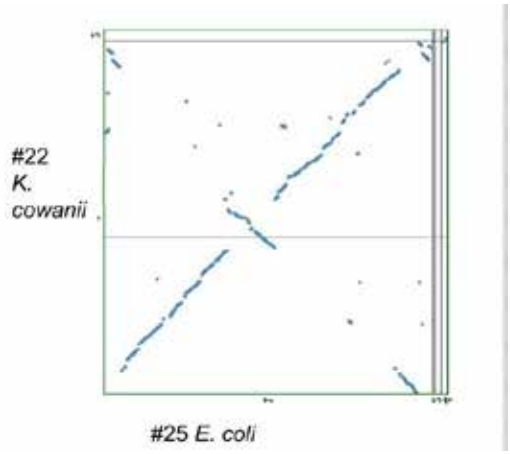


Figure 8. WEM25 and WEM22 synteny map

3.8 GeVo Analysis

GeVo was used to analyze the ends of the inversion site on the synteny maps, finding multiple repeat regions in the *E. coli* and the presence of transposase near the inversion sites. The origin of the major inversion in *E. coli* and *Kosakonia* WEM22 is centered around an origin near the excinuclease ABC subunit A (*uvrA*) locus. Similar results were found in the *Kosakonia WEM41* and *E. coli* inversion. This inversion was centered around an origin of replication near Putative sulfate permease in the *E. coli*. Again, there were repeating regions in the *E. coli* near the site of the inversion. Integrase was also found near the site. PATRIC (shown in Figures 12 and 13) shows the genomic inversion in more detail, showing the entire genomic inversion in the genome.

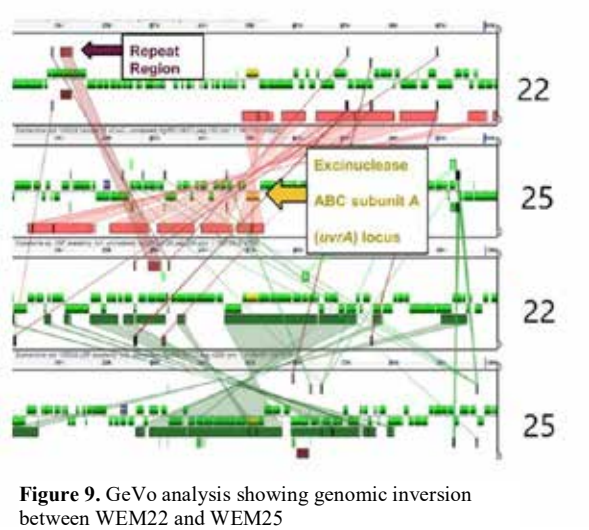


Figure 9. GeVo analysis showing genomic inversion between WEM22 and WEM25

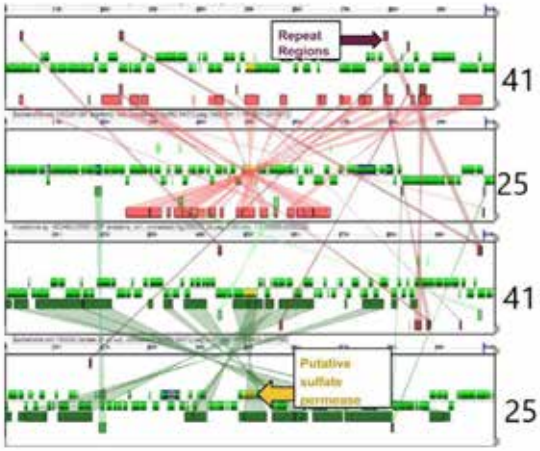


Figure 10. GeVo analysis showing genomic inversion between WEM25 and WEM41

3.9 PATRIC Analysis

PATRIC was also used to perform multiple sequence alignment, revealing areas of inversion within the whole genome and showing the places where key osmotolerance genes were exchanged. The osmotolerance two-component system regulator gene (*ompR*) is near the origin of replication (oriC), signaling its importance in the genome.

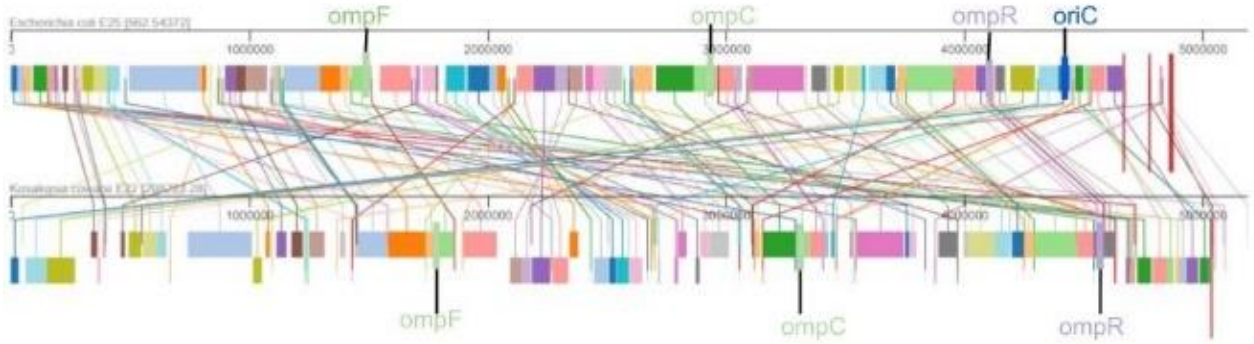


Figure 11. PATRIC genome alignment showing inversion between WEM25 and WEM22

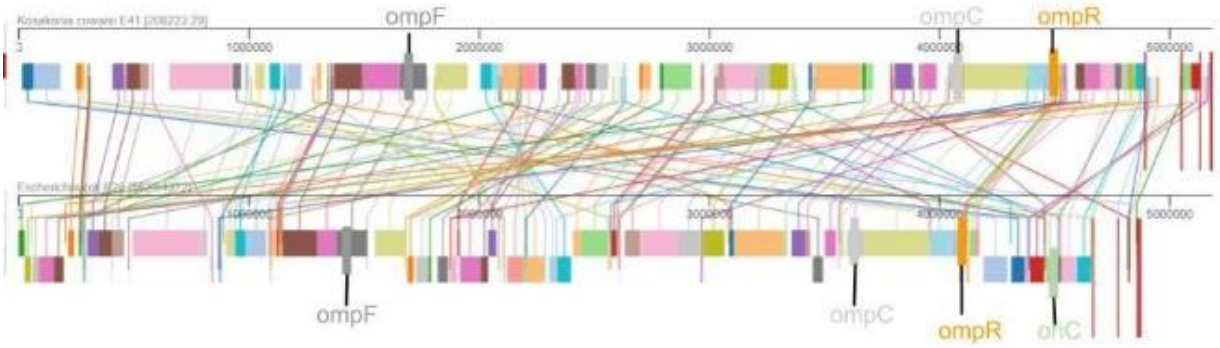


Figure 12. PATRIC genome alignment showing inversion between WEM25 and WEM41

Figure 13. WEM25 BLAST alignment against *E. coli* (red) and *Kosakonia* (green)

3.10 BLAST

BLAST revealed a large number of *Kosakonia* genes in the *E. coli*, many of which code for proteins relevant to osmotolerance such as GBT proteins, *ompA*, and potassium uptake proteins such as *TrkA*. Furthermore, other survival-related proteins such as drug resistance, heat and cold shock proteins were also identified.

4. Conclusion

Diene *et al*. (2012) describe the evolution of an *Enterobacter aerogenes* showing "genomic mosaicism" very similar to the case of *E. coli* isolate 25 here, with *Kosakonia* genes and operons widely distributed throughout the genome.

In addition, a large-scale genomic inversion may place emphasis on survival-related genes, such as transporters linked to halotolerance. This type of behavior was previously examined by Cui *et al*. (2012) and also observed in *Streptococcus mutans* (*Maruyama et al*., 2009) and in *Helicobacter pylori* (Lara-Ramírez *et al.,* 2011) as a mechanism for adaptation.

It is perhaps significant that Maruyama *et al*. (2009) and Lara-Ramírez *et al*. (2011), as well as Diene *et al.* (2012), link these genomic changes to virulence, so that the mechanisms driving pathogenicity and environmental adaptation may be one and the same.

Acknowledgments

Many thanks to Professor Fred Leung for his support in this project. Special thanks to the *Shuyuan* science team, in particular, Dr. Simon Griffin and Ms. Grace Lai.

Reference

Alanjary, M., Steinke, K., Ziemert, N. (2019). AutoMLST: an automated web server for generating multi-locus species trees highlighting natural product potential. *Nucleic Acids Research*, 47(W1), W276–W282. doi.org/10.1093/nar/gkz282

Arndt, D., Grant, J. R., Marcu, A., Sajed, T., Pon, A., Liang, Y., Wishart, D. S. (2016). PHASTER: a better, faster version of the PHAST phage search tool. *Nucleic Acids Research,* 44(W1), W16–W21. doi.org/10.1093/nar/gkw387

Buckingham, L., Hogan, J., Mann, S., Wirges, S. (2010). BLAST Atlas : a function based multiple genome browser. *Eprints.Qut.Edu.Au*. eprints.qut.edu.au/38334/

Clausen, P. T. L. C., Aarestrup, F. M., Lund, O. (2018). Rapid and precise alignment of raw reads against redundant databases with KMA. *BMC Bioinformatics*, 19(1). doi.org/10.1186/s12859-018-2336-6

Clausen, P. T. L. C., Zankari, E., Aarestrup, F. M., Lund, O. (2016). Benchmarking of methods for identification of antimicrobial resistance genes in bacterial whole genome data. *Journal of Antimicrobial Chemotherapy,* 71(9), 2484–2488. doi.org/10.1093/jac/dkw184

Cui, L., Neoh, H. -m., Iwamoto, A., Hiramatsu, K. (2012). Coordinated phenotype switching with large-scale chromosome flip-flop inversion observed in bacteria. *Proceedings of the National Academy of Sciences*, 109(25), E1647–E1656. doi.org/10.1073/pnas.1204307109

Diene, S. M., Merhej, V., Henry, M., El Filali, A., Roux, V., Robert, C., Azza, S., Gavory, F., Barbe, V., La Scola, B., Raoult, D., Rolain, J.-M. (2012). The Rhizome of the Multidrug-Resistant *Enterobacter* aerogenes Genome Reveals How New "Killer Bugs" Are Created because of a Sympatric Lifestyle. *Molecular Biology and Evolution*, 30(2), 369–383. doi.org/10.1093/molbev/mss236

Haug-Baltzell, A., Stephens, S. A., Davey, S., Scheidegger, C. E., Lyons, E. (2017). SynMap2 and SynMap3D: web-based whole-genome synteny browsers. *Bioinformatics*, 33(14), 2197–2198. doi.org/10.1093/bioinformatics/btx144

Lara-Ramírez, E. E., Segura-Cabrera, A., Guo, X., Yu, G., García-Pérez, C. A., Rodríguez-Pérez, M. A. (2011). New Implications on Genomic Adaptation Derived from the Helicobacter pylori Genome Comparison. *PLoS ONE*, 6(2), e17300. doi.org/10.1371/journal.pone.0017300

Luo, H., Zhang, C.-T., Gao, F. (2014). Ori-Finder 2, an integrated tool to predict replication origins in the archaeal genomes. *Frontiers in Microbiology*, 5(428). doi.org/10.3389/fmicb.2014.00482

Lyons, E., Freeling, M. (2008). How to usefully compare homologous plant genes and chromosomes as DNA sequences. *The Plant Journal,* 53(4), 661–673. doi.org/10.1111/j.1365-313x.2007.03326.x

Madeira, F., Park, Y., Lee, J. (2019). The EMBL-EBI search and sequence analysis tools APIs in 2019. *Nucleic Acids Research*, 47(1), 636–641. doi.org/10.1093/nar/gkz268

Maruyama, F., Kobata, M., Kurokawa, K., Nishida, K., Sakurai, A., Nakano, K., Nomura, R., Kawabata, S., Ooshima, T., Nakai, K., Hattori, M., Hamada, S., Nakagawa, I. (2009). Comparative genomic analyses of Streptococcus mutans provide insights into chromosomal shuffling and species-specific content. *BMC Genomics*, 10(1), 358. doi.org/10.1186/1471-2164-10-358

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

Munro, P. M., Gauthier, M. J., Breittmayer, V. A., Bongiovanni, J. (1989). Influence of osmoregulation processes on starvation survival of Escherichia coli in seawater. *Applied and Environmental Microbiology*, 55(8), 2017–2024. doi.org/10.1128/aem.55.8.2017-2024.1989

Smith, K. C., Castro-Nallar, E., Fisher, J. N., Breakwell, D. P., Grose, J. H., Burnett, S. H. (2013). Phage cluster relationships identified through single gene analysis. *BMC Genomics*, 14(1), 410. doi.org/10.1186/1471-2164-14-410

Tokishita, S., Yamada, H., Aiba, H., Mizuno, T. (1990). Transmembrane Signal Transduction and Osmoregulation in Escherichia coli: II. The Osmotic Sensor, EnvZ, Located in the Isolated Cytoplasmic Membrane Displays Its Phosphorylation and Dephosphorylation Abilities as to the Activator Protein, OmpR1. *The Journal of Biochemistry*, 108(3), 488–493. doi.org/10.1093/oxfordjournals.jbchem.a123226

Wang, X., Kim, Y., Ma, Q., Hong, S. H., Pokusaeva, K., Sturino, J. M., Wood, T. K. (2010). Cryptic prophages help bacteria cope with adverse environments. *Nature Communications*, 1(1). doi.org/10.1038/ncomms1146

Wattam, A. R., Davis, J. J., Assaf, R., Boisvert, S., Brettin, T., Bun, C., Conrad, N., Dietrich, E. M., Disz, T., Gabbard, J. L., Gerdes, S., Henry, C. S., Kenyon, R. W., Machi, D., Mao, C., Nordberg, E. K., Olsen, G. J., Murphy-Olson, D. E., Olson, R., … Stevens, R. L. (2016). Improvements to PATRIC, the all-bacterial Bioinformatics Database and Analysis Resource Center. *Nucleic Acids Research*, 45(D1), D535–D542. doi.org/10.1093/nar/gkw1017

Wick, R. R., Judd, L. M., Gorrie, C. L., Holt, K. E. (2017). Unicycler: Resolving bacterial genome assemblies from short and long sequencing reads. *PLOS Computational Biology*, 13(6), e1005595. doi.org/10.1371/journal.pcbi.1005595

Aminoglycoside Resistance in *Salmonella enterica* Isolated from Crested Geckos (*Correlophus ciliatus*)

Ines Belza-Garcia

Abstract

Salmonella enterica was identified and isolated from five crested geckos (*Correlophus ciliatus*) to determine its prevalence and also its pathogenicity. Eight isolates were found to possess virulence characteristics as well as aminoglycoside and other antibiotic resistance genes, suggesting that pet reptiles such as crested geckos can represent a risk of zoonotic infection. Nevertheless, it was noticed that while the *aac(6')*-laa aminoglycoside resistance gene was present in all isolates, none of the isolates presented this phenotypically as all were found to be susceptible to kanamycin. However, increased resistance to kanamycin could be induced by repeated exposure and resistant isolates produced in this way were subsequently characterised using genomic sequencing.

1. Introduction

Zoonoses, or diseases transferred from animals to humans, have long been a problem for human society. For example, SARS and COVID-19 outbreaks are both linked to viruses found in bats (Wang *et al*., 2006). Another example of zoonoses is in armadillos, who are known to act as reservoirs of *Mycobacterium leprae,* the bacterium that causes leprosy (da Silva *et al*., 2018). As the companionship of a bigger range of exotic animal pets has grown, the potential for zoonotic infections at home has increased (CDC, 2017).

In Hong Kong, reptiles, such as tortoises, geckos and even snakes are popular household pets because they require little space and exercise. Yet the close contact between humans and animals occurring in these settings also facilitates the transmission of zoonoses such as *Salmonella*.

Salmonella is a Gram-negative enteric coliform bacterium from the family *Enterobacteriaceae*, strains of which may be pathogenic to humans and other vertebrates. In humans, symptoms of salmonellosis include nausea, vomiting, diarrhoea, fever and even death (Shi *et al*., 2015; Giannella, 1996). However, birds and reptiles, including the crested gecko (*Correlophus ciliatus*), may act as reservoirs of the bacterium without presenting any symptoms of salmonellosis (CFSPH, 2013). If *Salmonella* is endemic in reptiles kept as pets, infection in humans could occur through direct contact with the animals or after direct or indirect contact with their faeces.

Antibiotic resistance in *Salmonella* makes salmonellosis harder to treat and is a growing problem worldwide. In their recent study into the spread of the multidrug resistant *Salmonella enterica* serovar Goldcoast in Taiwan, Liao *et al*. (2019) showed the *aadA* and *aac(6')*-laa genes – which give resistance to aminoglycosides, such as gentamicin and kanamycin – were present in all of the 128 isolates reported worldwide. Aminoglycosides are commonly used against Gram-negative bacteria and have previously been recommended as a useful alternative when β-lactams (*i.e.* penicillins) and fluoroquinolones (*e.g.* ciprofloxacin) need to be avoided (for instance due to allergy) and when bacteria are already resistant to more common antibiotics (Goodlet *et al*., 2018).

The research for this article was conducted in the *Shuyuan* Molecular Biology Laboratory and posters about this work were prepared for the AGU (American Geophysical Union) Fall Meeting in 2019 and the ASM Microbe conference in 2020.

Sample ID	1	2	3	4	5	6	7	8
Sample Source	Gecko 1 - Household	Gecko 1 - Household	Gecko 1 - Household	Gecko 1 - Household	Gecko 2 - petting zoo	Gecko 3 - petting zoo	Gecko 4 - petting zoo	Gecko 5 - petting zoo
No. of reads (aligned)	161335	445107	207363	726437	745632	283745	718146	477002
Average depth	9.0	26.0	11.0	38.0	45.0	17.0	44.0	29.0
No. of large contigs	1036	188	913	328	275	606	293	292
No. of large bases	4477059	4673854	4820386	5004002	4712716	4595122	4708153	4699886
N50 contig size	7077	53221	9626	30531	33756	14652	31354	32611
No. of all contigs	1246	592	1196	671	337	683	351	345
No. of all bases	4544349	4780035	4907920	5098483	4729586	4621761	4725518	4714750
Estimated genome size	5.4 MB	5.1 MB	5.6 MB	5.5 MB	4.9 MB	5.0 MB	4.7 MB	5.2 MB
Homology %	99.70%	99%	99.65%	97.95%	99.05%	99.95%	99.85%	99.50%
Identity	*Salmonella enterica*	*Salmonella enterica*	*Salmonella enterica*	*Salmonella enterica*	*Salmonella enterica*	*Salmonella enterica*	*Salmonella enterica*	*Salmonella enterica*
Serotype	*Arechavaleta*	*Tinda / Arechavaleta*	*II 35:m,t:- / Monschaui*	N/A	*Virginia / Muenchen*	*Virginia / Muenchen*	*Virginia / Muenchen*	*Virginia / Muenchen*

Figure 1. High quality draft sequences of the eight isolates and information on sequenced DNA material

[References and links to these on-line analytical platforms are given as supplementary data – please see the QR code at the end of this article.]

2. Methods

2.1 Isolation

Faecal samples were collected from five crested geckos: one household pet and four from a pet center in Hong Kong. Four *Salmonella* isolates from the household pet and one from each pet shop gecko were identified to be *salmonella* using CHROMagar Salmonella® and CHROMagar *Salmonella* Plus® agar plates. The chosen isolates were passaged ten times on LB agar and Gram-staining was used to confirm bacterial identity and purity.

2.2 Identification of Isolates

An Invitrogen™ PureLink™ Genomic DNA Mini Kit was used for genomic DNA extraction and shotgun sequencing was performed using the Illumina MiSeq platform. Draft sequences were assembled using Glimmer 3.02 and identified by NCBI BLAST. Isolates were serotyped using DengLab SeqSero1.0 and KAAS was used to map KEGG pathways. CARD and CGE ResFinder were used to identify antibiotic resistance genes. Prophage sequences were identified through CGE MetaPhinder-2.1 and *Salmonella* Pathogenic islands were identified through CGE SPIFinder 1.0 and human pathogenicity estimated using PathogenFinder. Phylogenetic analysis was made using PATRIC and antibiotic susceptibility was evaluated using the disk diffusion method.

2.3 Induced kanamycin resistance

Isolates were exposed to kanamycin in test strips on agar and, when left to over-grow, small colonies growing closest to the kanamycin concentration zone were picked and passaged on kanamycin-containing agar for a further three generations. The MIC[2] for kanamycin of each generation was determined on LB agar using MIC Test Strips (Liofilchem®) with a concentration gradient of 0.016 - 256 µg/mL.

3. Results

3.1 Identification of strains

Genomic sequence data for all eight isolates is shown in Figure 1.

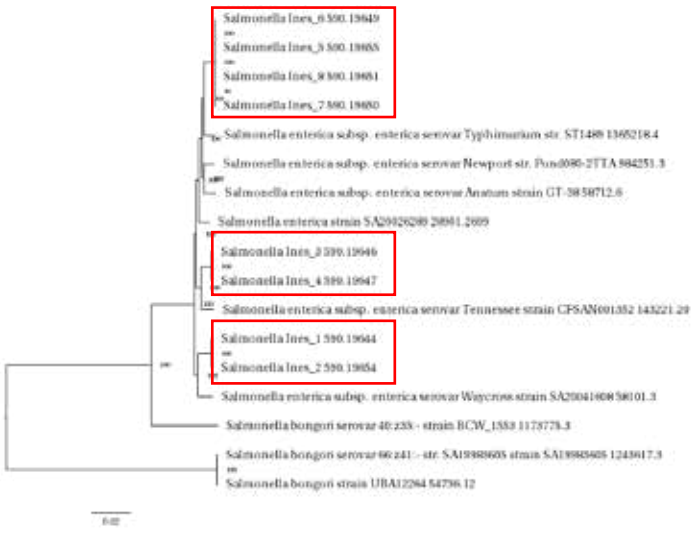


Figure 2. A phylogenetic tree, based on a comparison of 500 genes, generated using PATRIC.

3.2 Antibiotic Resistance

Functional tests for antibiotic resistance were carried out using disk diffusion (an example is shown in Figure 3). All eight isolates showed resistance to ampicillin (a beta-lactam antibiotic), but susceptibility to tetracycline and to kanamycin (an aminoglycoside).

Genomic analysis using CGE ResFinder and CARD, found the aminoglycoside resistance genes *aac(6')-laa* and *KpnF* present in all isolates, and *kdpE, baeR, KpnE* in all except isolate 1.

Figure 3. A disk diffusion assay indicates resistance to ampicillin but susceptibility to tetracycline and kanamycin.

3.3 Pathogenic Islands

Salmonella Pathogenic Islands (SPIs) in each isolate are shown in Figure 4. Isolates 1-4, all coming from the same household pet gecko, each hold fewer Pathogenicity Islands. Isolates 5-8, which are sourced from four individual petting zoo geckos, show a higher likelihood for pathogenicity.

Analysis of the draft sequences by PathogenFinder suggests that all eight isolates have a greater than 90% probability of being a human pathogen (Figure 5). It is also clear that isolates 1-4 do not match as many pathogenic families as isolates 5-8, which come from the petting zoo.

In addition to the *aac(6')-laa* and other resistance genes found in the isolates, some pathogenic islands might also favour aminoglycoside resistance. Mg^{2+} uptake genes (*MgtC, MgtB, MarT, MisL*) were present in all isolates. Mg^{2+} can inhibit the transport of aminoglycosides (Kang *et al*., 2000). Complementing this, electron transfer flavoprotein beta subunit (*gpiA/etfA*), normally part of SPI-14, was absent from all isolates except isolate 8. gpiA (*etfA*) is an important protein for the uptake of aminoglycosides as they require electron transport in order to pass through the inner membrane of the cell (Taber *et al.*, 1987; Voskuil *et al*., 2018).

3.4 Genomic analysis

The *aac(6')-laa* aminoglycoside resistance gene was found present in all isolates, yet after closer inspection of the sequence, it was found that all were pseudogenes (compromised genes) and likely to be inactive. Each isolate had some modifications to the certain gene segments as shown in Figure 6.

Figure 4. Pathogenicity Islands found in each isolate and their functions.

Isolate	1	2	3	4	5	6	7	8
Pathogenic Families Matched	569	656	601	653	763	696	747	747
Probability of being human pathogen	0.942	0.943	0.948	0.945	0.939	0.939	0.938	0.938

Figure 5. Matches to Pathogenic Families and probability of human pathogenicity in isolates.

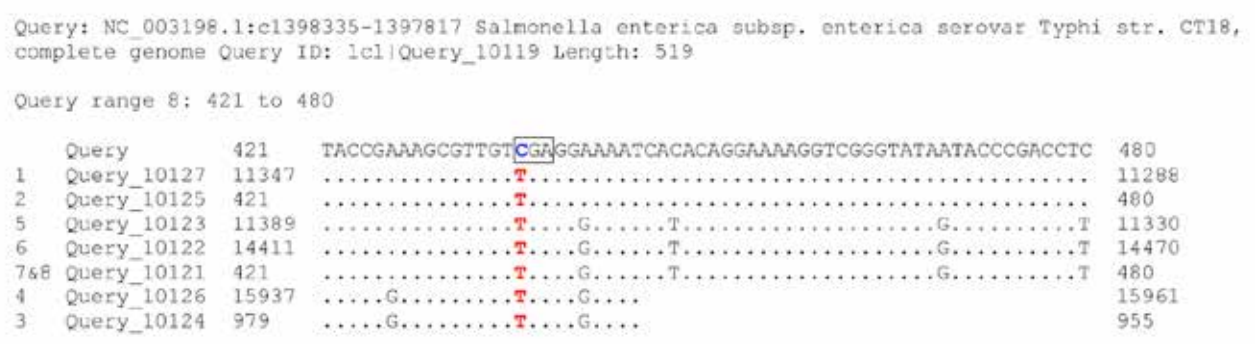


Figure 6. Alignment comparison of a section of the *aac(6')-laa* gene: all eight isolates present a modification from CGA (arginine) to the stop codon TGA

All eight isolates have a single nucleotide substitution near the C-terminal of the *aac(6')-laa* gene that changes the CGA codon (normally encoding arginine-146) to TGA, creating a premature STOP codon that results in the loss of the last 27 amino acids of the protein sequence. Isolates 1 and 2 have no other differences from the *aac(6')-laa* reference sequence other than this single nucleotide polymorphism.

Isolates 5-8, which all come from the petting zoo geckos, have a further deletion of 8 base pairs in the section beyond this. For isolates 3 and 4, most of the sequence following this STOP codon is deleted.

3.5 Induction of aminoglycoside resistance

Each isolate was exposed to kanamycin over four generations, with five isolates showing greatly increased resistance after four generations of successive growth on kanamycin-containing LB agar: 1(e), 2(h), 5(a), 7(b) and 8(b), each derivatives of isolates 1, 2, 5, 7 and 8 respectively. Isolates 2(h), 7(b) and 8(b), in particular, showed a 24-fold increase in MIC from 2 µg/mL to 48 µg/mL as shown in Figure 6.

Nevertheless, shotgun sequences of the generation 4 isolates showed that the early stop codon in the *aac(6')-laa* gene was still present in these five isolates.

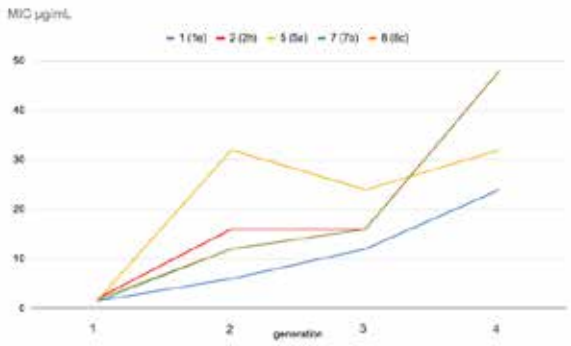


Figure 7. MIC increases within four generations in five isolates exposed to kanamycin.

4. Discussion

It is significant that all of the crested geckos sampled contained non-typhoidal strains of *Salmonella enterica*. Furthermore, genomic analysis shows that the petting zoo gecko *Salmonella* isolates contain around 20% more virulence genes than the isolates from the household pet gecko. A recent review by Halsby *et al*. (2014) found 57 incidents of zoonotic infections associated with pet shops (excluding rabies and zoonoses from pet foods) and warned that pet shops could act "as a nexus point for zoonotic disease".

The results above show the presence of *aac(6')-laa* as a pseudogene in all eight isolates, which may partly account for the lack of phenotypic resistance to kanamycin. Nevertheless, after exposure to kanamycin, selected isolates were able to show a large increase in tolerance of the aminoglycoside. This could suggest that the pseudogene might have been reactivated after exposure to the antibiotic.

One place for this to occur may have been in the early stop codon in *aac(6')-laa*. If this single nucleotide switch is a mechanism for gene silencing, then perhaps it could be changed back when the bacteria are exposed to the antibiotic. However, shotgun sequences showed that this codon did not change in the genomes of the generation 4 isolates, so that there must be another explanation for the increased resistance.

Another mechanism may have been the action of the ribosomes themselves. Hainrichson *et al*. (2007) explain that around 4% of the time, a stop codon will be read incorrectly by the ribosome, which means that it could then result in a functional enzyme product. Furthermore, because aminoglycosides interfere with the ribosome, they can make it more likely to read through a stop codon. In the case of *aac(6')-laa*, this would then increase the chances of the ribosome producing an active protein.

Since the eight isolates came from five different geckos from different environments, it seems unlikely that the same random mutation can account for the occurrence of the same stop codon in the *aac(6')-laa* gene, so perhaps there is a function for this. Since the pseudogene profile is thought to be matched to the host environment (Lopera *et al*., 2017), this may suggest that something in the crested gecko itself may have a relationship with the pseudogene found in all the isolates.

The induced resistance may also have occurred through transcriptional control, where the genes in the isolates remain the same, but there is more expression of the ones that are there. For example, the isolates may produce more membrane efflux pumps or allow the existing ones to work faster or more efficiently when exposed to the aminoglycoside antibiotic. Nicoloff *et al.* (2019) looked at the fraction of cells within a bacterial sample that seem to be more resistant to an antibiotic than others and concluded that gene amplifications (*i.e.* more copies of a gene) often explained the difference.

5. Conclusion

Pet geckos are a consistent source of *Salmonella* and reptiles that have spent time in pet shops may present a health hazard if people do not look after their own personal hygiene (such as washing hands and avoiding contact with faeces). In addition, *Salmonella enterica* may not be as susceptible to antibiotic treatment as it first appears on an agar plate because it seems easy to induce resistance to an antibiotic after repeated exposure.

6. Next Steps

Since this study has worked with draft sequences, it is not easy to see if genes have been rearranged in a way that can explain the induced resistance after exposure. Full genome sequences would be able to show if the positions of the genes change to regulate expression after exposure. RNA sequencing (RNAseq) would also be a good way to look at gene expression to see how *Salmonella* responds to an antibiotic.

Following up on the pet shop evidence, a closer survey of microbes in pet shops and animals housed there – as well as some of the management practices – might be able to reveal the factors that are driving the increased virulence and antibiotic resistance in these environments.

Acknowledgements

Thank you to Jurassic Garage Hong Kong for providing the gecko faecal samples. Special thanks to Prof. Fred Leung and the *Shuyuan* Molecular Biology Laboratory, particularly to Ms. Grace Lai and Dr. Simon Griffin.

Supplementary Data

Supplementary data for this project is available at https://bit.ly/3htmwy8 or by scanning the QR code:

References

CDC (2017*) Zoonotic Diseases*. Retrieved from www.cdc.gov/onehealth/basics/zoonotic-diseases.html

CFSPH (2013) *Salmonellosis Associated with Reptiles*, Iowa State University. Retrieved from www.cfsph.iastate.edu/FastFacts/pdfs/salmonella_reptile_F.pdf

da Silva, M.B., Portela, J.M., Li, W., Jackson, M., Gonzalez-Juarrero, M., Hidalgo, A.S., Belisle, J.T., Bouth, R.C., Gobbo, A.R., Barreto, J.G., Minervino, A.H.H., Cole, S.T., Avanzi, C., Busso, P., Frade, M.A.C., Geluk, A., Salgado, C.G., Spencer, J.S. (2018) Evidence of zoonotic leprosy in Pará, Brazilian Amazon, and risks associated with human contact or consumption of armadillos*. PLoS Neglected Tropical Diseases*, vol. 12(6): e0006532. doi.: 10.1371/journal.pntd.0006532

Giannella, R.A. (1996) Salmonella. In: Baron, S. (Ed.) *Medical Microbiology*. 4th edition. Galveston (TX): University of Texas Medical Branch at Galveston; Ch. 21. Available from: www.ncbi.nlm.nih.gov/books/NBK8435/

Goodlet, K.J., Benhalima, F.Z., Nailor, M.D. (2018) A Systematic Review of Single-Dose Aminoglycoside Therapy for Urinary Tract Infection: Is It Time To Resurrect an Old Strategy?, *Antimicrobial Agents and Chemotherapy,* vol. 63(1): e02165-18. doi: 10.1128/AAC.02165-18

Hainrichson, M., Nudelman, I., Baasov, T. (2008) Designer aminoglycosides: the race to develop improved antibiotics and compounds for the treatment of human genetic diseases, *Organic and Biomolecular Chemistry*, vol. 6, pp. 227-239. doi: 10.1039/b712690p

Halsby, K.D., Walsh, A.L., Campbell, C., Hewitt, K., Morgan, D. (2014) Healthy Animals, Healthy People: Zoonosis Risk from Animal Contact in Pet Shops, a Systematic Review of the Literature. *PLoS ONE*, 9(2): e89309. doi:10.1371/journal.pone.0089309

Kang, H.S., Kerstan, D., Dai, L., Ritchie, G., Quamme, G.A. (2000) Aminoglycosides inhibit hormone- stimulated Mg2+ uptake in mouse distal convoluted tubule cells, *Canadian Journal of Physiology and Pharmacology*, vol. 78(8), pp. 595-602.

Lopera, J., Miller, I.J., McPhail, K.L., Kwan, J.C. (2017) Increased Biosynthetic Gene Dosage in a Genome-Reduced Defensive Bacterial Symbiont, *mSystems,* 2(6): e00096-17. doi: 10.1128/mSystems.00096-17

Nicoloff, H., Hjort, K., Levin, B.R., Andersson, D.I. (2019) The high prevalence of antibiotic heteroresistance in pathogenic bacteria is mainly caused by gene amplification*, Nature Microbiology*, vol. 4, pp. 504–514. doi: 10.1038/s41564-018-0342-0

Shi, C., Singh, P., Ranieri, M.L., Wiedmann, M., Moreno Switt, A.I. (2015) Molecular methods for serovar determination of *Salmonella, Critical Reviews in Microbiology*, vol. 41(3), pp. 309-325, doi: 10.3109/1040841X.2013.837862

Taber, H. W., Mueller, J. P., Miller, P. F., & Arrow, A. S. (1987) Bacterial uptake of aminoglycoside antibiotics, *Microbiological Reviews*, vol. 51(4), pp. 439-457.

Voskuil, M.I., Covey, C.R., Walter, N.D. (2018) Antibiotic Lethality and Membrane Bioenergetics, *Advances in Microbial Physiology*, vol. 73, pp. 77-122. doi:10.1016/bs.ampbs.2018.06.002

Wang, L.F., Shi, Z., Zhang, S., Field, H., Daszak, P., Eaton, B.T. (2006) Review of bats and SARS, *Emerging Infectious Diseases*, vol. 12(12), pp. 1834-1840. doi:10.3201/eid1212.060401

How does the Voltage across an Electrostatic Precipitator affect its Performance?

Marco Lau

Abstract

Particulate matter (PM) is the mixture of particles and liquid droplets in the air (Particulate Matter (PM) Basics, 2018). PM10 and PM2.5, particulate matters that are generally 10 micrometers and 2.5 micrometers or smaller respectively, can be harmful to the human body. For every 10 $\mu g/m^3$ increase in PM10 concentration, there will be a rise of 0.58% of respiratory disease; and for every 10 $\mu g/m^3$ increase in PM2.5 concentration, there will be a rise of 8% in 'hospitalization rate'(Xing, *et al*, 2016). PM10 and PM2.5 are commonly found in construction sites, fires, and roads (Particulate Matter (PM) Basics, 2018).

As the world population increases, landfills will eventually run out. A solution to waste disposal is central incineration (Bump, 2012). The problem with incineration, however, is the air pollution. With a 'growing public awareness' of air quality, there is a need for more efficient air pollution control (Bump, Robert L, 2012). According to *The Use of Electrostatic Precipitators on Municipal Incinerators*, the electrostatic precipitator has the highest efficiency (percentage of PM filtered) to cost per ton ratio (the cost to filter a ton of PM), compared to other air pollution controls such as: baffled spray chamber, wet scrubber, and spray chamber fabric collector (Bump, Robert L, 2012). This makes the electrostatic precipitator the best air pollution control in terms of efficiency and cost.

The PM2.5 concentration standard of World Health Organization (WTO) is 10 μgm^3. As the PM2.5 concentration outdoors, on average, is 100 μgm^3, electrostatic filters must operate at above 90% efficiency, in order to maintain the WTO standard of PM2.5 concentration indoors (Tang, *et al*, 2018). However, six electrostatic precipitators in commercial offices in cities such as Beijing and Shanghai were found to be only operating at 20% - 55% efficiency (Tang, *et al*, 2018).

Many factors affect the performance of an EP. Face velocity - 'the measure of airspeed entering or exiting the filter' - and the charge of the particle, for instance, are two examples (Walsh, *et al*, 1998). According to Tang, *et al*, the increase in face velocity will result in a decrease in filter efficiency. In contrast, the filter will be more efficient when particles have a higher charge (Tang, *et al.* 2018).

In this study, the voltage across the electrostatic precipitator will be varied in order to optimize the performance of a self-made electrostatic precipitator.

1. Working Mechanism of an EP

An EP uses the electric field generated from the potential difference across it to filter particulate matter from the air. In an electrostatic precipitator, there are two main parts. The first is an ionization net, which is a metal net that ionizes the particulate matter in the air (Helbing, 2019). The second part is the collector, which is a metal surface. A high voltage will exist between the collector (high potential) and the ionization net (low potential). From Figure 1, it can be seen that the air is initially filled with clean air particles and neutral particulate matter. First, the ionization net ionizes PM in air with the electric field generated from the low electric potential. This will not affect clean air particles because oxygen, and nitrogen particles have a high ionisation energy due to its nonpolarity (Helbing, 2019). Above the ionization net, it can be seen that the air is composed of clean air particles and ionized PM. The collector will then

This article was written in partial fulfillment of the Extended Essay for the IB Diploma, May 2019.

attract ionized PM with the electric field generated from the high electric potential. Oxygen and nitrogen remains neutral, therefore, will not be attracted by the collector. In theory, the output of the EP will be pure air (mostly O_2 and N_2) and without PM. Essentially, the EP uses a high voltage across the ionization net and the collector to generate electric fields to ionize and capture PM from the air and hence filter air.

Past research has shown that variables such as face

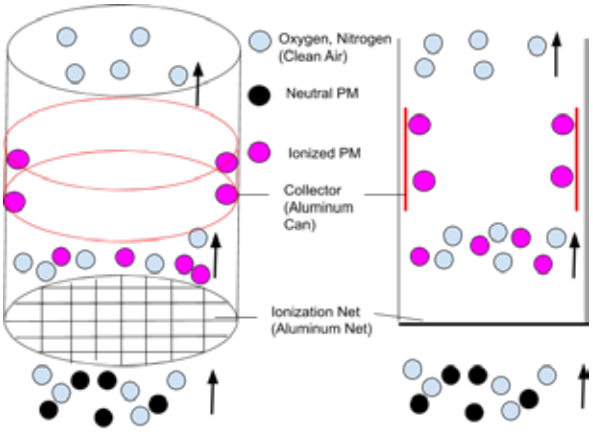


Figure 1. The Schematic Diagram of The Electrostatic Precipitator.

velocity, particle size, and humidity can affect the performance of EPs (Walsh, D.C., *et al*, 1998)(Tang, Min, *et al*, 2018). To further explore the variables that can vary the performance of the filter, the voltage across the ionization net and the collector of the EP will be experimented on. The voltage can be changed on the high voltage power supply.

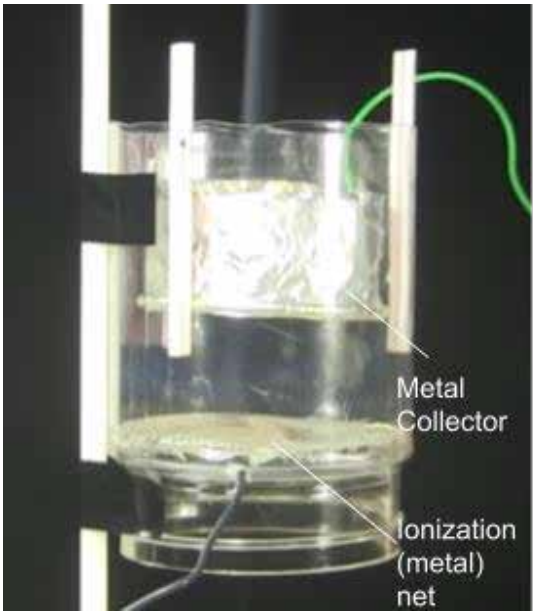


Figure 2. The Referenced Design of the EP.

2. Methodology

2.1 Self-made EP

A self-made EP will be used for the experiment. The benefit of a self-made EP is that the components can be changed easily and the voltage across the device is controlled. The design of the EP is based on an existing design found on the internet, as shown in Figure 2 (RimstarOrg, 2014).

Upon the existing design, several modifications were made to optimize the performance and lower the difficulty in production. First, the aluminum ring connected to the green wire is replaced with a modified soft drink can, which made the production of the EP easier. A bottle with a narrow opening is used instead of the cylindrical bottle. The metal net is balanced on the narrow section of the bottle, where the input of air enters from. This is beneficial for two reasons. First, as the smoke (source of PM) enter through a narrow opening, they will be more concentrated. Second, the metal net can be easily changed, as it is not attached to the mainframe of the device.

2.2 Measuring the Performance

The data collection method initially planned, was to measure the change of mass in the EP after it operates for a period of time. The source of smoke chosen was Chinese incense because they are cheap and common. The mass of the filter will be measured on an electronic scale, which is accurate up to four decimal points, before and after the EP operation.

Figure 3a. The Brand of Chinese Incense

After some preliminary tests, it can be seen that the electronic balance is not precise enough to measure the mass of filtered particles. This imprecision is large enough to ruin the data, as the mass change of the filter is very small.

2.3 Light Spectrum

Figure 3b. The Chinese Incense

Another method to measure the EP's performance is through analyzing the absorption spectrum. A light beam in the visible region will be shone through the output funnel of the filter where a spectrometer will collect the result - of the light - on the other side. The intensity or amplitude of the light can be measured to analyze the density of smoke, and hence, quantify the performance of the filter.

However, the lack of a precise spectrometer and experience in light experimentation limits the use of this method.

2.4 Image Analysis

In the preliminary tests, it can clearly be seen that the colour of smoke changes before and after the EP. If the human eye can observe a difference, so can cameras. This inspired the idea of image analysis.

ImageJ is an image analysis software which can process images for scientific purposes (ImageJ, n.d.). ImageJ is often used to analyze images of bacterial growth and count bacteria population through color recognition. ImageJ can also analyze black and white images with a function called histogram.

Histogram analyzes an image and represents the image with a graph, where the x-axis is the possible color values and the y-axis is the pixels for that specific color in the image. The x-axis, possible color values, depends on the type of image. For example, a grayscale 8-bit image has 2^8=256 possible color values. Therefore, the x-axis will range from 0 to 255, where 0 is black and 255 is white.

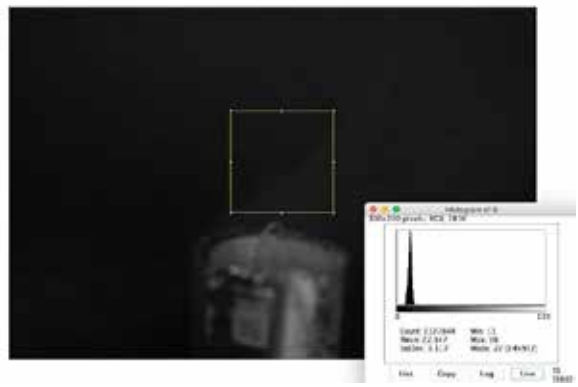


Figure 4. Image Analysis Method in ImageJ

In this grayscale 8 bit image, for example, the pixels of each color in the yellow box is counted and plotted in the graph. 0 represents pure black and 255 represents pure white. A wave-shape is plotted on the graph, where the mean (average) grayscale value is 22.347 on the grayscale. In the preliminary tests, 8-bit images were sufficient to show the performance of the EP. 16-bit and 32-bit images were also tested. However, they require more computational time to analyze and does not show a significant improvement in precision. Therefore, the 8-bit image was chosen. A grayscale image eliminates the influence of other colors and allows Histogram to focus on the smoke concentration (white) and background (black) of the image. This will simplify the analysis of the images. Hence, the 8-bit grayscale images were chosen for the experiment.

To enhance the color of the smoke, the experiment will be set up in front of a black background. A Canon Powershot G7 X camera will be used to take the pictures. The ISO of the camera is set to 1000, exposure and aperture is set to manual, and the images are taken in color. The images will be converted to grayscale with ImageJ. The mean grayscale value will be taken as data. As the smoke determines the mean grayscale value (otherwise the pure black background will have a 0 mean grayscale value), the mean

grayscale value can be interpreted as the relative density of smoke. The denser the smoke is, the higher the mean grayscale value.Therefore, in this study, the term 'density of smoke' will replace 'mean grayscale value', in order to give a clearer physical picture of the data.

2.5 Concentrated smoke beam

To measure a more obvious result, multiple incense will be used in a squeeze bottle to produce a concentrated smoke beam. When more concentrated smoke enters the filter, the difference between input and output air will be maximized. An Aqua One 2500 Precision Air Pump is attached to the bottle to create a steady air flow, and hence, a steady flow of smoke.

3. Experimentation

Figure 5a. Set Up of the Experiment

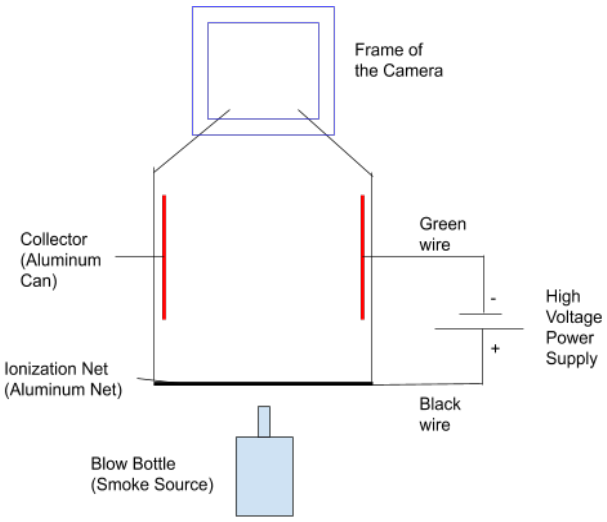


Figure 5b. The Schematic Diagram of The Set-Up of the Electrostatic Precipitator

The experiment will be set up as shown in Figure 5a and 5b. The plastic bottle with the aluminum soda can inside is the self-made EP. The green wire connects the collector to the positive DC output, while the black wire connects the ionization net to the negative DC output. The power source is a high voltage power supply. The blow bottle with incense inside is located below the EP.

In this experiment, the effect of voltage on the performance of the EP will be investigated. The voltage will be provided by a high voltage power supply, ranging from 0 to 5000 volts with 500 volts increments. The performance of the EP will be measured through ImageJ's histogram function - the mean grayscale value of the image in a consistent selected area (directly above the EP with a size of 1000x1000 pixels).

As smoke can be easily affected by turbulent air or the combustion process of the incense, a total of 8 trials, with 11 data points per trial, will be conducted. For each trial, the performance of the EP at 0 to 5000 volts will be measured before beginning the next trial. A total of 8 trials will diminish the uncertainty of smoke and provide accurate data.

3.1 Examples of how the images are analyzed

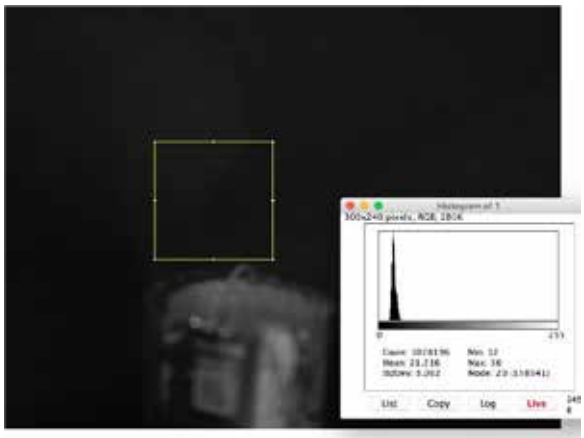


Figure 6. Data Collection of Trial 1 at 1000V

The selected area (yellow box, 1000x1000 pixels) is directly above the EP and has a mean grayscale value of 21.216.

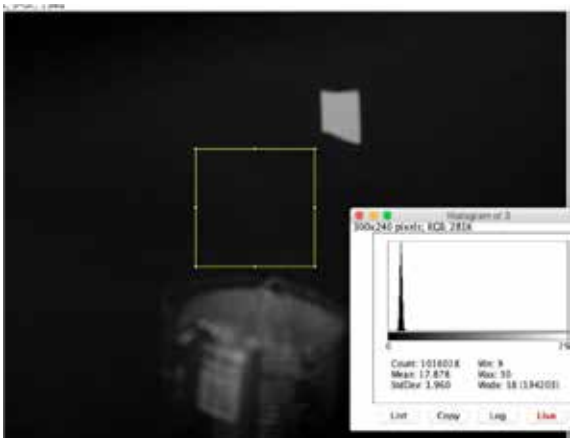


Figure 7. Data Collection of Trial 5 at 3000V

The selected area (yellow box, 1000x1000 pixels) is directly above the EP and has a mean grayscale value of 17.878

3.2 Experimentation Results

How Voltage Affects the Performance of the EP (Mean Color Value)								
	Density of Smoke (mean grayscale value +/- 0.0005)							
Voltage (volts) +/- 5 volts	Trial 1	Trial 2	Trial 3	Trial 4	Trial 5	Trial 6	Trial 7	Trial 8
0	22.636	21.134	22.417	25.484	20.451	20.239	17.311	17.259
500	21.082	20.146	21.982	25.095	18.521	20.001	16.716	15.637
1000	20.673	19.753	21.637	24.619	18.309	20.350	15.903	16.395
1500	19.987	18.645	21.347	24.135	18.342	19.626	15.612	16.083
2000	19.529	17.964	20.973	23.628	18.166	19.668	15.629	16.440
2500	19.231	17.779	20.120	23.043	18.008	19.228	14.124	16.277
3000	18.808	17.543	19.870	23.138	17.878	19.117	16.608	16.113
3500	18.103	17.299	18.972	22.886	17.329	19.248	15.774	16.566
4000	17.733	17.270	18.351	22.779	16.982	18.971	16.655	16.545
4500	17.438	17.031	17.638	22.036	16.463	18.679	16.159	16.890
5000	16.954	16.991	17.234	21.942	16.818	18.383	15.509	15.713

3.3 Data Processing

3.3.1 Uncertainty of measurement

The histogram function presents the number of pixels in each grayscale value. The data resembles the shape of a wave. The peak, where there is the most pixels of a certain grayscale value, is approximately the mean grayscale value. Grayscales above and below the peak will have gradually less pixels. As this study extrapolates the mean value of the wave-like data, the width of the wave, which is the distribution of data, can be interpreted as the uncertainty. After research, the method 'Full Width Half Maximum' seemed suitable to analyze the width of the wave.

Full width half maximum (FWHM) is a method to process data involved in waves or peaks. It is the width of a wave/curve at its half maximum point. It can be used to measure pulses for optical communications (A Glossary of Spectroscopic Terms, n.d.). FWHM also shows how wide the peak is, in this case, the distribution of grayscale values within the selected area. By dividing the width of the wave by the total number of grayscale values in a grayscale 8-bit image - 256 - the uncertainty can be found.

The data is exported from histogram and processed on Google Spreadsheet. The following math equation is used to determine the FWHM value:

$$FWHM = x_2 - x_1$$

Where : x_1 *and* x_2 *are points that have a y value of* $0.5 \times y_{max}$, y_{max} *is the y value of the peak*

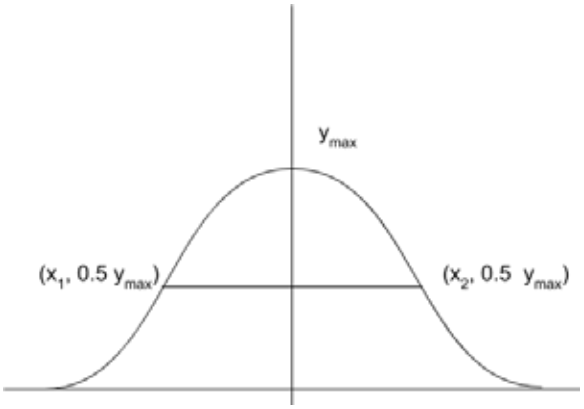


Figure 8. Full Width Half Maximum Diagram

The value calculated from the Google Spreadsheet is then compared with another equation provided by an online source, (Aspelin, 2012). The FWHM of each data point in each trial is then divided by 256 - the total number of color values - and averaged to find the overall percentage error, which is 1.79%.

FWHM (wave width)	Voltage (V)	Trial 1	Trial 2	Trial 3	Trial 4	Trial 5	Trial 6	Trial 7	Trial 8
	0	6	4	7	5	5	5	5	4
	500	5	3	6	7	4	3	4	4
	1000	4	4	6	9	5	4	3	3
	1500	4	5	5	5	4	3	3	3
	2000	3	3	7	6	5	4	3	3
	2500	3	5	6	5	4	3	5	4
	3000	3	5	6	6	6	4	3	3
	3500	4	5	7	5	5	4	4	3
	4000	4	4	6	5	4	4	4	3
	4500	5	6	7	6	6	4	5	4
	5000	7	5	7	4	5	3	3	3
FWHM (%)	Voltage (V)	Trial 1	Trial 2	Trial 3	Trial 4	Trial 5	Trial 6	Trial 7	Trial 8
	0	2.34%	1.56%	2.73%	1.95%	1.95%	1.95%	1.95%	1.56%
	500	1.95%	1.17%	2.34%	2.73%	1.56%	1.17%	1.56%	1.56%
	1000	1.56%	1.56%	2.34%	3.52%	1.95%	1.56%	1.17%	1.17%
	1500	1.56%	1.95%	1.95%	1.95%	1.56%	1.17%	1.17%	1.17%
	2000	1.17%	1.17%	2.73%	2.34%	1.95%	1.56%	1.17%	1.17%
	2500	1.17%	1.95%	2.34%	1.95%	1.56%	1.17%	1.95%	1.56%
	3000	1.17%	1.95%	2.34%	2.34%	2.34%	1.56%	1.17%	1.17%
	3500	1.56%	1.95%	2.73%	1.95%	1.95%	1.56%	1.56%	1.17%
	4000	1.56%	1.56%	2.34%	1.95%	1.56%	1.56%	1.56%	1.17%
	4500	1.95%	2.34%	2.73%	2.34%	2.34%	1.56%	1.95%	1.56%
	5000	2.73%	1.95%	2.73%	1.56%	1.95%	1.17%	1.17%	1.17%
									1.79%

Figure 9. The Calculation of Error Bar with the Full Width Half Maximum Method on Google Spreadsheet

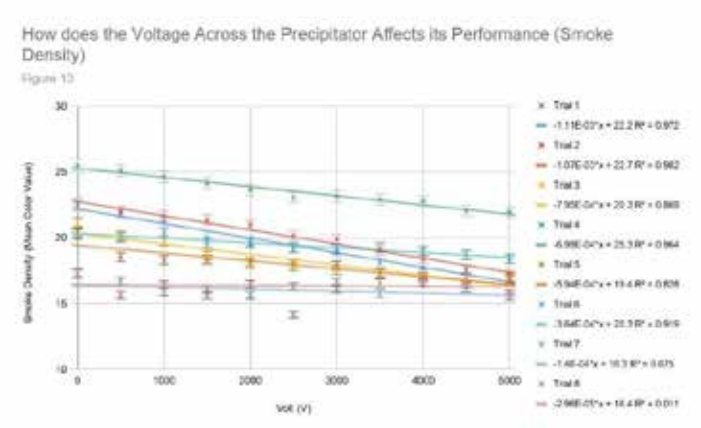


Figure 10.

3.4 Data Analysis

Figure 10 shows the relationship between the voltage across the precipitator and the smoke density after the smoke leaves the EP. As shown, there is a general decrease in smoke density as the voltage increases, which suggests a negative correlation between the two variables. For example, in trial 4, the smoke density is 25.484 at 0 volts. At 2000 volts, the smoke density is 23.628. At 5000 volts, the smoke density is 21.942. The trendline of trial 4 is -0.699x+25.3, which suggests an inversely proportional relationship. A similar negative trend can be observed across all 8 trials, although less obvious in trials 7 and 8. Therefore, it can be concluded that as the voltage increases, the smoke density decreases. A lower smoke density is represents a higher performance of the EP. Therefore, it can be concluded that as the voltage increases, the performance of the EP also increases.

The R^2 value of the trendline in Figure 10 is the coefficient of determination, which represents how well the trendline reflects the actual data. In trial 1- 6, the R^2 values are around 0.9, which suggests an accurate representation of the actual data. In trial 7 and 8, however, the R^2 values are 0.075 and 0.011, which suggests a less accurate representation of the actual data. Focusing on the data from trial 7 and 8:

Considering the large uncertainty in measuring smoke density, the fluctuation in data suggests no significant correlation between the voltage of the precipitator and smoke density, in contrast with the first 6 trials.

Initially, 8 trials were conducted in order to produce accurate results. However, even after redoing the 8 trials, a similar loss of trend can still be observed in the last two trials. This suggests that the lack of correlation between the voltage of the EP and its performance is not an experimental error, but instead, a scientific phenomenon. Puzzled by the data, I read more scientific papers and found a possible explanation: saturation point.

As EPs attract negatively charged PM ions in the air, the surface of the metal collector will be covered by these particles. These negatively charged particles on the surface will lower the electric potential of the collector, which will reduce the attractive force of the precipitator . After a certain point, the attraction force will diminish until it is unable to attract PM ions, which will be the saturation point.

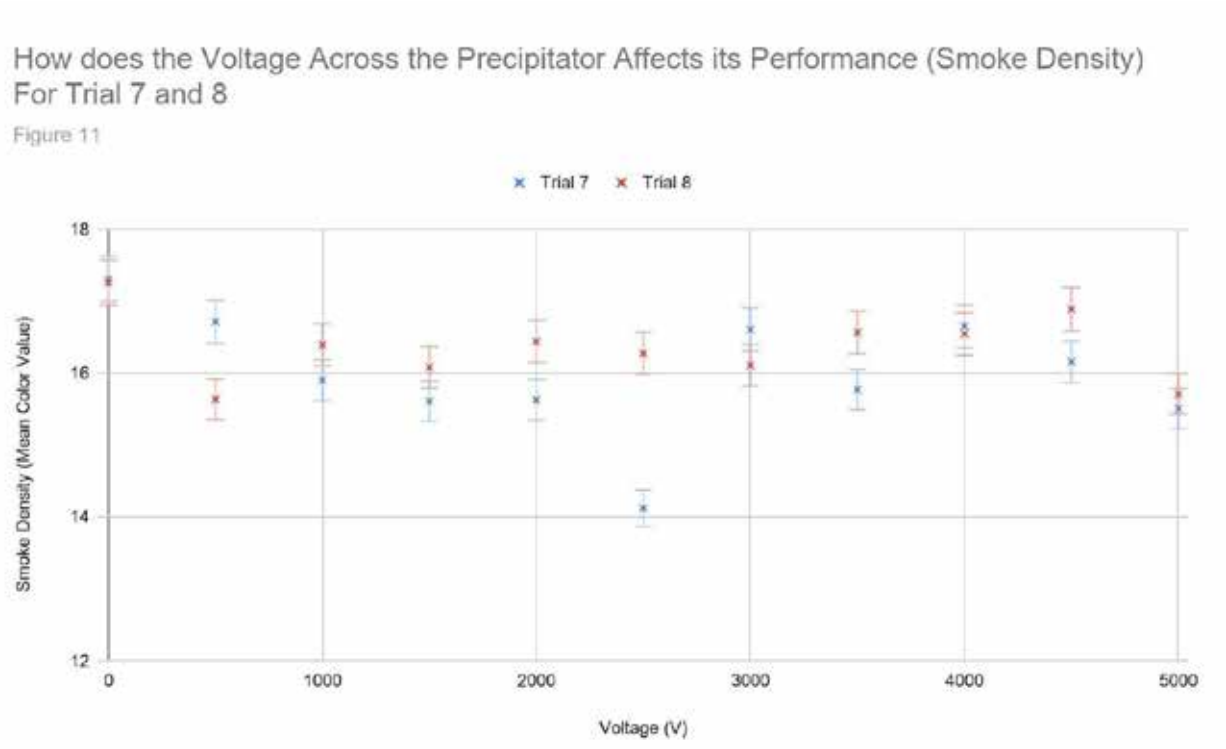


Figure 11. It can be seen that the smoke density in trial 7 and 8 are fluctuating.

To investigate deeper into the saturation point encountered in trial 7 and 8, another experiment is set up. The surface area of the collector (aluminum soda can) will be changed - by changing the height of the cylinder (8.5 cm, 7 cm, 5.5 cm) and keeping the diameter constant. The performance of the EP with different sized collectors will be measured every 20 seconds for 3 minutes at a constant 5000V voltage. The saturation point is when the output smoke has a mean color value identical or similar to when the EP has no voltage across it.

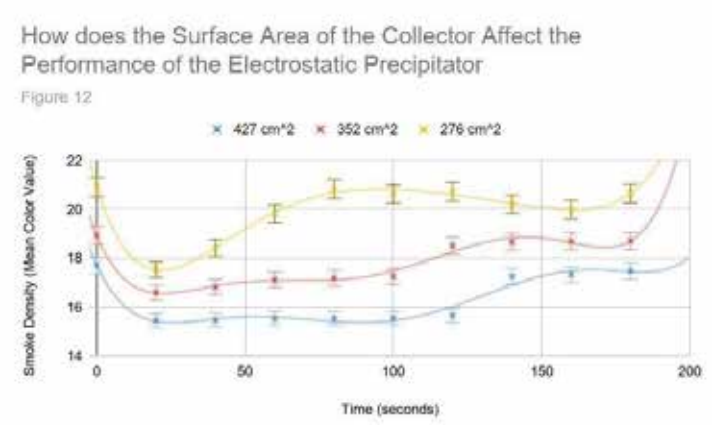


Figure 12.

*Surface area = $\pi r^2 h$, where the radius of the aluminum can is 4 cm

In Figure 12, the smoke density of a selected area is plotted for 3 minutes, with 20 seconds interval. The EP is operating at 5000 volts. The height of the metal collector is varied (8.5 cm, 7 cm, 6.5 cm) while the diameter is kept constant, which causes the surface area of the metal collector to be varied (427 cm^2, 352 cm^2, 276 cm^2). The data points are connected with a smooth line to more clearly illustrate the relationship between saturation point and surface area of the metal collector. As can be seen, in the trial with a surface area of 427 cm^2, the saturation point occurs at 140 seconds. This can be seen from the plateau of the smooth line. At 0 seconds, when the EP is not on, the smoke density is at 17.706. After the EP is turned on, the smoke density stayed relatively low, at around 15.5, which shows that the EP is operating normally. At 140 seconds, the smoke density increased significantly, back to 17.255, approaching the value before the EP is on. Therefore, the saturation point of an EP with a collector of 427 cm^2 will be 140 seconds running at 5000 volts. In the trial with a surface area of 352 cm^2, the same increase in smoke density can be seen at 120 seconds. This shows that the saturation point for 352 cm^2 occurs at 5000 volts at 120 seconds. For the trial with 276 cm^2 of surface area, the smoke density dropped from 20.921 to 17.532 once the EP is turned on. The number slowly increased to 18.419 and then to 19.823 at 60 seconds. The peak, or the saturation point, can be seen at 80 seconds, where the smoke density reached 20.837, extremely close to the initial value. The smoke density after 80 seconds remains somewhat similar.

The data suggests that the surface area of the collector of the EP is positively correlated with the operation time. The larger the surface area, the longer the EP can operate before reaching its saturation point. This can be explained with electric potential and electric fields.

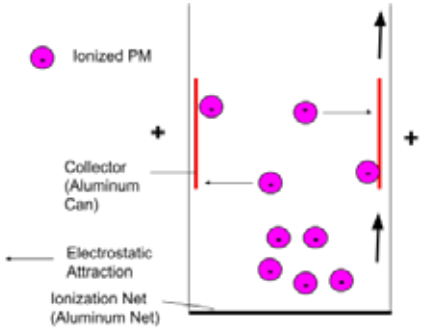


Figure 13a. EP Functioning.

Figure 13a shows the EP when there is a high voltage across the precipitator. The high potential of the collector attracts negatively charged PM. The electrostatic attraction between the collector and the ionized PM captures and filters the particles. The attractive force between the collector and the PM is described by this formula:

$$F = Eq$$

Where F is the attractive force, E is the electric field generated from the high electric potential of the collector, and q is the charge of the PM.

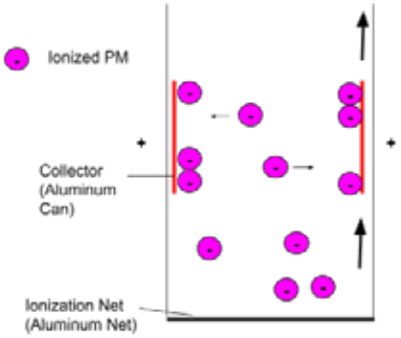


Figure 13b. EP After a Period of Time

Figure 13b shows the EP after a period of time. The surface of the collector is filled with negatively charged PM ions it attracted. The negatively charged PM will diminish the high electric potential of the collector and cause the electric field of it to weaken. This can be shown by an equation, where E_I and F_I is the initial electric field and attraction force of the collector, E_F and F_F is the electric field and attraction force of the collector after some time, respectively.

$$F_I = E_I \times q$$

$$F_F = E_F \times q$$

As the PM ions attracted onto the surface of the collector will diminish the electric potential and weaken the electric field,

$$E_I > E_F$$

Therefore,

$$F_I > F_F$$

Which shows that the attractive force of the electric field generated by the collector will decrease as more PM ions are attracted.

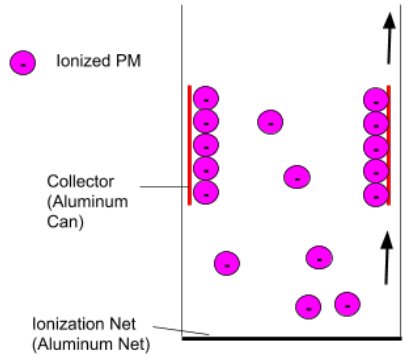


Figure 13c. EP Saturated

In Figure 13c, PM ions on the collector reduced the high electric potential to 0.

The electric field will also equal 0. This is when the EP reaches its saturation point, where the collector has no electric potential and will no longer be able to attract PM ions and filter air.

This can be observed in the data of Figure 12.

4. Conclusion

Overall, it can be concluded that the voltage across the precipitator and its performance has a positively correlated relationship.

Moreover, from the anomaly of trial 7 and 8 (Figure 10 and 11), another experiment has been conducted, which investigates the relationship between the surface area of the collector and its saturation point. The results suggest that the surface area of the EP and the saturation point is positively correlated. An EP with a larger surface area will take longer to reach its saturation point, meaning the total units of PM filtered will be greater compared to filters will smaller surface area.

The investigation of saturation point also reveals the limitation of the EP. The self-made EP with a collector of 427 cm^2 surface area could only operate effectively for approximately 140 seconds at 5000 Volts. A commercial air pollution control device must be able to function for a much longer time than 140 seconds. This suggests that despite the benefits of the EP, further research must be conducted to increase its saturation point in order to commercialize the EP as a viable air pollution control method.

5. Evaluation

5.1 Strengths

In contrast with literature methods in measuring the performance of EPs (light absorbance/transmittance), the experiment uses image analysing software instead. The measurements were made with precision and accuracy. This is reflected in the consistency in measurement methods and the small uncertainty of the experiment. In addition, to adapt to the new data collection method, the traditional equipment uncertainty calculation was replaced by the full width half maximum method. The experimentation overall demonstrates critical thinking and creativity.

5.2 Limitation

A limitation is the imprecision in the actual experimentation. Although the overall conclusions drawn from the results are consistent with theories and other scientific papers, the imprecision of the experiment can be seen from the Figure 10. For example, the first three trials are conducted on 3rd September, 2019, and the remaining five trials are conducted on 10th September, 2019. The separation of experiment dates can lead to imprecision in data, as the humidity can change. Humidity can affect the effectiveness of the ionisation net. This can be seen in the spread of data points at 0 volts (Figure 1). In theory, at 0 volts, the smoke density should be consistent, which is not the case in Figure 1.

On the other hand, despite the imprecision of the data, the percentage uncertainty is only 1.79%. From the graphs, it can be seen that a 1.79% error bar is acceptable and does not affect the overall trend.

Moreover, a limitation of the experiment is the use of ImageJ. Although the area analyzed by ImageJ is consistently right above the EP and with dimensions 1000 x 1000 pixels, the area is still selected manually. This is a problem because first, it is very time consuming; and second, it increases uncertainty. To improve upon this, a code could be written to perform histogram analysis within certain coordinates of the image. Assuming the picture taken every trial is consistent in terms of size and angle, this method will increase the precision and reduce the time for data analysis.

5.3 Implication

Based on these results, several conclusions can be made about EPs in real life. First, to enhance performance in the short term, the voltage can be increased. An increase in voltage will increase the particulate matter filtered per unit time, which means cleaner air.

In the long run, however, the performance of EP depends on its saturation point. Once the saturation point is reached, the EP will not be able to filter particulate matters as efficiently. At this point, the collector either has to be washed or changed completely. Combining these results, it can be concluded that although increasing the voltage will increase performance in the short run, it can shorten the time for the EP to reach its saturation point. Therefore, to optimize performance and maintain low cost, both voltage and surface area of the collector must be considered.

References

Bump, R. L.. The Use of Electrostatic Precipitators on Municipal Incinerators. *Journal of the Air Pollution Control Association*, 18(12), 803-809. doi:10.1080/00022470.1968.10469219

Aspelin, L. Å. Finding Full With at Half Maximum (on a graph) in Excel? Retrieved December 1, 2019, from answers.microsoft.com/en-us/office/forum/office_2010-excel/finding-full-wit h-at-half-maximum-on-a-graph-in/e3b1b2a0-2d90-4bf4-b583-3ed05399e6e5?db=5.

A Glossary of Spectroscopic Terms. (n.d.). Retrieved from oceanoptics.com/glossary/

ImageJ. (n.d.). Retrieved from https://imagej.net/Welcome

Mayer, H. Air pollution in cities. *Atmospheric Environment*, 33(24-25), 4029-4037. doi:10.1016/s1352-2310(99)00144-2

RimstarOrg. Electrostatic Precipitator/Smoke Precipitator - How it Works/How to Make. Retrieved from www.youtube.com/watch?v=x5YFK8mmeRQ

Helbing, B. Electrostatic Air Filters: How Do They Work? Retrieved from www.airtro.com/air-filters-2/electrostatic-air-filters-work

Particulate Matter (PM) Basics. (2018, November 14). Retrieved from www.epa.gov/pm-pollution/particulate-matter-pm-basics

Tang, M., Thompson, D., Chang, D., Chen, S., Pui, D. Y. Filtration efficiency and loading characteristics of PM2.5 through commercial electret filter media. *Separation and Purification Technology,* 195, 101-109.

doi:10.1016/j.seppur.2017.11.067

Walsh, D. C., Stenhouse, J. I. Parameters Affecting the Loading Behavior and Degradation of Electrically Active Filter Materials. *Aerosol Science and Technology*,29(5), 419-432. doi:10.1080/02786829808965580

Which Electrostatic Air Filter is the Best? Retrieved from howtohome.com/hvac/which-electrostatic-air-filter-is-the-best

Xing, Y., Xu, Y., Shi, M., Lian, Y. The impact of PM2.5 on the human respiratory system. Retrieved from www.ncbi.nlm.nih.gov/pmc/articles/PMC4740125/

為什麼燒香祭拜神不靈驗？隱藏了數千年的秘密！. (n.d.). Retrieved from bldaily.com/culture/p-376281.html

Isolation and Characterisation of Gut *E. coli* and *Klebsiella* sp. from *Psittaculidae*

Miriam M.-C. Cheng

Abstract

An increasing number of birds are being captured, bred and traded as companion animals (Grant *et al*., 2016), where they increase the potential for multidrug-resistant zoonotic[1] pathogens (*e.g.* Harada *et al*., 2016; Machado, 2018). In the wild, forest-dwelling birds do not normally stand in the faeces of other birds and tend to maintain nests carefully to reduce parasites. Caged birds, especially when kept at high density for commerce, frequently encounter a mixture of faeces, feathers and food scraps, which favors disease transmission and microbial proliferation (Doneley, 2009). Multidrug resistant strains have been reported to affect caged birds, infecting wounds that are hard to treat (Paraschiv *et al*., Guenther *et al*., 2010; Borges *et al*., 2017). Attaching and effacing *Escherichia coli* has been previously reported in budgerigars (Seeley, 2014). In a domestic setting, close confinement increases risk of exposure for both birds and humans (Pontes, 2018). This study uses a genomic analysis of faecal isolates from three *Psittaculidae* to provide an insight into hygiene amongst caged birds.

1. Introduction

Zoonotic pathogens refers to diseases or infections capable of transferring from vertebrate animals to humans. Societies are increasingly expanding into more rural areas. This increases animal-human-environment interactions, compounded with the exotic animal trade and consumption, leads to an increased risk of zoonotic diseases reaching humans. Recent outbreaks have often been traced to zoonotic sources. Studying and monitoring these pathogens are an important measure to protect the safety of both humans and pets.

Attaching and effacing *Escherichia coli* are dangerous strains known to cause lesions. They are dangerous to humans and pets. Plasmids are Plasmids are circular extrachromosomal DNA strands found in eukaryotes. They can be transmitted to other cells by horizontal gene transfer. Prophages are bacteriophage genomes inserted into bacterial chromosomal DNA. They show genetic transferral between samples. The gut microbiome is the collection of genomes of microorganisms that are present in an animal's gut. This is typically analyzed through faecal samples, which were collected from the birds individually.

2. Method

Source: Fresh faecal samples were collected from an 8-month-old eclectus parrot in a pet shop, a budgerigar bought from the same shop 4 months earlier, and a pet budgerigar from a different source purchased 16 months before. Though both species originate from Australia, the eclectus parrot was sourced from Indonesia. The light blue budgie and the eclectus parrot were both hatched and hand-reared by the shopkeeper.

Figure 1. Image of host birds and description of species, age, gender, status, source, possession status *etc.* Budgerigars on the left are pets, from which two *E. coli* samples were isolated. The eclectus parrot on the right is a store bird, from which an *E. coli* sample was isolated. The eclectus parrot on the right is a store bird, from which an *E. coli* sample and a *Klebsiella quasipneumoniae* sample was isolated.

The research for this article was conducte in the *Shuyuan* Molecular Biology Laboratory and posters about this work were prepared for the ASM Microbe Conference in 2020.

[1] Disease or infection capable of transferring from vertebrate animals to humans

Isolation and Extraction: After serial dilution in Phosphate-buffered saline, extracts were screened for *E. coli* and coliforms using 3M Petrifilm® plates and for *Salmonella* using CHROMagar *Salmonella* plus and also with CHROMagar Orientation medium. Selected colonies were transferred to LB agar and passaged for at least ten times. DNA was extracted from four isolates and shotgun sequencing was performed using the Illumina MiSeq platform.

Genomic Analysis: Draft sequences were assembled by Prinseq Lite and Newbler; isolates were identified using NCBI BLAST and autoMLST. Resistance, virulence genes and plasmids were found using CARD and CGE ResFinder, VirulenceFinder and PlasmidFinder. Prophages were found using PHASTER. Sequences were annotated using Patric. A type strain genomic phylogenetic tree was made by comparing *E. coli* strains to Avian Pathogenic Escherichia Coli (AVEC) strains found on NCBI Genbank.

3. Results

3.1 Identification of Strains:

Isolates from the budgerigars were confirmed to *be E. coli;* the eclectus parrot yielded both *E. coli* and *Klebsiella quasipneumoniae* (Johnson *et al*., 2008). Isolates from the two budgerigars were found to be *E. coli*; the eclectus parrot yielded one *E. coli* isolate and one *Klebsiella quasipneumoniae*. VirulenceFinder uggested that each *E. coli* isolate had over 0.93 probability of being pathogenic for humans (Cosentino *et al*., 2013), although no Shiga toxin genes were present (Joensen *et al*., 2014).

3.2 Phylogenetic Tree:

According to Type (Strain) Genome Server analysis of budgie E. coli isolates and Avian Pathogenic *E. coli* samples isolated in Australia (Cummins *et al*., 2019), budgie *E. coli* isolates showed high similarity to each other, scoring over 99% for all digital DNA:DNA hybridization scores. This strain may have been shared core microbiota in budgerigars, or transmission may have occurred between budgies within their 4 months of cage sharing.

Protein sequences in *E. colis* are identical though they contain small variations with other *E. colis* found in other HK animals, such as a difference in a beta-glucosidase protein between horse *E. coli* and parrot *E. coli* (TSI instead of TST), which was more similar to isolates from pandas and rats.

A phylogenetic tree comparing the *E. coli* samples to Avian Pathogenic *Escherichia Coli* (APEC) isolates, K-12 and an isolate from a Hong Kong panda HKUOPY1 shows that *E. coli* from budgies was on a close cluster branch to the panda isolate, but did not share much similarity with APEC. The *E. coli* sample from the eclectus parrot shared a branch with AVC235 and AVC38 isolated from Australia and APEC O1 *E. coli* isolated from a chicken (Sayers, 2019).

Sample ID	36	37	38	39
Source	Pet Budgie #1	Pet Budgie #2	Store eclectus	Store eclectus
No. of reads	238239	232647	248982	195649
Average depth	15	14	12	10
No. of large contigs	571	600	509	1174
No. of bases (large)	4695289	4667874	5252392	4788278
No. of all contigs	662	680	567	1174
N50 contig size	15768	15266	20086	6545
GC (%)	51.06	51.12	58.06	51.18
Est. genome size	4.8 MB	5.1 MB	6.3 MB	5.9 MB
Homology %	99.99%-100%	100%	99.1-99.83%	99.28-99.6
Identity	*Escherichia coli*	*Escherichia coli*	*Klebsiella quasipneumoniae*	*Escherichia coli*

Table 1. Isolate source and identification. 1 *E.* coli samples were isolated from each bird, and another *Klebsiella quasipneumoniae* sample was isolated from the eclectus store bird.

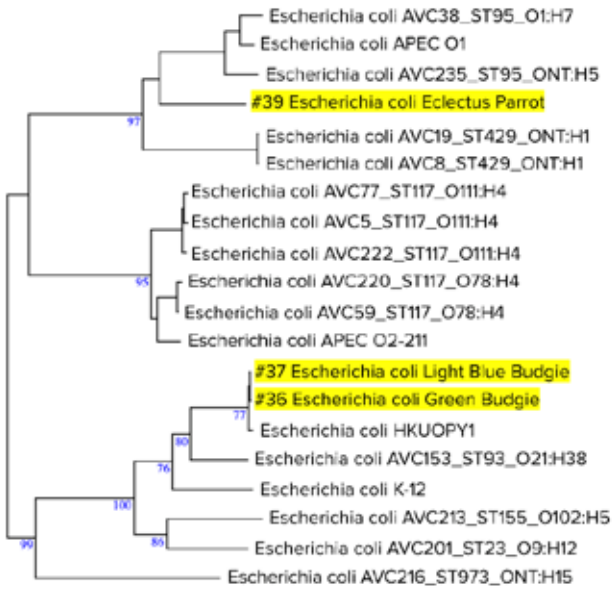


Figure 2. A phylogenetic tree uses DNA to display genetic relationship between organisms. Samples #36, #37 and #39 isolated from the green budgerigar, white budgerigar and eclectus parrot respectively are compared to similar *E. coli* samples. AVC are avian pathogenic strains isolated from Australian birds. HKUOPY1 is a panda *E. coli* strain. K-12 is a human *E. coli* strain

3.3 Virulence and Pathogenicity:

VirulenceFinder suggested that each *E. coli* isolate had over 0.93 probability of being pathogenic for humans, although no Shiga toxin genes were present.

Isolate	36	37	38	39
Virulence	cia, gad, ompT, terC	cia, gad, ompT, terC	--	cba, cea, chuA, cia, cma, fyuA, gad, ibeA, irp2, iss, kpsE, ompT, papC, pic, sitA, terC, traT, usp, vat, yfcV
Plasmid	IncI1, IncI2	IncI1, IncI2	IncKIB (F)	Col156, ColRNAI, IncFIB(AP001918), InFII(pSE11), IncX1

Table 2. Virulence and Pathogenicity of Isolates according to CGE analysis. Many plasmids were found in eclectus Parrot samples (#38 and #39), especially in #39. #36 and #37 have identical profiles. #38 has plasmid IncKIB(F), a plasmid also reported in other *Klebsiella* samples.

According to CGE analysis on pathogenicity, *E. coli* samples all had over 0.93 probability of being pathogenic for humans. No Shiga-toxin releasing genes were found. As shown in Table 2., plasmids show more similarities between budgerigar samples, likely revealing horizontal gene transfer. The *E. coli* isolate from the eclectus parrot had an abundance of virulence genes and plasmids.

Also shown in Table 2., the plasmid IncI was found in budgie *E. coli*, with genes increasing resistance to drug classes ceftriaxone and colistin (Rozwandowicz *et al.*, 2018). Plasmids IncFIB, IncFII and IncX1, identified in the eclectus isolates, are associated with resistance to drug classes tigecycline, β-lactams, aminoglycosides, carbapenems, and ciprofloxacin (Johnson, 2009). Plasmids were similar to those reported in Australian APEC strains (Cummins, 2019). The eclectus *E. coli* #39 shared many plasmids as Carbapenem-resistant *K. Pneumoniae* strains isolated at the Italian ASST Fatebenefratelli Sacco Hospital (Rimoldi, 2017).

3.4 Prophages:

According to PHASTER analysis, more intact phages were found in *K. quasipneumonia*, though eclectus *E. coli* contained more segments labelled questionable or incomplete segments. eclectus isolates had more phages despite the parrot being individually caged. This may show early acquisition from the parents or the environment; but given this bird was hand-reared from an egg by the shop owner, the latter is more likely.

Phage Completeness	36	37	38	39
Intact	1	1	3	
Incomplete	2			11
Questionable				2

Table 3. Number of Phages found within Isolates according to PHASTER analysis. eclectus *E. coli* contains various phages, though they are incomplete or questionable. eclectus *K. quasipneumoniae* contains three intact phages.

3.5 Antibiotic Resistance:

CARD Analysis

CARD analysis matched genes of isolates to known drug resistance genes.

CARD analysis showed similar multi-drug resistance (MDR) profiles in the budgerigar *E. coli* isolates: with several efflux pumps, a range of mdt genes, including mdtH, and, beta-lactamases ampC and ampH, as well as PBP3. MDR in the *Klebsiella quasipneumoniae* was more extensive, with a range of genes giving low susceptibility to cephalosporin, penam, fluoroquinolone and tetracycline drug classes; including TEM-1 and OKP-B-6 beta-lactamases which have been repeatedly associated with multidrug resistance in clinical isolates (Halsby *et al*., 2014).

The resistance profile of *K. quasipneumonia* shows multidrug resistance with efflux pumps and antibiotic target alteration with low susceptibility to cephalosporin, penam, fluoroquinolone and tetracycline drug classes. Resistance to drug classes OKP-B-6 and TEM-1 was also found in another *K. quasipneumonia* found in India (Shankar *et al*., 2019) and also in Brazil (Nicolás *et al*., 2018).

E. coli from budgerigars had 22 shared antibiotic resistance properties with one having an additional evgA resistance gene. *K. quasipneumonia* displayed less resistance, but contained the unique resistance gene tet(D). *E. colis* had more perfect and strict hits though budgerigar samples had more perfect hits (Zankari *et al*., 2012). This shows greater genetic alignment to resistance genes.

Kirby-Bauer Testing

Susceptibility testing of isolates with 20 antibiotics show *K. quasipneumonia* is resistant to various antibiotics, including ampicillin, clindamycin, erythromycin, rifamycin and vancomycin. *E. coli* isolates showed resistance to clindamycin and erythromycin. Budgerigar samples showed low susceptibility to vancomycin and rifamycin. #36 showed resistance to cefepime whilst #37 shows resistance to streptomycin.

Figure 4. #36 *E. coli* (white section) from the green pet budgerigar shows resistance to Ampicillin (left) as *E. coli* is able to grow near the disk. It is susceptible to Augmentin, Cefepime, and Azithromycin (shown by a clear zones, where the bacteria is unable to continue to grow).

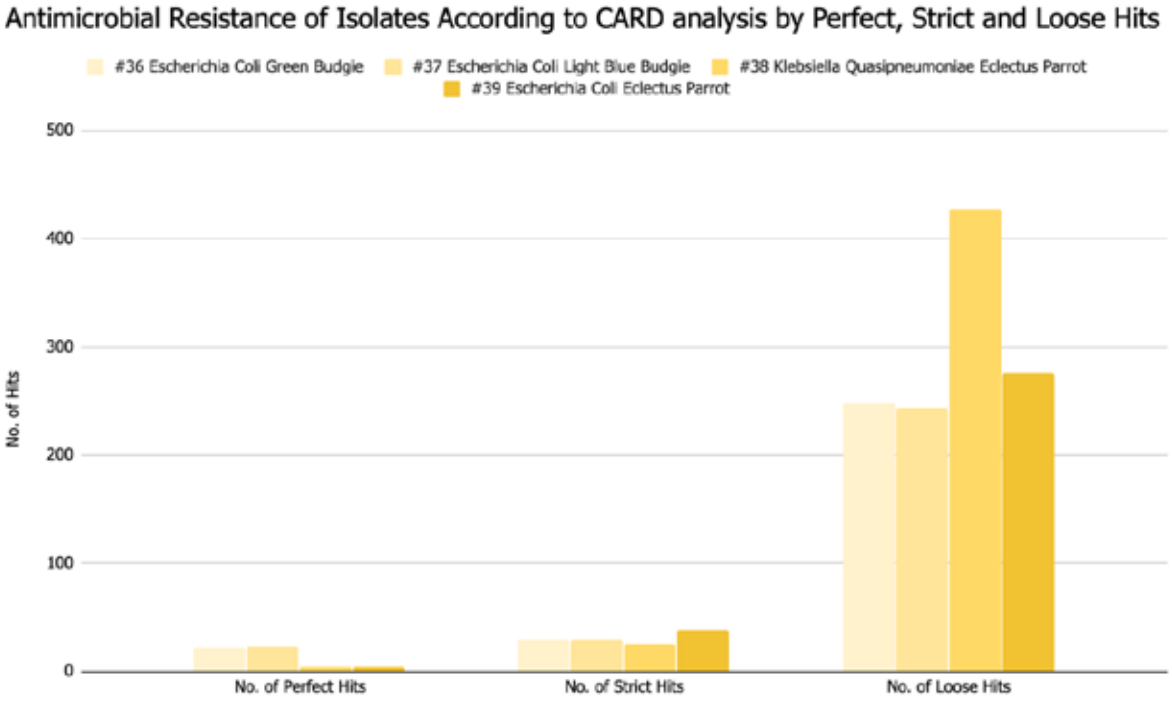


Figure 3. Analysis by The Comprehensive Antibiotic Resistance Database (CARD) on sample genomes shows that the *Klebsiella* isolate (#37) from the eclectus parrot contained many weak matches of antimicrobial resistance genes. #36 and #37, both *E. colis* from pet budgies similarity to each other and have many perfect matches of antimicrobial resistant genes.

Acquired Antibiotic Resistance

According to CGE analysis, *E. coli* isolates show acquired mdf(A) genes for macrolide drug classes. *K. quasipneumonia* shows a larger variety of acquired genes, including Aminoglycoside, Beta-lactam, Fosfomycin, Quinolone, Sulphonamide, Tetracycline and Trimethoprim (Zankari *et al*.,2012).

Sample ID	36	37	38	39
Acquired Antimicrobial Genes	mdf(A)	mdf(A)	aph(3")-Ib, aph(6)-Id, blaOKP-B-15, blaOKP-B-2, blaTEM-1B, fosA, oqxA, oqxB, sul2, tet(D), dfrA14	mdf(A)

Table 4. Acquired Antimicrobial Genes According to CGE Analysis. K. quasipneumoniae from the store eclectus Parrot #38 contains a high amount of acquired antibiotic resistance.

The bla_{TEM} genes confer resistance to broad-spectrum β-lactams, tet(D) is associated with resistance to tetracycline, and dfrA with Trimethoprim (Partridge *et al*., 2018).

3.6 *Klebsiella quasipneumoniae*:

The *Klebsiella quasipneumoniae* in the parrot was surprising because it carried an even larger number of resistance genes: beta-lactamases ampH, OmpK37, OKP-B-6 and TEM-1 (Alcock, 2020); as well as resistance to aminoglycosides, sulphonamides and tetracycline via APH(3")-Ib and APH(6)-Id; sul2 and tetD. Similar multidrug resistant *Klebsiella* samples have been reported in companion animals in Japan (Harada *et al*., 2016) and companion birds in Istanbul (Sigirci *et al*., 2020). The parrot isolates also contained a large number of plasmids. Given that this bird was very young, it is clear that this MDR *Klebsiella* strain had been acquired and not generated within the bird.

4. Conclusion

Budgerigar *E. coli* isolates are highly similar despite different origins, suggesting a spread of resistance between closely confined individuals. eclectus isolates contain more phages, likely from environmental acquisition given its young age, though resistance resistance genes carried more mutations. Isolation may be an effective method of controlling Avian pathogenic transmissions but more research is needed.

5. Discussion

Figure 6. Hong Kong Wanchai Bird Store X, source of budgerigars and eclectus parrot.

The incidence of MDR infections in avians and companion animals has been increasing and bird stores may be reservoirs for pathogenic strains (Nicolás, 2018). This may be due to a variety of factors: close proximity of birds in shared accommodation may allow for sharing of pathogens; trading of foreign species, particularly wild birds which carry plasmids antibiotic resistance (Dolejska *et al*., 2018) and antibiotic resistant pathogens (Parker *et al*., 2016), through routes such as Singapore (Aloysius *et al*., 2020), may introduce MDR pathogens to domestic birds. Hand rearing of altricial birds to increase tameness may prevent natural gut microbiome transfer, increasing susceptibility to environmental pathogens; and, stress from isolating social birds may also have unforeseen outcome on gut microbiome (Noguera *et al*., 2018).

Acknowledgements

Many thanks to Professor Fred Leung for his support in this project. Special thanks to the *Shuyuan* science team, in particular, Dr. Simon Griffin and Ms. Grace Lai.

References

Alcock, B.P., Raphenya, A.R., Lau, T.T.Y., *et al.* (2020). CARD 2020: Antibiotic resistome surveillance with the Comprehensive Antibiotic Resistance Database, *Nucleic Acids Research*, vol. 48, D517-D525 doi: 10.1093/nar/gkz935

Almeida, M., Cangussú, Í. M., Carvalho, A., Brito, I., Costa, R.A. (2017). Drug resistance, AmpC-β-lactamase and extended-spectrum β-lactamase-producing Enterobacteriaceae isolated from fish and shrimp. *Revista do Instituto de Medicina Tropical de Sao Paulo*, 59, e70. doi: 10.1590/S1678-9946201759070

Aloysius, S., Yong, D., Lee, J., Jain, A. (2020). Flying into extinction: Understanding the role of Singapore's international parrot trade in growing domestic demand, *Bird Conservation International*, 30(1), 139-155. doi: 10.1017/S0959270919000182

Arndt, D., Grant, J., Marcu, A., Sajed, T., Pon, A., Liang, Y., Wishart, D.S. (2016). PHASTER: a better, faster version of the PHAST phage search tool, Nucleic *Acids Research*, 44(W1): W16-21. doi: 10.1093/nar/gkw387

Borges, C.A., Beraldo, L.G., Maluta, R.P., Cardozo, M.V., Barboza, K.B., Guastalli, E.A., Kariyawasam, S., DebRoy, C., Ávila, F.A. (2017). Multidrug-resistant pathogenic *Escherichia coli* isolated from wild birds in a veterinary hospital, *Avian Pathology*, 46(1), 76–83. doi: 10.1080/03079457.2016.1209298

Brettin, T., Davis, J.J., Disz, T., Edwards, R.A., Gerdes, S., Olsen, G.J., Olson, R., Overbeek, R., Parrello, B., Pusch, G.D., Shukla, M., Thomason, J.A., 3rd, Stevens, R., Vonstein, V., Wattam, A.R., Xia, F. (2015). RASTtk: a modular and extensible implementation of the RAST algorithm for building custom annotation pipelines and annotating batches of genomes. *Scientific Reports*, 5, 8365. doi: 10.1038/srep08365

Carattoli, A., Zankari, E., García-Fernández, A., Voldby Larsen, M., Lund, O., Villa, L., Møller Aarestrup, F., Hasman, H. (2014). In silico detection and typing of plasmids using PlasmidFinder and plasmid multilocus sequence typing, *Antimicrobial Agents and Chemotherapy*, 58(7), 3895–3903. doi: 10.1128/AAC.02412-14

Cosentino, S., Voldby Larsen, M., Møller Aarestrup, F., Lund, O. (2013). PathogenFinder – distinguishing friend from foe using bacterial whole genome sequence data, *PloS One*, 8(10), e77302. doi: 10.1371/journal.pone.0077302

Cummins, M. L., Reid, C. J., Roy Chowdhury, P., Bushell, R. N., Esbert, N., Tivendale, K. A., Noormohammadi, A. H., Islam, S., Marenda, M. S., Browning, G. F., Markham, P. F., Djordjevic, S. P. (2019). Whole genome sequence analysis of Australian avian pathogenic *Escherichia coli* that carry the class 1 integrase gene, *Microbial Genomics*, 5(2), e000250. doi: 10.1099/mgen.0.000250

Dolejska, M., and Papagiannitsis, C.C. (2018). Plasmid-mediated resistance is going wild, *Plasmid*, 99, 99–111. doi: 10.1016/j.plasmid.2018.09.010

Doneley R.J. (2009) Bacterial and parasitic diseases of parrots, *Veterinary Clinics of North America: Exotic Animal Practice*, 12(3), 417-432. doi: 10.1016/j.cvex.2009.06.009

Gerlach, H. (1999). Bacteria. In: Ritchie, B.W., Harrison, G.J., Harrison, LR. (Eds.), *Avian Medicine: Principles and Application*. HBD International, Inc. (Delray Beach, FL, USA); Ch. 33, 949-983.

Goldstein B.P. (2014). Resistance to rifampicin: a review, *The Journal of Antibiotics*, 67(9), 625–630. doi: 10.1038/ja.2014.107

Grant, R.A., Montrose, V.T., Wills, A.P. (2017). ExNOTic: Should We Be Keeping Exotic Pets?. *Animals*, 7(6), 47. doi: 10.3390/ani7060047

Guenther, S., Grobbel, M., Lübke-Becker, A., Goedecke, A., Friedrich, N.D., Wieler, L.H., Ewers, C. (2010). Antimicrobial resistance profiles of Escherichia coli from common European wild bird species, *Veterinary Microbiology*, 144(1-2), 219–225. doi: 10.1016/j.vetmic.2009.12.016

Guerra, B., Fischer, J., & Helmuth, R. (2014). An emerging public health problem: acquired carbapenemase-producing microorganisms are present in food-producing animals, their environment, companion animals and wild birds. *Veterinary Microbiology*, 171(3-4), 290–297. doi.org/10.1016/j.vetmic.2014.02.001

Halsby, K.D., Walsh, A.L., Campbell, C., Hewitt, K., Morgan, D. (2014). Healthy animals, healthy people: zoonosis risk from animal contact in pet shops, a systematic review of the literature, *PLoS One*, 9(2): e89309. doi: 10.1371/journal.pone.0089309

Harada, K., Shimizu, T., Mukai, Y., Kuwajima, K., Sato, T., Usui, M., Tamura, Y., Kimura, Y., Miyamoto, T., Tsuyuki, Y., Ohki, A., Kataoka, Y. (2016). Phenotypic and Molecular Characterization of Antimicrobial Resistance in *Klebsiella* spp. Isolates from Companion Animals in Japan: Clonal Dissemination of Multidrug-Resistant Extended-Spectrum β-Lactamase-Producing *Klebsiella pneumoniae, Frontiers in Microbiology,* 7, 1021. doi: 10.3389/fmicb.2016.01021

Henderson, T.A., Young, K.D., Denome, S.A., Elf, P.K. (1997). AmpC and AmpH, proteins related to the class C beta-lactamases, bind penicillin and contribute to the normal morphology of *Escherichia coli, Journal of Bacteriology*, 179(19), 6112–6121. doi: 10.1128/jb.179.19.6112-6121.1997

Joensen, K.G., Scheutz, F., Lund, O., Hasman, H., Kaas, R.S., Nielsen, E.M., Aarestrup, F.M. (2014). Real-time whole-genome sequencing for routine typing, surveillance, and outbreak detection of verotoxigenic *Escherichia coli, Journal of Clinical Microbiology*, 52(5), 1501–1510. doi: 10.1128/JCM.03617-13

Johnson, M., Zaretskaya, I., Raytselis, Y., Merezhuk, Y., McGinnis, S., Madden, T.L. (2008). NCBI BLAST: a better web interface, *Nucleic Acids Research*, 36 (Web Server issue): W5-W9. doi: 10.1093/nar/gkn201

Johnson, T.J., Nolan, L.K. (2009). Pathogenomics of the virulence plasmids of *Escherichia coli, Microbiol Mol Biol Rev*., 73(4): 750-774. doi: 10.1128/MMBR.00015-09 [published correction appears in *Microbiol Mol Biol Rev*. 2010 Sep; 74(3): 477–478].

Johnson, T.J., Aziz, M., Liu, C.M., Sokurenko, E., Kisiela, D.I., Paul, S., Andersen, P., Johnson, J.R., Price, L.B. (2016). Complete Genome Sequence of a CTX-M-15-Producing *Escherichia coli* Strain from the H30Rx Subclone of Sequence Type 131 from a Patient with Recurrent Urinary Tract Infections, Closely Related to a Lethal Urosepsis Isolate from the Patient's Sister, *Genome Announcements*, 4(3), e00334-16. doi: 10.1128/genomeA.00334-16

Köck, R., Daniels-Haardt, I., Becker, K., Mellmann A., Friedrich, A.W., Mevius, D., Schwarz, S., Jurke, A. (2018). Carbapenem-resistant *Enterobacteriaceae* in wildlife, food-producing, and companion animals: a systematic review, *Clinical Microbiology and Infection*, 24(12), 1241-1250. doi: 10.1016/j.cmi.2018.04.004

Machado, D.N., Lopes, E.S., Albuquerque, A.H., Horn, R.V., Bezerra, W.G.A., Siqueira, R.A.S., Lopes, I.T., Nunes, F.P., Teixeira, R.S.C., Cardoso, W.M. (2018). Isolation and Antimicrobial Resistance Profiles of *Enterobacter*ia from Nestling Grey-Breasted Parakeets (*Pyrrhura griseipectus*), *Brazilian Journal of Poultry Science*, 20(1), 103-110. doi: 10.1590/1806-9061-2017-0551

Meguenni, N., Chanteloup, N., Tourtereau, A., Ahmed, C.A., Bounar-Kechih, S., Schouler, C. (2019). Virulence and antibiotic resistance profile of avian *Escherichia coli* strains isolated from colibacillosis lesions in central of Algeria, *Veterinary World*, 12(11), 1840–1848. doi: 10.14202/vetworld.2019.1840-1848

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

Nicolás, M.F., Ramos, P., Marques de Carvalho, F., Camargo, D., de Fátima Morais Alves, C., Loss de Morais, G., Almeida, L., Souza, R. C., Ciapina, L. P., Vicente, A., Coimbra, R.S., Ribeiro de Vasconcelos, A.T. (2018). Comparative Genomic Analysis of a Clinical Isolate of *Klebsiella quasipneumoniae* subsp. *similipneumoniae*, a KPC-2 and OKP-B-6 beta-Lactamases Producer Harboring Two Drug-Resistance Plasmids from Southeast Brazil, *Frontiers in Microbiology,* 9, 220. doi: 10.3389/fmicb.2018.00220

Noguera, J.C., Aira M, Pérez-Losada, M, Domínguez, J, Velando, A. (2018). Glucocorticoids modulate gastrointestinal microbiome in a wild bird, *Royal Society Open Science*, 5(4): 171743. doi: 10.1098/rsos.171743

Paraschiv, I., Stoian, A., Tasbac, B., Soare, T., Visoui, C., Militaru, M. (2017.). Inflammatory lesions in cases of birds kept in captivity. *Scientific Works. Veterinary Medicine*, LXIII(1), c, 17-20.

Parker, D., Sniatynski, M.K., Mandrusiak, D., Rubin, J.E. (2016). Extended-spectrum β-lactamase producing *Escherichia coli* isolated from wild birds in Saskatoon, Canada. *Letters in Applied Microbiology*, 63(1), 11–15. doi: 10.1111/lam.12589

Partridge, S.R., Kwong, S.M., Firth, N., Jensen, S.O. (2018). Mobile Genetic Elements Associated with Antimicrobial Resistance, *Clinical Microbiology Reviews*, 31(4), e00088-17. doi: 10.1128/CMR.00088-17

Pontes, P.S., Coutinho, S., Iovine, R.O., Cunha, M., Knöbl, T., Carvalho, V.M. (2018). Survey on pathogenic Escherichia coli and *Salmonella* spp. in captive cockatiels (*Nymphicus hollandicus*), *Brazilian Journal of Microbiology*, 49(Suppl 1), 76–82. doi: 10.1016/j.bjm.2018.05.003

Rimoldi, S.G., Gentile, B., Pagani, C., Di Gregorio, A., Anselmo, A., Palozzi, A.M., Fortunato, A., Pittiglio, V., Ridolfo, A.L., Gismondo, M.R., Rizzardini, G., Lista, F. (2017). Whole genome sequencing for the molecular characterization of carbapenem-resistant *Klebsiella pneumoniae* strains isolated at the Italian ASST Fatebenefratelli Sacco Hospital, 2012-2014, *BMC Infectious Diseases*, 17(1), 666. doi: 10.1186/s12879-017-2760-7

Rozwandowicz, M., Brouwer, M.S.M., Fischer, J., Wagenaar, J.A., Gonzalez-Zorn, B., Guerra, B., Mevius, D.J., Hordijk, J. (2018). Plasmids carrying antimicrobial resistance genes in *Enterobacteriaceae, Journal of Antimicrobial Chemotherapy*, 73(5), 1121–1137. doi: 10.1093/jac/dkx488

Sayers, E. W., Agarwala, R., Bolton, E. E., Brister, J. R., Canese, K., Clark, K., Connor, R., Fiorini, N., Funk, K., Hefferon, T., Holmes, J. B., Kim, S., Kimchi, A., Kitts, P. A., Lathrop, S., Lu, Z., Madden, T. L., Marchler-Bauer, A., Phan, L., Schneider, V. A., … Ostell, J. (2019). Database resources of the National Center for Biotechnology Information. *Nucleic Acids Research*, 47(D1), D23–D28. doi.org/10.1093/nar/gky1069

Seeley, K.E., Baitchman, E., Bartlett, S., DebRoy, C., Garner, M.M. (2014). Investigation and control of an attaching and effacing *Escherichia coli* outbreak in a colony of captive budgerigars (Melopsittacus undulatus), *Journal of Zoo and Wildlife Medicine*, 45(4), 875–882. doi: 10.1638/2012-0281.1

Shankar, C., Karunasree, S., Manesh, A., Veeraraghavan, B. (2019). First Report of Whole-Genome Sequence of Colistin-Resistant *Klebsiella quasipneumoniae* subsp. *similipneumoniae* Producing KPC-9 in India, *Microbial Drug Resistance*, 25(4), 489–493. doi: 10.1089/mdr.2018.0116

Sigirci, B.D., Celik, B., Halac, B., Adiguzel, M.C., Kekec, I., Metiner, K., Ikiz, S., Funda Bagcigil, A., Yakut Ozgur, N., Ak, S., Kahraman, B.B. (2020). Antimicrobial resistance profiles of *Escherichia coli* isolated from companion birds, *Journal of King Saud University - Science*, 32(1), 1069-1073. doi: 10.1016/j.jksus.2019.09.014

Wattam, A.R., Davis, J.J., Assaf, R., Boisvert, S., Brettin, T., Bun, C., Conrad, N., Dietrich, E. M., Disz, T., Gabbard, J.L., Gerdes, S., Henry, C.S., Kenyon, R.W., Machi, D., Mao, C., Nordberg, E.K., Olsen, G.J., Murphy-Olson, D.E., Olson, R., Overbeek, R., Parrello, B., Pusch, G.D., Shukla, M., Vonstein, V., Warren, A., Xia, F., Yoo, H., Stevens, R.L. (2017). Improvements to PATRIC, the all-bacterial Bioinformatics Database and Analysis Resource Center, *Nucleic Acids Research*, 45(D1), D535–D542. doi: 10.1093/nar/gkw1017

Zankari, E., Hasman, H., Cosentino, S., Vestergaard, M., Rasmussen, S., Lund, O., Aarestrup, F.M., Larsen, M.V. (2012). Identification of acquired antimicrobial resistance genes. *The Journal of Antimicrobial Chemotherapy*, 67(11), 2640–2644. doi: 10.1093/jac/dks261

Zhou, Y., Liang Y., Lynch, K.H., Dennis, J.J., Wishart, D.S. (2011). PHAST: a fast phage search tool, *Nucleic Acids Research*, 39 (Web Server issue): W347-W352. doi: 10.1093/nar/gkr485

Antimicrobial Resistance in Shiga Toxin-producing Strains of *Escherichia coli* from the New Territories, Hong Kong

Scarlett A. Wright and Valerie M. Stacey

Abstract

Whilst *Escherichia coli* mostly occurs as a benign gut commensal in warm-blooded animals, pathogenic strains contaminating food are a frequent cause of nausea, vomiting and even death. Shiga toxin-producing strains of *E. coli* (STECs), in particular, are associated with haemolytic uraemic syndrome (HUS), causing acute kidney injury in many of those infected. This study uses draft and complete genomic sequencing to investigate the antimicrobial resistance and virulence of six STECs isolated from cows and goats at a small-holding in the New Territories, Hong Kong. Whilst frequently characterised as highly pathogenic, these STEC isolates proved less aggressive than expected despite possessing the means to become so. It is possible that the functional characteristics of these isolates reflect the fact that their host animals are grass-fed and raised organically.

1. Introduction

Escherichia coli (E. coli) is a Gram-negative bacterium that resides in the gastrointestinal tracts of humans and most warm-blooded animals. Whilst most *E. coli* is harmless and has a mutually-beneficial relationship with the host, certain strains, such as those that produce Shiga toxin, are pathogenic (Kaper *et al.,* 2004). Shiga toxin-producing strains of *E. coli* (STECs) and of *Shigella dysenteriae* can cause severe illness in humans, which can lead to large community outbreaks particularly if the source of infection is not quickly identified (Melton-Celsa, 2014).

STECs are associated with ruminant animals, especially cattle, which have been identified as a reservoir of STECs within their guts (Sandhu and Gyles, 2002). Symptoms of human infection are diarrhoea, haemorrhagic colitis, and the life-threatening haemolytic-uraemic syndrome (HUS) (Nataro and Kaper, 1998). HUS causes acute kidney injury and often develops in children, around half of whom will require dialysis (Karpman *et al*., 2017; Scheiring *et al*., 2010).

Large-scale infections by STECs are well-documented, for example an O157:H7 strain of *E. coli* in Oregon and Michigan in 1982, led to hamorrhagic colitis in 47 people (with symptoms including abdominal cramps and bloody diarrhoea) and was linked to under-cooked burgers at a fast-food chain (Riley *et al*., 1983). An even larger outbreak involving O157:H7 took place across 7 neighbouring counties in Jiangsu and Anhui, China in 1999. This outbreak resulted in 195 cases of HUS and 177 deaths. No single source of the outbreak was determined, instead the STEC strain was traced to farm animals living alongside villagers, who practised poor hygiene, and also to flies as vectors transferring the bacterium to food (Xiong *et al*., 2012).

More recently, in 2011 in Germany, a Shiga-toxin producing strain of *E. coli* was associated with the largest outbreak ever recorded: a total of 3,842 cases, including 855 cases of HUS and 53 deaths. In this case, the O104:H4 strain was eventually traced to contaminated beansprouts, but also possible secondary infections (Muniesa *et al.,* 2012). However, the pathogenicity of the strain was not only due to the presence of the *stx2* Shiga toxin gene, but also a set of virulence genes that made the bacterium stick better to the walls of the intestine (adherence factors). These were: *iha, lpfO26, lpfO113*, as well *as aggA, aggR, set1, pic, aap*. The bacterium also contained extended-spectrum beta-lactamase (ESBL) genes, which made it harder to treat with antibiotics (Bielaszewska *et al*., 2011).

The research for this article was conducted in the *Shuyuan* Molecular Biology Laboratory and a poster about this work was prepared for the ASM Microbe Conference in 2020.

In fact, previous research has found that STECs generally have higher frequencies of antimicrobial drug resistance than other strains of *E. coli* and this can contribute to an increased severity of diseases (Mukherjee *et* al., 2017). For example, Li *et al*. (2011), sampling meat products from Jilin in China, showed that seventeen of twenty STEC isolates tested were multiresistant to a range of antibiotics including ampicillin, ciprofloxacin, tetracycline, gentamycin, and streptomycin. All twenty samples were also resistant to sulfamethoxazole-trimethoprim. These antimicrobial resistant bacteria would not only be harder to combat with antibiotics, but also costlier and more persistent.

This paper examines six STEC isolates found amongst eight *E. coli* isolates recovered from a herd of cattle (5 of which were Shiga toxigenic) and six isolates from a herd of goats (including 1 STEC) living in the same small-holding in Hong Kong. It seems surprising to find Shiga toxin-producing bacteria in this context: all of these animals are pasture-fed and organically-raised (and have never been prescribed antibiotics or any other medication) so that there is no obvious selection pressure driving virulence.

2. Methods

8 cow faecal samples and 6 goat faecal samples were obtained from healthy animals in a small, organically-raised pasture-fed herd of Braunvieh-Brahman cattle and Asian black goats respectively. Both herds of animals live in the same area in New Territories, Hong Kong.

After suspension in PBS and serial dilution, extracts were screened for *E. coli* using 3M Petrifilm® plates after which chosen isolates were streaked for ten generations on LB agar plates. Purified isolates were then used for further study; their DNA was extracted via an Invitrogen PureLink® DNA Mini Kit and sequenced via the Illumina MiSeq platform to give draft genomes. Long-read sequences obtained via the Nanopore MinIon were then combined with the Illumina short-reads using Unicycler (Wick *et al.,* 2017) to generate complete genomes and plasmids. VirulenceFinder2.0 (Joensen *et al*., 2014), PlasmidFinder (Carattoli *et al.,* 2014), PHASTER (Arndt et al., 2016) and Atom were used to identify Shiga Toxin-producing strains and other potential sources of virulence; Uniprot (The UniProt Consortium, 2019) and NCBI BLAST (Altschul *et al.,* 1990) confirmed Shiga variants and mapped protein alignments; RGI 5.1.0 at CARD (Alcock *et al*., 2020) was used to identify antimicrobial resistance (AMR) genes and KEGG pathways were mapped using GhostKoala (Kanehisa *et al*., 2016). Agar diffusion assays were used to determine antibiotic susceptibility.

3. Results

3.1 Overview of Strains

Sample ID	C2(I1)	C2(I2)	C3(I1)	C4(I1)	C4(I2)	G2(12)
Source	Juvenile Cow	Juvenile Cow	Adult Cow	Calf	Calf	Goat
Serotype	O146:H21	O-:H21	O146:H21	O-:H12	O-:H12	O179:H8
Shiga toxin	stx1d	stx1d	stx1d	stx1a	stx1a	stx2b
Other virulence factors	gad lpfA	gad lpfA	gad lpfA	gad lpfA iss cdtB espP iha	gad lpfA iss	gad lpfA celb espI subA

Figure 1. The characteristics of the six Stx-positive isolates

Isolates were identified by comparing their sequences to NCBI's nucleotide database: NCBI BLAST identified all isolates as *Escherichia coli*. Identification of virulence factors and Stx variants was carried out using VirulenceFinder2.0 (Joensen *et al*., 2014); serotypes were found using SerotypeFinder2.0 (Joensen *et al*., 2015).

3.2 KEGG pathways and Serotypes

KEGG results showed that all 14 isolates (including non-STEC isolates) had similar if not identical biochemical pathways available.

The serotypes of each Stx isolate showed more promising results regarding not only virulence potential, but also sub-species difference. All Stx isolates were found to be non-O157 STECs. Non-O157 STECs can cause sickness as severe as O157 and can be considered to be 'emerging pathogens' (Mathusa *et al*., 2010). Previous research has also shown that antimicrobial resistance was found to be more common in non-O157 STECs which could contribute to more severe cases of disease outbreaks (Mukherjee *et al*., 2017).

While each cow isolate had similar O and H antigens with H21, H12 and O146 appearing in more than one isolate and the goat isolate had completely different antigens. Although, since there is only one goat STEC isolate, a conclusive comparison cannot be drawn. Regarding the cattle isolates, in previous studies these serotypes and serogroups have not been reported and may be related to the fact that our bovine isolates were extracted from grass-fed and 'organically' raised cows while other studies focused on commercially-raised beef cattle (Fan *et al*., 2019; Mekata *et al*., 2014).

3.3 Shiga Toxin and other Virulence Factors

Further investigation into the specific Shiga Toxin genes once again showed a distinction between the cow isolates and goat isolate. All cow isolates had the Stx1 genes (2 with Stx1a; 3 with Stx1d) while the goat isolate had the Stx2b gene. While this seems to indicate a difference between goat and cow isolates with regard to Stx variants, more goat STEC isolates will need to be isolated and examined to confirm this.

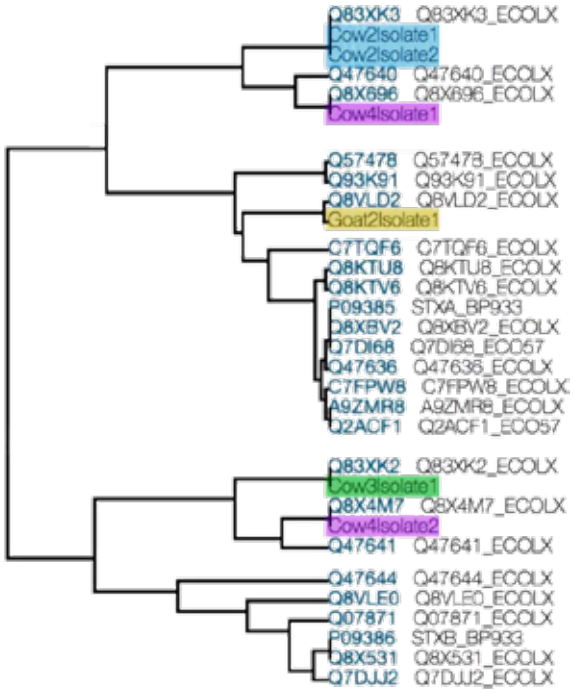


Figure 2. Phylogenetic tree based on Shiga toxin genes (made using Uniprot)

Figure 2 is a phylogenetic tree with our isolates highlighted. The other isolates which we compared our isolates to are STECs from the Uniprot (The UniProt Consortium, 2019) database. It is interesting to note that the goat isolate is genetically more similar to the other cow isolates than expected. It was also an unexpected result that certain cow isolates (C3(I1) and C4(I2)) were genetically distant to other cow isolates. In the case of the isolates C4(I1) and C4(I2), even though they were from the same cow, they were relatively distant on the phylogenetic tree.

Other virulence factors were present in the isolates including the enzyme *gad* and the long polar fimbriae *IpfA* which were present in all isolates. Virulence factors are described as molecules that aid a bacterium to colonize its host. Previous research has shown that *gad* (glutamate decarboxylase) could be used as a screener for pathogenic *E. coli* as it is present in most if not all enterovirulent *E. coli* isolates (Grant *et al*., 2001). This is not surprising regarding the STEC isolates, however our non-STECs isolates also had this enzyme. The long polar fimbriae, *IpfA*, is also a common virulence factor associated with human diarrhoeal diseases (Afset *et al*., 2006; Torres *et al*., 2009).

Despite the presence of these virulence factors, these STEC isolates are not as pathogenic as those strains reported in major outbreaks since they do not possess a large number of adherence factors. Most isolates have fewer than 4 virulence factors (all isolates excluding C4(I1) and G2(I2)), which greatly diminishes their ability to remain in the intestine long enough to cause harm.

3.4 Antibiotic Resistance and Plasmids

Antimicrobial resistance (AMR) genes can be detected in genomes using the Resistance Genes Identifier (RGI 5.1.0) at CARD (Alcock *et al*., 2020). However, for the determination of functional resistance or susceptibility to antibiotics, agar diffusion assays were used. Results of these are shown respectively in Figure. 3 and Figure. 4.

C2(I1)	C2(I2)	C3(I1)	C4(I1)	C4(I2)	G2(12)
Juvenile Cow	Juvenile Cow	Adult Cow	Calf	Calf	Goat
acrA	acrA	acrA	acrA	acrA	acrA
acrB	acrB	acrB	acrB	acrB	
AcrE	AcrE	AcrE	AcrE	AcrE	AcrE
					AcrS
ampH	ampH	ampH	ampH	ampH	ampH
baeR	baeR	baeR			baeR
					baeS
	cpxA (x2)	cpxA	cpxA	cpxA	cpxA
emrB	emrB	emrB	emrB	emrB	emrB
emrR	emrR	emrR	emrR	emrR	emrR
evgA	evgA	evgA	evgA	evgA	evgA
H-NS		H-NS	H-NS		H-NS
marA	marA	marA	marA	marA	marA
mdtE	mdtE	mdtE	mdtE	mdtE	
mdtG	mdtG	mdtG	mdtG	mdtG	mdtG
			mdtH	mdtH	
msbA	msbA	msbA	msbA	msbA	msbA
TolC	TolC	TolC	TolC	TolC	TolC

Figure 3. Table documenting all complete AMR genes identified by CARD

	C2(I1)	C2(I2)	C3(I1)	C4(I1)	C4(I2)	G2(12)
Ampicillin	s	s	s	s	s	s
Kanamycin	s	s	s	s	s	s
Tetracycline	s	s	s	s	s	s
Clindamycin	**R**	**R**	**R**	**R**	**R**	**R**
Erythromycin	**R**	s	s	**R**	s	s
Vancomycin	s	s	s	s	s	**R**

Figure 4. Table documenting selected results from antibiotic susceptibility tests. (R - resistant; s - sensitive)

In Figure 3, perfect hits were identified by the Resistance Genes Identifier (RGI 5.1.0) at CARD (Alcock *et al*., 2020). Interestingly, the goat isolate shows resistance to vancomycin, although this was not shown by the others. Vancomycin is often used as a last resort for bacterial infections and vancomycin resistant bacteria can be detrimental in a clinical environment (Mcdonald, 1997; Rice, 2001) as medical practitioners have limited resources to combat the bacteria. Another aspect that caught our attention was that even though each isolate contains abundant chromosome-encoded AMR genes, they still proved susceptible to many antibiotics. Notably, even though all isolates contain the beta-lactamase *ampH* gene, they all still proved sensitive to ampicillin. *AmpH* is a protein related to class C beta-lactamases and it binds penicillin (Henderson *et al.,* 1997) which is why this result is curious. Investigation into the specific *ampH* genes in the isolates led us to the discovery that in all isolates, the *ampH* protein sequence began with leucine (L) instead of the typical methionine (M) as it was reading a TTG start codon instead of an ATG start codon (Koonin and Novozhilov, 2009). A start codon sets the reading frame which determines which amino acids are to be added to the protein. An unusual start codon may decrease gene expression as it does not get recognised as often by the ribosome which means less mRNA is translated (Louten, 2016), lowering the potency of the whole gene.

Profiles of the plasmids in each isolate also help to explain the limited AMR each isolate has. Plasmids are autonomous DNA molecules and are typically seen as carriers for AMR genes (Veeraraghavan *et al*., 2019; Villa *et al*., 2010). Plasmids can transfer these genes in multiple ways via horizontal gene transfer. However, in our isolates, all of the AMR genes were found to be chromosome-encoded and not plasmid-encoded.

Isolate	C2(I1)	C2(I2)	C3(I1)	C4(I1)	C4(I2)	G2(I2)
IncFIB	✓	✓	✓	✓	✗	✗
IncFIC	✓	✓	✓	✗	✗	✗
IncFII	✗	✗	✗	✓	✗	✓

Figure 5. IncF plasmids found in each isolate

Out of all plasmid types, IncF plasmids are reported as the type that usually carry the greatest variety of AMR genes (Rozwandowicz *et al*., 2018). Initially during the research, an assumption was made that as these isolates had IncF plasmids, this must be where the AMR genes originated from. However, full genome sequencing later showed that this was not the case.

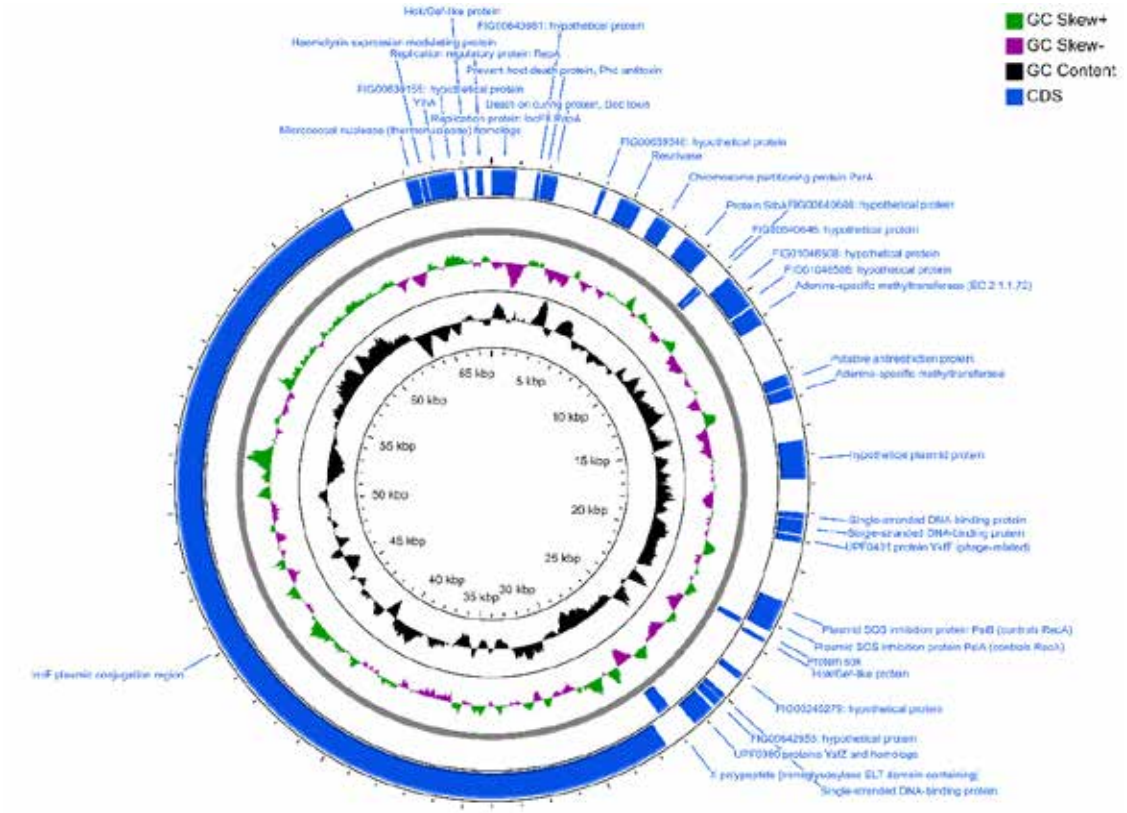


Figure 6. One of the two plasmids in isolate C3(I1) mapped out; This plasmid is a hybrid plasmid (Full images can be found in supplementary resources)

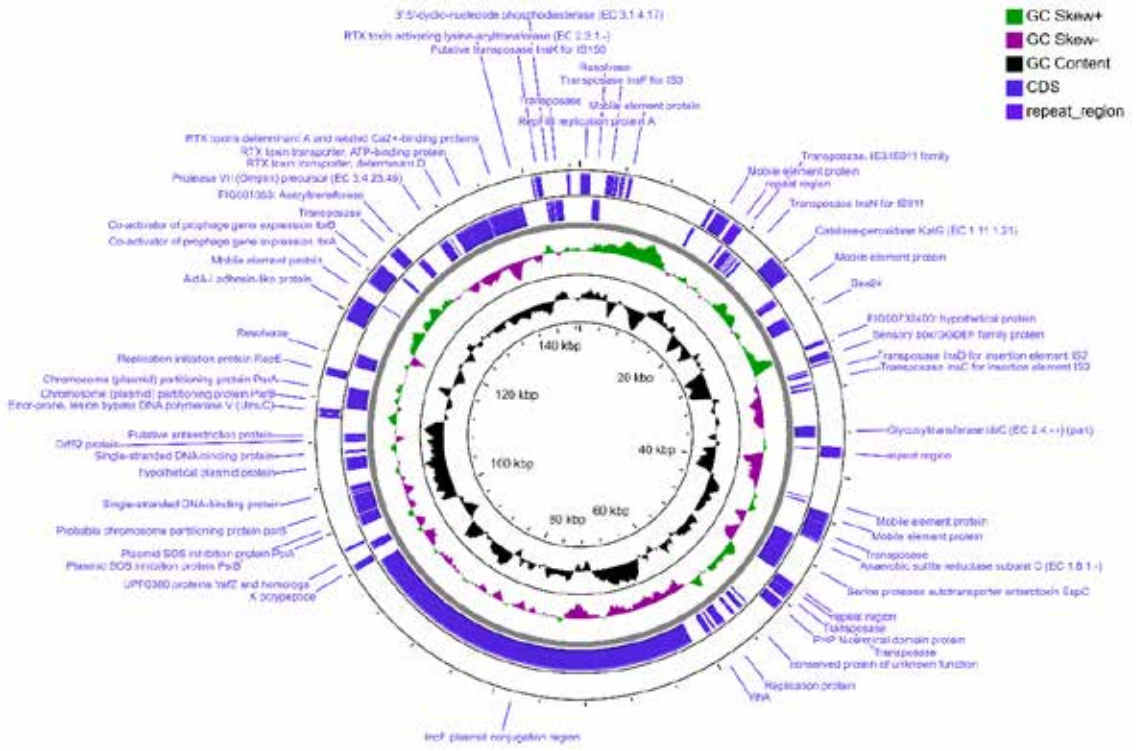


Figure 7. The IncF plasmid found in G2(I2) mapped out (Full image can be found in supplementary resources)

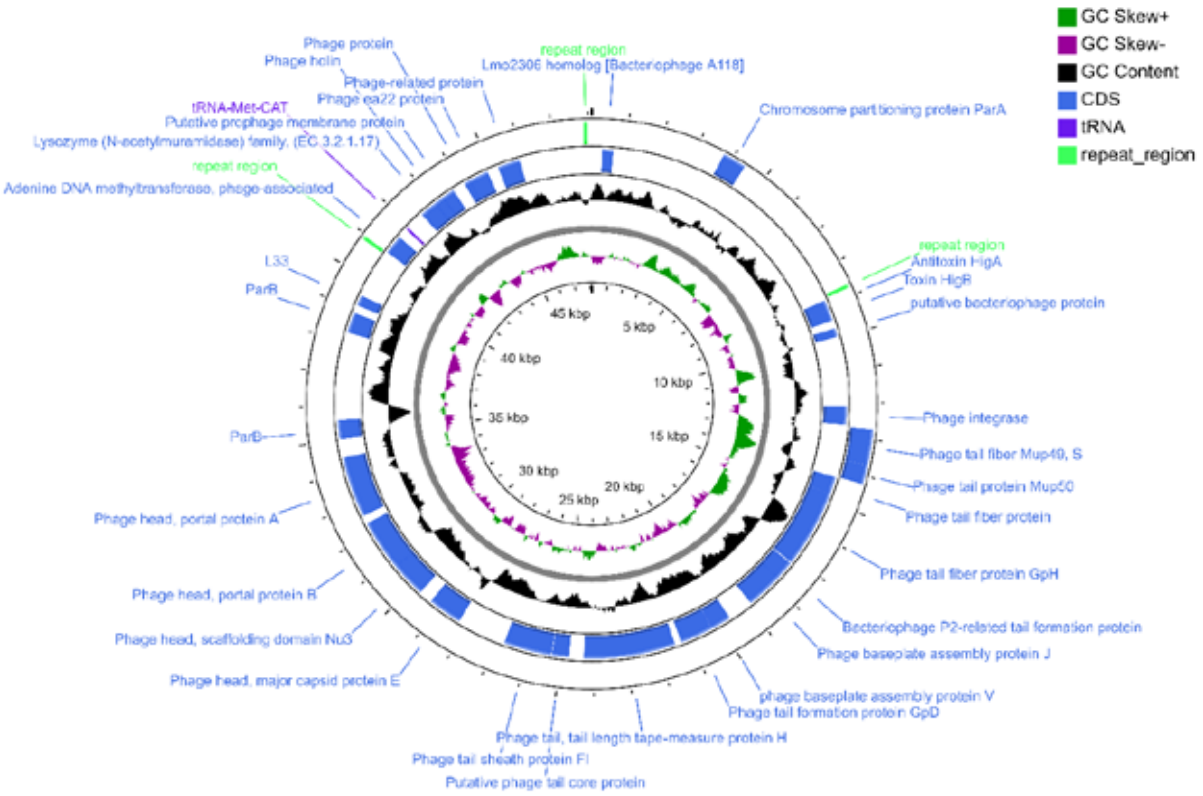


Figure 9. The phage-like plasmid mapped out (Full image can be found in supplementary resources)

Among these IncF plasmids, no AMR were present and few virulence genes. Instead, all the plasmids carried toxin-antitoxin (TA) systems (shown in Figures 6 &7). These plasmids carried these TA systems in order to secure their survival and make them indispensable. In Figure 6, a plasmid isolated from C3(I1), TA systems such as Doc-Phd, CcdA-CcdB and Repeats-in-toxin (RTX), are prominent. The TA system Hok/Gef-Sok was present in the goat isolate (shown in Figure 7). These TA systems not only help keep the plasmid stable, but it also helps the plasmid proliferate. It does this by producing the protein toxins which are then exported out of the cells. In the cells with the plasmids, antitoxins are created so that the producer cell stays safe. However, neighbouring cells do not have these antitoxins and get killed (Unterholzner *et al*., 2013; Melderen and Bast, 2009). Doc-Phd contributes to the perpetuation of its plasmid by having the toxin Doc inhibit translation via blocking the ribosomal A site, this is then countered by the antitoxin Phd (Garcia-Pino et *al*., 2010). CcdA-CcdB contributes to the stability of the plasmid by post-segregational killing (PSK) (Coray *et al*., 2017) daughter cells which contain no plasmids (Unterholzner *et al*., 2013; Ogura and Hiraga, 1983). RTX form sites for Ca2+ ions to bind which then pulls the RTX out of the cell and drives the folding of the toxin (Linhartová *et al*., 2010). They can also play a role in virulence of the pathogen (Linhartová *et al*., 2010). Many plasmids also carry colicin, a protein which attacks other bacteria reducing competition from other bacterial strains and keeping the host safe (Cascales *et al*., 2007).

The plasmids in our isolates are primed to pick up and transfer genes with numerous transposases to transfer genes from plasmid to/from chromosomes. A resolvase was also present in one of the C3(I1) plasmids. Resolvase keeps plasmids in a monomeric state to reduce plasmid loss during growth (Grindley, 2001). Within these plasmids, some adhesins were also found which is a consistent feature of STECs (Toma *et al*., 2004). Adhesins are an important factor in the colonisation and/or persistence of STECs (Farfan and Torres, 2011).

3.4 In Isolate C3(I1): a Novel Phage-like Plasmid and an IncF hybrid

In isolate C3(I1), two IncF plasmids were predicted by PlasmidFinder (IncFIB and InFIC). However, when the full genome sequence was obtained, including the plasmids, it was shown that both RepFIB and RepA2 were found in the same 144,932 bp plasmid. In addition, by using PHASTER to detect prophage insertions, it was found that the smaller plasmid (48,103 bp) was almost entirely a single phage. Figure 8 below shows the prophage regions present in the phage plasmid and the hybrid IncF plasmid.

Figure 8. Prophage regions identified in the two plasmids from isolate C3(I1): the 48,103 bp phage plasmid (left) and the hybrid IncFIB-IncFIC plasmid (144,932 bp) (right).

This phage-like plasmid encodes no AMR or virulence genes but it does carry the HigA-HigB TA system (Xu *et al*., 2019). It also carries the chromosome partitioning proteins ParA and ParB (Funnell, 2016; Bignell and Thomas, 2001). When this phage-like plasmid's contig is run in NCBI BLAST and on ProphageHunter, it has <10% homology with any other recorded plasmids and phages and so constitutes a novel phage plasmid.

4. Discussion

Shiga toxin-producing strains of *E. coli* (STEC) can cause a range of human diseases. STECs can also carry a range of antimicrobial resistance (AMR) genes which makes combating illnesses an even harder challenge. In this investigation, 6 STEC isolates were isolated from a number of organically-raised cattle and goats in New Territories, Hong Kong. Both herds of animals live and feed in the same areas. These animals have never received medication or other veterinary interventions and forage-feed. The aim of this investigation was to investigate the AMR and virulence genes present in these STEC isolates, on the assumption that, like documented STECs, would be very pathogenic. Although the isolates carry *stx* genes, we find they contain fewer virulence genes than expected, their IncF plasmids do not carry AMR genes and that even the AMR genes they have are not strongly expressed.

The isolates present in this investigation all had chromosome-encoded AMR genes and despite being present in the genomes, each isolate was still susceptible to a large range of antibiotics. For example, our isolates all have the *ampH* resistance gene which binds to penicillins but the isolates themselves are not resistant to ampicillin. We find that the unusual start codon of TTG for this gene means that the gene is not likely to be translated as often and thus, the gene is expressed only at a low level. In addition, the plasmids within each isolate, while being IncF plasmids which are often specialised to transfer resistance genes, mainly contained transposases for transferring genes and TA systems to ensure the survival of the host plasmid.

From these results, we can conclude that while STECs are generally a dangerous strain of *Escherichia coli*, our isolates do not exhibit the same dangers. Not only do the lack of multiple virulence factors render these isolates less dangerous than other reported STECs but the weaker AMR genes and transfer focused plasmids mean that these isolates cannot defend themselves as easily against antibiotics. This is not to say that STECs in general are not dangerous, as our isolates also have the potential to be as such, it is just that our isolates are not as potent as they could be. This is likely from the fact that our isolates are from cows and goats that are grass-fed and raised organically so that they have not needed to be super-virulent.

If plasmids can be seen as the major facilitators of transfer of advantageous genes, then in these organically-raised, grass-fed animals, AMR genes may not offer immediate advantage. Similarly, frequent diarrhoea or rapid intestinal transit times would certainly favour virulence genes associated with adhesion, so that relative abundance may simply reflect relative usefulness.

This research also yielded a novel phage-like plasmid found in a STEC isolated from a cow. This phage-like plasmid had a < 10% homology when compared to all other recorded samples on NCBI BLAST and ProphageHunter. As of the writing of this paper, this phage-like plasmid has yet to be named.

5. Improvement and Further Research

Overall, increasing the number of isolates would result in more accurate, precise data and thus, increase the reliability of our results. In this paper, only one goat isolate was found with *Stx* genes and thus, we had no other goat-derived isolates to compare it with. A further investigation of the full DNA sequences of all of the plasmids present will also help to understand why the IncF plasmids here are unlike those reported, which have always been shown to carry many AMR genes.

Acknowledgements

Many thanks to Dr Simon Griffin, Professor Fred Leung and Ms Grace Lai for their continuous support in this project. Their advice and guidance has been paramount to this project's completion and success. Special thanks to the *Shuyuan* science team for providing equipment and the means to complete this project.

Supplementary Resources

1. Antimicrobial Susceptibility Test Results
   [url: http://tinyurl.com/y4envh9e]
2. G2(I2) IncF Plasmid Map
   [url: http://tinyurl.com/yyxf7nu4]
3. C3(I1) Hybrid Plasmid Map
   [url: http://tinyurl.com/yxfcer84]
4. C3(I1) Novel Phage-like Plasmid Map
   [url: http://tinyurl.com/y54njd2e]

References

Afset, J. E., Bruant, G., Brousseau, R., Harel, J., Anderssen, E., Bevanger, L., Bergh, K. (2006). Identification of Virulence Genes Linked with Diarrhea Due to Atypical Enteropathogenic *Escherichia coli* by DNA Microarray Analysis and PCR. *Journal of Clinical Microbiology*, 44(10), 3703-3711. doi: 10.1128/jcm.00429-06

Alcock, B. P., Raphenya, A. R., Lau, T., Tsang, K. K., Bouchard, M., Edalatmand A., Huynh, W., Nguyen, A. V., Cheng, A. A., Liu, S., Min, S. Y., Miroshnichenko, A., Tran, H. K., Werfalli, R. E., Nasir, J. A., Oloni, M., Speicher, D. J., Florescu, A., Singh, B., Faltyn, M., … McArthur, A. G. (2020). CARD 2020: antibiotic resistome surveillance with the comprehensive antibiotic resistance database. *Nucleic acids research*, 48(D1), D517–D525. doi: 10.1093/nar/gkz935

Altschul, S. F.,Gish, W., Miller, W., Myers, E. W., Lipman, D. J. (1990). Basic lcal alignment search tool, *Journal of Molecular* Biology, 215(3, 403-410. doi: 10.1016/s0022-2836(05)80360-2

Arndt, D., Grant, J. R., Marcu, A., Sajed, T., Pon, A., Liang, Y., Wishart, D. S. (2016) PHASTER: A better, faster version of the PHAST phage search tool. *Nucleic Acids Research*, 44(W1): W16-21 doi: 10.1093/nar/gkw387

Bielaszewska, M., Mellmann, A., Zhang, W., Köck, R., Fruth, A., Bauwens, A., Peters, G., Karch, H. (2011). Characterisation of the *Escherichia coli* strain associated with an outbreak of haemolytic uraemic syndrome in Germany, 2011: a microbiological study*, The Lancet Infectious Diseases*, 11(9), 671-676. doi: 10.1016/S1473-3099(11)70165-7

Bignell, C., Thomas, C. M. (2001). The bacterial ParA-ParB partitioning proteins. *Journal of Biotechnology*, 91(1), 1-34. doi: 10.1016/s0168-1656(01)00293-0

Carattoli, A., Zankari, E., García-Fernández, A., Voldby Larsen, M., Lund, O., Villa, L., Møller Aarestrup, F., Hasman, H. (2014). In silico detection and typing of plasmids using PlasmidFinder and plasmid multilocus sequence typing, *Antimicrobial Agents and Chemotherapy*, 58(7), 3895–3903. doi:.10.1128/AAC.02412-14

Cascales, E., Buchanan, S. K., Duché, D., Kleanthous, C., Lloubès, R., Postle, K., . . . Cavard, D. (2007). Colicin Biology. *Microbiology and Molecular Biology Reviews*, 71(1), 158-229. doi:.10.1128/mmbr.00036-06

Coray, D. S., Kurenbach, B., Heinemann, J. A. (2017). Exploring the parameters of post-segregational killing using heterologous expression of secreted toxin barnase and antitoxin barstar in an *Escherichia coli* case study. *Microbiology*, 163(2), 122-130. doi:.10.1099/mic.0.000395

Fan, R., Shao, K., Yang, X., Bai, X., Fu, S., Sun, H., . . . Xiong, Y. (2019). High prevalence of non-O157 Shiga toxin-producing *Escherichia coli* in beef cattle detected by combining four selective agars. *BMC Microbiology*, 19(1). doi:.10.1186/s12866-019-1582-8

Farfan, M. J., Torres, A. G. (2011). Molecular Mechanisms That Mediate Colonization of Shiga Toxin-Producing *Escherichia* coli *Strains. Infection and Immunity*, 80(3), 903-913. doi:.10.1128/iai.05907-11

Funnell, B. E. (2016). ParB Partition Proteins: Complex Formation and Spreading at Bacterial and Plasmid Centromeres. *Frontiers* in, *Molecular Biosciences*, 3. doi:.10.3389/fmolb.2016.00044

Garcia-Pino, A., Balasubramanian, S., Wyns, L., Gazit, E., Greve, H. D., Magnuson, R. D., . . . Loris, R. (2010). Allostery and Intrinsic Disorder Mediate Transcription Regulation by Conditional Cooperativity. *Cell*, 142(1), 101-111. doi:.10.1016/j.cell.2010.05.039

Grant, M. A., Weagant, S. D., Feng, P. (2001). Glutamate Decarboxylase Genes as a Prescreening Marker for Detection of Pathogenic *Escherichia coli* Groups. *Applied and Environmental Microbiology,* 67(7), 3110-3114. doi:.10.1128/aem.67.7.3110-3114.2001

Grindley, N. (2001). Resolvase. *Encyclopedia of Genetics*, 1687-1688. doi:.10.1006/rwgn.2001.1104

Henderson, T. A., Young, K. D., Denome, S. A., Elf, P. K. (1997). AmpC and AmpH, proteins related to the class C beta-lactamases, bind penicillin and contribute to the normal morphology of *Escherichia coli*. *Journal of Bacteriology*, 179(19), 6112-6121. doi: 10.1128/jb.179.19.6112-6121.1997

Joensen, K. G., Scheutz, F., Lund, O., Hasman, H., Kaas, R. S., Nielsen, E. M., Aarestrup, F. M. (2014) Real-Time Whole-Genome Sequencing for Routine Typing, Surveillance, and Outbreak Detection of Verotoxigenic *Escherichia coli*. *Journal of Clinical Microbiology*, 52(5), 1501-1510. doi: 10.1128/jcm.03617-13

Joensen, K. G., Tetzschner, A. M., Iguchi, A., Aarestrup, F. M., Scheutz, F. (2015). Rapid and Easy In Silico Serotyping of *Escherichia coli* Isolates by Use of Whole-Genome Sequencing Data. *Journal of Clinical Microbiology*, 53(8), 2410-2426. doi: 10.1128/jcm.00008-15

Kanehisa, M., Sato, Y., Morishima, K. (2016). BlastKOALA and GhostKOALA: KEGG tools for functional characterization of genome and metagenome sequences. *Journal of Molecular Biology*, 428(4), 726-731. doi: 10.1016/j.jmb.2015.11.006

Kaper, J. B., Nataro, J. P., Mobley, H. L. (2004). Pathogenic *Escherichia coli*. *Nature Reviews Microbiology*, 2(2), 123-140. doi: 10.1038/nrmicro818

Karpman, D., Loos, S., Tati, R., Arvidsson, I. (2017). Haemolytic uraemic syndrome. *Journal of Internal Medicine*, 281(2), 123–148. doi: 10.1111/joim.12546

Koonin, E. V., Novozhilov, A. S. (2009). Origin and evolution of the genetic code: The universal enigma. *IUBMB Life*, 61(2), 99-111. doi: 10.1002/iub.146

Li, M., Wang, F., Li, F. (2011). Identification and Molecular Characterization of Antimicrobial-Resistant Shiga Toxin–Producing *Escherichia coli* Isolated from Retail Meat Products. *Foodborne Pathogens and Disease*, 8(4), 489-493. doi:.10.1089/fpd.2010.0688

Linhartová, I., Bumba, L., Mašín, J., Basler, M., Osička, R., Kamanová, J., . . . Šebo, P. (2010). RTX proteins: A highly diverse family secreted by a common mechanism. *FEMS Microbiology Reviews,* 34(6), 1076-1112. doi:.10.1111/j.1574-6976.2010.00231.x

Louten, J. (2016). Features of Host Cells. *Essential Human Virology*, 31-48. doi:.10.1016/b978-0-12-800947-5.00003-x

Mathusa, E. C., Chen, Y., Enache, E., Hontz, L. (2010). Non-O157 Shiga Toxin–Producing *Escherichia coli* in Foods. *Journal of Food Protection*, 73(9), 1721-1736. doi:.10.4315/0362-028x-73.9.1721

Mcdonald, L. C. (1997). Vancomycin-Resistant *Enterococci* Outside the Health-Care Setting: Prevalence, Sources, and Public Health Implications. *Emerging Infectious Diseases*, 3(3), 311-317. doi: 10.3201/eid0303.970307

Mekata, H., Iguchi, A., Kawano, K., Kirino, Y., Kobayashi, I., Misawa, N. (2014). Identification of O Serotypes, Genotypes, and Virulotypes of Shiga Toxin–Producing *Escherichia coli* Isolates, Including Non-O157 from Beef Cattle in Japan. *Journal of Food Protection*, 77(8), 1269-1274. doi:.10.4315/0362-028x.jfp-13-506

Melderen, L. V., Bast, M. S. (2009). Bacterial Toxin–Antitoxin Systems: More Than Selfish Entities? *PLoS Genetics*, 5(3). doi: 10.1371/journal.pgen.1000437

Melton-Celsa, A. R. (2014). Shiga Toxin (Stx) Classification, Structure, and Function. *Microbiology Spectrum*, 2(4). doi: 10.1128/microbiolspec.ehec-0024-2013

Mukherjee, S., Mosci, R. E., Anderson, C. M., Snyder, B. A., Collins, J., Rudrik, J. T., Manning, S. D. (2017). Antimicrobial Drug–Resistant Shiga Toxin–Producing *Escherichia coli* Infections, Michigan, USA. *Emerging Infectious Diseases*, 23(9), 1609-1611. doi: 10.3201/eid2309.170523

Muniesa, M., Hammerl, J.A., Hertwig, S., Appel, B., Brüssow, H. (2012). Shiga Toxin-Producing *Escherichia coli* O104:H4: a New Challenge for Microbiology, *Applied and Environmental Microbiology*, 78(12), 4065-4073. doi: 10.1128/AEM.00217-12

Nataro, J. P., Kaper, J. B. (1998). Diarrheagenic *Escherichia coli*. *Clinical Microbiology Reviews,* 11(1), 142-201. doi: 10.1128/cmr.11.1.142

Ogura, T., Hiraga, S. (1983). Mini-F plasmid genes that couple host cell division to plasmid proliferation. *Proceedings of the National Academy of Sciences,* 80(15), 4784-4788. doi:.10.1073/pnas.80.15.4784

Pettengill, E. A., Pettengill, J. B., Binet, R. (2016). Phylogenetic Analyses of *Shigella* and Enteroinvasive *Escherichia coli* for the Identification of Molecular Epidemiological Markers: Whole-Genome Comparative Analysis Does Not Support Distinct Genera Designation. *Frontiers in Microbiology*, 6. doi: 10.3389/fmicb.2015.01573

Rice, L. (2001). Emergence of Vancomycin-Resistant *Enterococci*. *Emerging Infectious Disease*s, 7(2), 183-187. doi: 10.3201/eid0702.010205

Riley, L. W., Remis, R. S., Helgerson, S. D., Mcgee, H. B., Wells, J. G., Davis, B. R., . . . Cohen, M. L. (1983). Hemorrhagic Colitis Associated with a Rare *Escherichia coli* Serotype. *New England Journal of Medicine*, 308(12), 681-685. doi:.10.1056/nejm198303243081203

Rozwandowicz, M., Brouwer, M. S., Fischer, J., Wagenaar, J. A., Gonzalez-Zorn, B., Guerra, B., . . . Hordijk, J. (2018). Plasmids carrying antimicrobial resistance genes in *Enterobacteriaceae*. *Journal of Antimicrobial Chemotherapy*, 73(5), 1121-1137. doi:.10.1093/jac/dkx488

Sandhu, K. S., Gyles, C. L. (2002). Pathogenic Shiga toxin-producing Escherichia coli in the intestine of calves. *Canadian Journal of Veterinary Research = Revue Canadienne De Recherche Veterinaire*, 66(2), 65-72. URL:.www.ncbi.nlm.nih.gov/pmc/articles/PMC226985/

Scheiring, J., Rosales, A., Zimmerhackl, L. B. (2010). Clinical practice. Today's understanding of the haemolytic uraemic syndrome. *European Journal of Pediatrics*, 169(1), 7–13. doi:.10.1007/s00431-009-1039-4

Schlum, K. Jun, S.-R., Udaondo, Z., Ussery, D.W., Emrich, S.J. (2019). Improved bacteria population structure analysis on thousands of genomes using unsupervised methods, *bioRxiv preprint*. doi:.10.1101/599944

The UniProt Consortium (2019). UniProt: a worldwide hub of protein knowledge, *Nucleic Acids Research*, 47: D506-515. doi: 10.1093/nar/gky1049. URL: www.uniprot.org/

Toma, C., Espinosa, E. M., Song, T., Miliwebsky, E., Chinen, I., Iyoda, S., . . . Rivas, M. (2004). Distribution of Putative Adhesins in Different Seropathotypes of Shiga Toxin-Producing *Escherichia coli*. *Journal of Clinical Microbiology*, 42(11), 4937-4946. doi: 10.1128/jcm.42.11.4937-4946.2004

Torres, A. G., Blanco, M., Valenzuela, P., Slater, T. M., Patel, S. D., Dahbi, G., . . . Blanco, J. (2009). Genes Related to Long Polar Fimbriae of Pathogenic *Escherichia coli* Strains as Reliable Markers To Identify Virulent Isolates. *Journal of Clinical Microbiology*, 47(8), 2442-2451. doi:.10.1128/jcm.00566-09

Unterholzner, S. J., Poppenberger, B., Rozhon, W. (2013). Toxin–antitoxin systems. *Mobile Genetic Elements*, 3(5). doi: 10.4161/mge.26219

Veeraraghavan, B., Ragupathi, N. D., Bakthavatchalam, Y., Mathur, P., Pragasam, A., Walia, K., Ohri, V. (2019). Plasmid profiles among some ESKAPE pathogens in a tertiary care centre in south India. *Indian Journal of Medical Research*, 149(2), 222. doi:.10.4103/ijmr.ijmr_2098_17

Villa, L., García-Fernández, A., Fortini, D., Carattoli, A. (2010). Replicon sequence typing of IncF plasmids carrying virulence and resistance determinants. *Journal of Antimicrobial Chemotherapy*, 65(12), 2518-2529. doi: 10.1093/jac/dkq347

Wick, R.R., Judd, L.M., Gorrie, C.L., Holt, K.E. (2017). Unicycler: Resolving bacterial genome assemblies from short and long sequencing reads. *PLOS Computational Biology*, 13(6), e1005595. doi: 10.1371/journal.pcbi.1005595

Xiong, Y., Wang, P., Lan, R., Ye, C., Wang, H., Ren, J., . . . Xu, J. (2012). A Novel Escherichia coli O157:H7 Clone Causing a Major Hemolytic Uremic Syndrome Outbreak in China. *PLoS ONE*, 7(4). doi: 10.1371/journal.pone.0036144

Xu, B., Liu, M., Zhou, K., Geng, Z., Gao, Z., Dong, Y., . . . Liu, Q. (2019). Conformational changes of antitoxin HigA from Escherichia coli str. K-12 upon binding of its cognate toxin HigB reveal a new regulation mechanism in toxin-antitoxin systems. *Biochemical and Biophysical Research Communications*, 514(1), 37-43. doi: 10.1016/j.bbrc.2019.04.061

Bibliography

Chen, L., Wang, L., Yassin, A. K., Zhang, J., Gong, J., Qi, K., . . . Wang, C. (2018). Genetic characterization of extraintestinal Escherichia coli isolates from chicken, cow and swine. *AMB Express*, 8(1). doi: 10.1186/s13568-018-0646-8

Day, M., Doumith, M., Jenkins, C., Dallman, T. J., Hopkins, K. L., Elson, R., . . . Woodford, N. (2016). Antimicrobial resistance in Shiga toxin-producing *Escherichia coli* serogroups O157 and O26 isolated from human cases of diarrhoeal disease in England, 2015. *Journal of Antimicrobial Chemotherapy*, 72(1), 145-152. doi: 10.1093/jac/dkw371

Hecht, A., Glasgow, J., Jaschke, P. R., Bawazer, L. A., Munson, M. S., Cochran, J. R., . . . Salit, M. (2017). Measurements of translation initiation from all 64 codons *in E. coli. Nucleic Acids Research*, 45(7), 3615-3626. doi: 10.1093/nar/gkx070

Johnson, T. J., Wannemuehler, Y. M., Nolan, L. K. (2008). Evolution of the iss Gene in *Escherichia coli. Applied and Environmental Microbiology,* 74(8), 2360-2369. doi:.10.1128/aem.02634-07

Kobayashi, N., Lee, K., Yamazaki, A., Saito, S., Furukawa, I., Kono, T., . . . Hara-Kudo, Y. (2013). Virulence Gene Profiles and Population Genetic Analysis for Exploration of Pathogenic Serogroups of Shiga Toxin-Producing *Escherichia coli. Journal of Clinical Microbiology,* 51(12), 4022-4028. doi: 10.1128/jcm.01598-13

Michino, H., Araki, K., Minami, S., Takaya, S., Sakai, N., Miyazaki, M., . . . Yanagawa, H. (1999). Massive Outbreak of *Escherichia coli* O157: H7 Infection In Schoolchildren in Sakai City, Japan, Associated with Consumption of White Radish Sprouts. *American Journal of Epidemiology*, 150(8), 787-796. doi: 10.1093/oxfordjournals.aje.a010082

Peterson, J. W. (1996). Bacterial Pathogenesis. In S. Baron (Ed.), *Medical Microbiology*. (4th ed.). University of Texas Medical Branch at Galveston.

Phan, M. D., Forde, B. M., Peters, K. M., Sarkar, S., Hancock, S., Stanton-Cook, M., . . . Schembri, M. A. (2015). Molecular Characterization of a Multidrug Resistance IncF Plasmid from the Globally Disseminated *Escherichia coli* ST131 Clone. *PLOS One*, 10(4). doi:.10.1371/journal.pone.0122369

Toshima, H., Yoshimura, A., Arikawa, K., Hidaka, A., Ogasawara, J., Hase, A., . . . Nishikawa, Y. (2007). Enhancement of Shiga Toxin Production in Enterohemorrhagic *Escherichia coli* Serotype O157:H7 by DNase Colicins. *Applied and Environmental Microbiology*, 73(23), 7582-7588. doi: 10.1128/aem.01326-07

Tozzoli, R., Grande, L., Michelacci, V., Ranieri, P., Maugliani, A., Caprioli, A., Morabito, S. (2014). Shiga toxin-converting phages and the emergence of new pathogenic *Escherichia coli*: A world in motion. *Frontiers in Cellular and Infection Microbiology*, 4. doi:.10.3389/fcimb.2014.00080

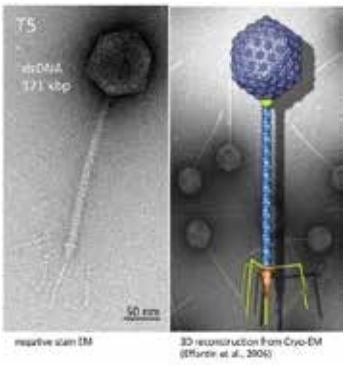
Yamaguchi, Y., Park, J., Inouye, M. (2011). Toxin-Antitoxin Systems in Bacteria and Archaea. *Annual Review of Genetics*, 45(1), 61-79. doi:.10.1146/annurev-genet-110410-132412

Bacteriophages Drive Survival and Virulence of *Escherichia coli* in a Soil Environment

Tian Yu Tatiana Zhang

Abstract

Although *Escherichia coli* is commonly known as a commensal bacterium of the intestines of warm-blooded animals and therefore useful as a sign of faecal contamination, strains of *E. coli* have been found to persist in water and soil. This project uses whole genome sequencing and longitudinal sampling to investigate the role of bacteriophage (viruses that targets bacteria) in the adaptation and survival of enteric strains of *E. coli* in a soil environment.



1. Introduction

Escherichia coli is a rod-shaped, Gram-negative bacterium of the family *Enterobacteriaceae* that is resident in the lower intestine of warm-blooded animals, from which it is typically discharged into the environment through faeces and can be detected in wastewater effluent. As a result, the presence of *E. coli* in environmental waters is typically considered an indicator of recent faecal pollution (Odonkor and Ampofo, 2013) because it is thought to require body temperatures to thrive (Broeze *et al*., 1978)

Nevertheless, numerous recent studies have reported that certain strains of *E. coli* can survive for long periods of time in extraintestinal environments and it is notable that major pathogenic outbreaks have often been associated with fresh vegetables and water at ambient temperatures. This phenomenon raises two main concerns: first, whether *E. coli* can continue to function as a reliable Faecal Indicator Bacterium (FIB).

(Anderson and Harwood, 2005); second, whether the mechanisms that allow *E. coli* to adapt to new environments might also lead to to evolution of the virulent strains that can cause human diseases (Jang *et al*., 2017).

There are a number of ways in which an organism can adapt to a new environment. For example, it could control genes already present or the species could evolve by random mutation and natural selection. However, the accumulation of random mutations is a gradual process that takes many generations and there do seem to be some very different genes present in environmental *E. coli* and intestinal (enteric) strains (Luo *et al*., 2011).

The research for this article was conducted in the *Shuyuan* Molecular Biology Laboratory and posters about this work were prepared for the AGU (American Geophysical Union) Fall Meeting in 2019, and the ASM Microbe Conference in 2020.

Another way in which bacterial cells can adapt is to acquire new genes from others, which is known as *horizontal gene transfer* (HGT). HGT in bacteria can occur by incorporating stray pieces of cell-free DNA from the environment (Sun, 2018); by taking up plasmids (self-replicating collections of genes) from compatible bacteria (Harrison and Brockhurst, 2012); or, by surviving an attack from a bacteriophage (a bacteria-specific virus), which would then leave many of its genes incorporated into the bacterial genome as *prophage* (Mosig and Calendar, 2002) (see Figure. 1). Since the survival of enteric strains in the environment depends on their rapid adaptation, and because bacteriophages are known to be very numerous in water and soil (Clokie *et al*., 2011), a bacteriophage-mediated pathway would probably be the fastest.

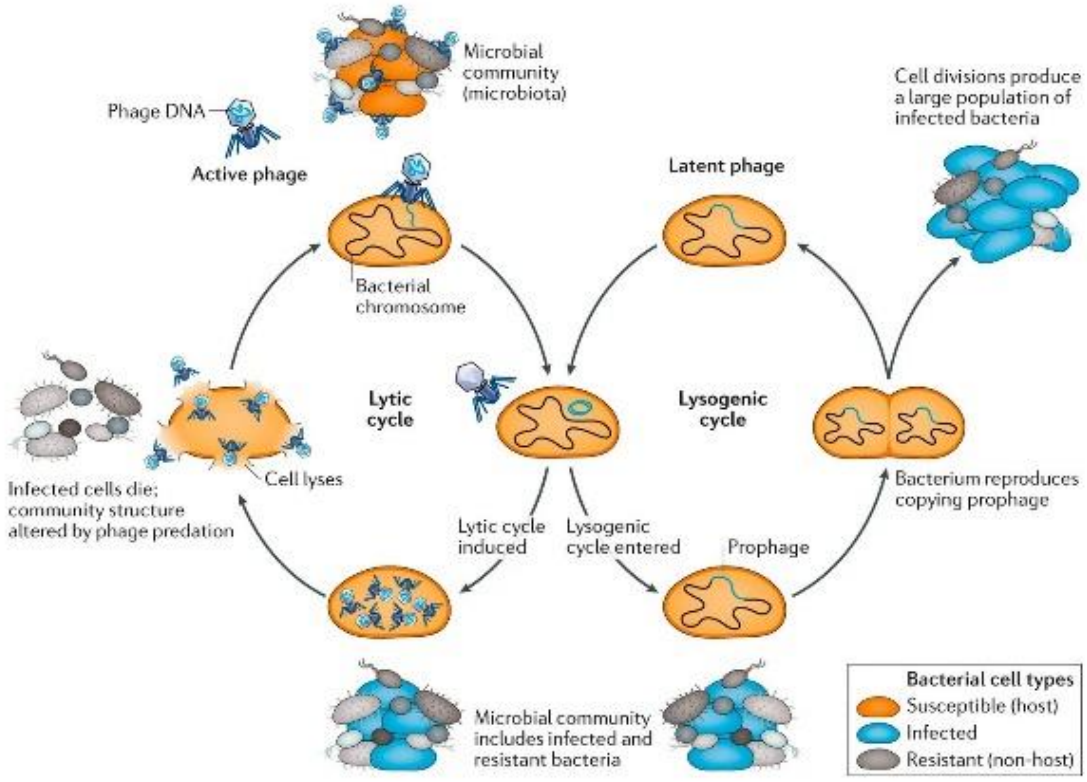


Figure 1. Active phages have two ways of propagation: lytic and lysogenic. In the former, the immediate replication of the virion lyses the bacterial cells. In the latter, they integrate with host DNA and remain dormant as prophages until the conditions of the host deteriorate. At that point, the *prophage* can become active and come back out of the bacterial chromosome, triggering the remaining steps of the lytic cycle [from "Figure. 1: The Lytic and Lysogenic Cycles of Bacteriophages", Garreto *et al*., 2019; p. 425]

Although *E. coli* is typically believed to be a specialist of the gut with poor growth at temperatures below 37 °C (Broeze *et al*., 1978), certain strains can survive and grow in environmental conditions for prolonged periods. This is particularly concerning, because environmentally-adapted *E. coli* strains are indistinguishable from faecally-derived strains based on current water quality testing methods and may not indicate recent faecal contamination at all. Therefore, the presence of these populations severely confounds the use of the species as a FIB. If high concentrations of environmentally-adapted *E. coli* are widespread, then FIB-based methods need to be improved to specifically quantify only *E. coli* (or another species) of purely faecal origin, in order to be a better predictor risks to human health (Jang *et al*., 2017).

The long-term survival and growth of *E. coli* in the environment is also of concern if such strains carry significant numbers of virulence factors or antibiotic resistance genes. Since these genes can be easily transferred between *Enterobacteriaceae*, including the environmental strains of *E. coli*, this raises questions for public health. As the pathogenic agent of

gastrointestinal and water-related diseases, *Escherichia coli* is thought to be responsible for over 73 thousand deaths each year in the United States alone (Lim *et al*., 2010). And, since levels of antimicrobial resistance have seen a steady increase making infections harder to treat, this number may be expected to rise (Aslam *et al*., 2018). As a result, it is imperative to understand the mechanisms behind the adaptation and survival of *Escherichia coli*, one of the most commonly found bacteria on the planet (Ishii and Sadowsky, 2008)

2. Methods

2.1 Soil sample collection and preparation of a 'phage solution'

Five cm^3 samples of soil were collected from the grounds of the ISF Academy and placed in three separate 15 mL centrifuge tubes labeled, respectively, "A", "B", and "C". Tubes A and B were autoclaved to sterilise the soil and to remove bacteria and bacteriophages. The sample of soil in Tube C was used to prepare the bacteriophage-containing solution: 20 mL of sterile phosphate buffered saline (PBS) was added and the mixture vortexed until homogeneous. After allowing larger solids to settle, a portion of the solution was passed through a 0.22 µm syringe filter (which removes microbes, but not the smaller bacteriophages). 10 mL of filtered solution to the autoclaved soil in Tube A, while 10 mL of sterile PBS was added to Tube B. Both tubes were then inoculated with a 10 µL aliquot of a solution of enteric *E. coli* K-12 strain and incubated at 37 °C on an orbital shaker.

Over the course of 10 days (on days 0, 2, 6, and 10), 100 µL samples of supernatant were recovered from tubes A and B. These samples were serial-diluted to give a countable number of colonies on the 3M Petrifilm® plates (which also allowed the relative concentrations of *E. coli* in each tube to be tracked). Two colonies were picked from each 3M plate and initially streaked onto MacConkey agar plates followed by four further passages on LB before DNA extraction.

2.2 DNA Extraction and Sequencing

An Invitrogen™ PureLink™ Genomic DNA Mini Kit was used for genomic DNA extraction. DNA concentration and purity were estimated by A_{260}/A_{230} and A_{260}/A_{280} ratios, measured using a Biodrop µLite spectrophotometer. Following DNA extraction, shotgun sequencing was performed using the Illumina MiSeq system to give draft sequences. Long-read sequences were then obtained using Nanopore MinIon and hybrid assembly of both long- and short-reads by Unicycler generated complete genome sequences.

2.3 Genomic Analysis

Synmap (Haug-Baltzell *et al*., 2017) was used to create synteny maps. Clustal Omega (Madeira *et al*., 2019) and PATRIC (Wattam *et al*., 2016) were used for multiple sequence alignment (MSA) comparisons. PATRIC was also used to find virulence and other specialty genes. Prophage Hunter (Song *et al*., 2019) was used to find prophages and their respective genes. Ori-Finder 2 (Luo *et al*., 2014) was used to find the origin of replication.

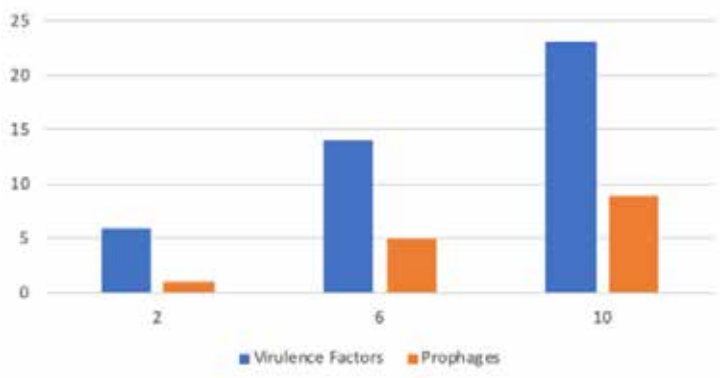


Figure 2. The number of prophages and virulence genes increases steadily in the sample incubated with the phage solution.

3. Results and Discussion

3.1 Correlation between prophages and virulence factors

Results show a steady increase in both virulence factors and prophages over time in the isolate incubated in the presence of phage. Values in Figure 2 were obtained by calculating the change in the number of genes over time, as compared to the day 0 isolate. This was conducted using analysis by PATRIC (Wattam *et al*., 2016) and Prophage Hunter (Song *et al*., 2019).

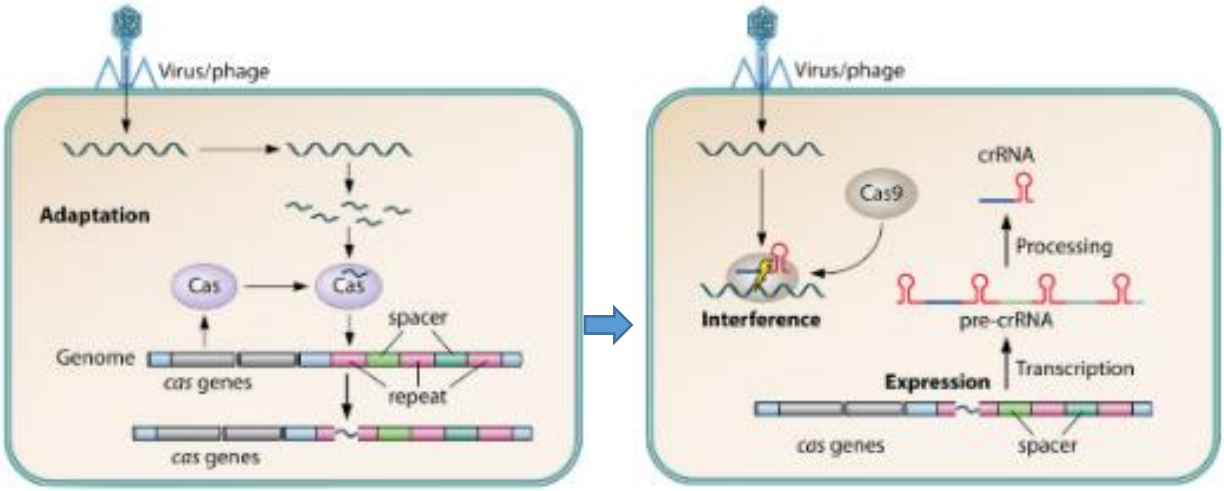


Figure 3. The CRISPR-Cas system is a prokaryotic defence mechanism against invading mobile genetic elements (MGE). In the bacterial genome, the CRISPR array is composed of alternating identical repeats and unique spacers. These spacer sequences match plasmid and phage genomes. CRISPR-Cas serves as an archive of previous infections. Through acquisition of phage (or other MGE) derived spacers, the bacterium immunises against future infection (Barrangou *et al*., 2007). There are three stages of immunity: adaptation or spacer acquisition, crRNA biogenesis, and target interference. During adaptation, the DNAs of foreign elements are integrated into the CRISPR array to provide memory of infection. Memory is retrieved when the locus is transcribed to produce pre-crRNA, a precursor of crRNA (CRISPR RNA) (Hille *et al.,* 2018). The pre-crRNA is soon processed into mature crRNA, smaller units corresponding to a single spacer flanked by a partial repeat. Upon invasion of MGEs, the crRNA, bound to Cas protein(s), detects complementary sequences, protospacers, in the foreign nucleic acids and cleaves them (Rath *et al*., 2015) [*from* "Fig. 4: Process of CRISPR-Cas acquired immune system", Ishino *et al*., 2018; p. 50]

The CRISPR-Cas system provides immunity and natural defense to the bacterium against foreign invading MGEs, in this case, bacteriophage (see Figure. 3). Analysis in PATRIC reveals four additional pairs of CRISPR spacers in the day 10 isolate with phage. Such spacer acquisition is indicative of new phage-derived DNA and evidence of increased bacteria-phage interactions (Figure 4).

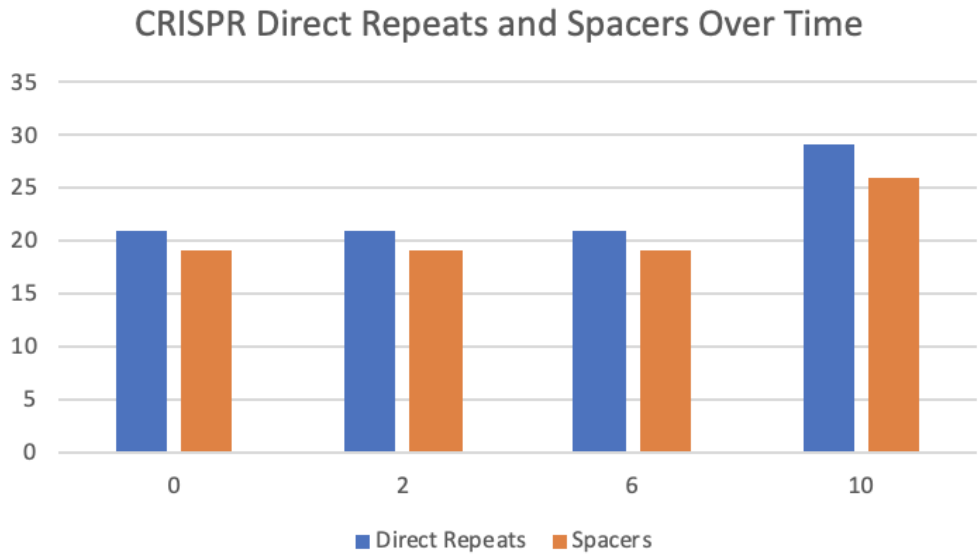


Figure 4. The number of CRISPR repeats and spacers shows a significant increase by day 10.

3.2 PATRIC Genome Alignment and Annotation

Isolates incubated in the presence of phage

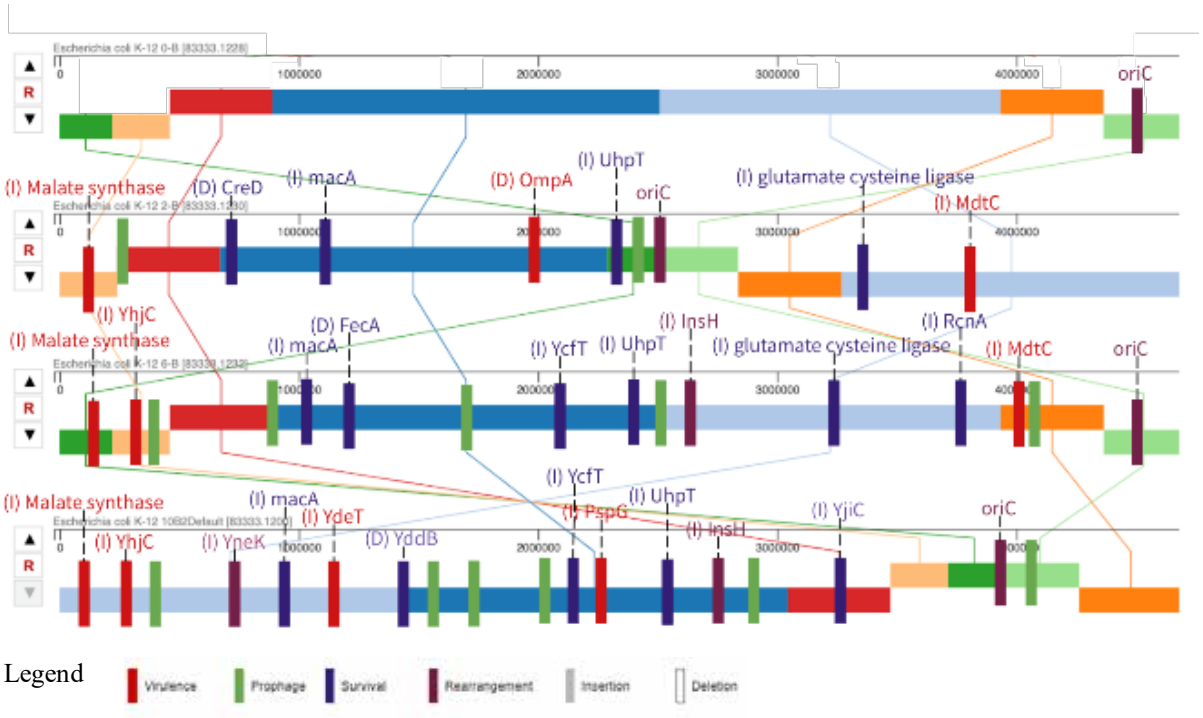


Figure 5. The rows show results for days 0, 2, 6, and 10, respectively. Around 25% more molecular events occur in the presence of phages.

PATRIC was also used to perform multiple sequence alignment, revealing areas of inversion within the whole genome and showing the places where key virulence and survival genes were exchanged. Isolates incubated in the presence of phages show approximately 25% more gene insertions, duplications, and deletions; in other words, there is considerably more gene transfer activity in the four experimental isolates than in the four control isolates. The bacterial genome in the presence of phage appears to grow more dynamic and hence more proactive in naturalisation, since the number of gene insertions only increases over time. In addition, most acquired genes in day 6 are related to either survival or adaption (as indicated in Figure. 5 by the large portion of purple tags). On the other hand, most changes in day 10 are related to virulence (as indicated by the large portion of red tags). In both cases, these new genes were found proximal to prophage insertions (marked by the green tags).

3.3 Large-scale chromosomal rearrangements

Large-scale chromosomal rearrangements also seem to play a significant role in helping this enteric strain naturalise to the soil environment. This is seen from the dramatic reshuffling of genomic regions in the day 10 isolate. Moreover, the position of the origin of replication (*oriC*) changes in day 10, signalling a large-scale genomic inversion, and further suggesting that as a means of adaptation. Genomic inversions refer to a chromosome rearrangement in which one or more segments change position relative to each and may become reversed end to end – for example, ABCDEFGH may become ABCFEDGH.

Isolates incubated without additional phage

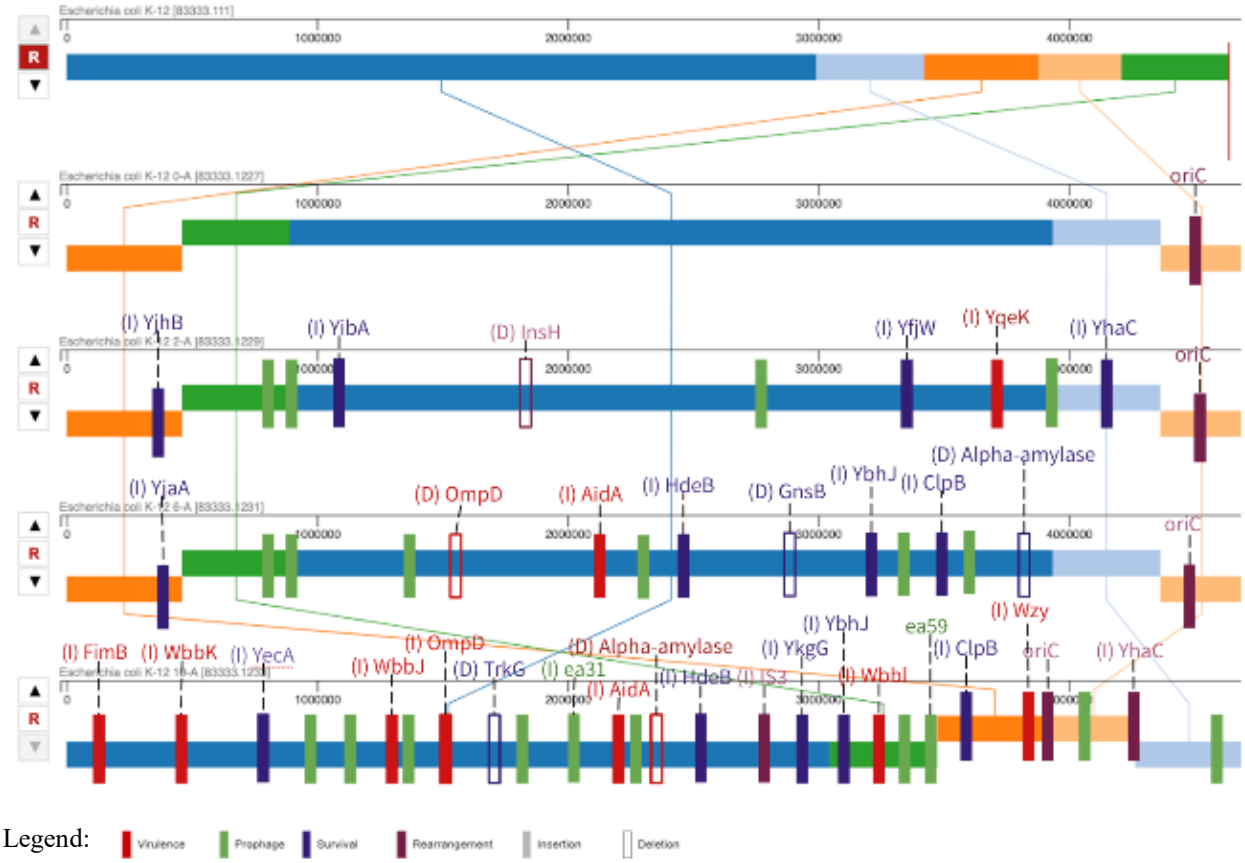


Figure 6. Some significant changes – including prophage insertions, gene duplication, as well as large-scale rearrangements – can also be seen in the control.

It was surprising to note many changes could also be seen in the control sample, incubated without additional phage, even though this was to a lesser extent than the phage-treated sample. These included prophage insertions, and gene duplications and deletions. For instance, the gene *PspG*, a phage shock protein, is believed to manage membrane stress in all *Enterobacteriaceae* and has recently emerged as being important for virulence in several pathogenic species of this phylum (Wallrodt *et al*., 2014). While limited activity might be expected in the control, gene exchanges like this one suggest that the bacterium on its own, without the aid of additional phage, can still adapt to a foreign environment, although with less success. It is nevertheless important to remember that some of the changes seen may be explained by the presence of pre-existing prophage in the isolate that may be re-activated.

3.4 -scale chromosomal rearrangements

Large-scale genomic rearrangements were also seen in the control (see Figure. 6). This is most significant in day 2 and day 10 isolates. On day 2, genome recombination occurs, as demonstrated by the relocation of the *oriC* and the reshuffling of the different sections (more specifically, the dark green, orange, light blue, and light green). On day 10, both experimental and control isolates showed similar genome rearrangements. This suggests that genome inversion is not something unique to isolates with phage, but could be a common form of adaptation in *E. coli*.

Although the large-scale rearrangements seen in both day 10 isolates (*i.e.* with additional phage and the control) are identical in where they occur in the genome, the positions of individual gene insertions are not the same in the control and treated samples. The inference here is that the large-scale rearrangement is sequence-dependent so that particular positions in the

genome are recognised. On the other hand, the fact that many new genes appear in different positions in the treated and untreated isolates indicates that this phenomenon is not necessarily guided by specific sequences.

4. Discussion

The results above further suggest that the presence of bacteriophages contributes extensively to the survival and pathogenicity of their bacterial hosts. The strong correlation between the number of prophage insertions and virulence genes supports the idea that these virulence genes are transferred between bacteria or induced in some way via transduction and/or other phage-mediated mechanisms. Similarly, phage-encoded genes frequently undergo replication and transcriptional activation following prophage induction (Wagner and Waldor, 2002).

4.1 Key genes that change in copy number

Certain genes are particularly significant due to the prominent role they play in either driving virulence or survival (see Table 1). *amyA*, α-Amylase, is an enzyme that catalyses the hydrolysis of the internal α-1,4-glycosidic linkages in starch, sucrose and other alpha-glucosides. It can convert starch into glucose, maltose, and maltotriose units (Rajagopalan and Krishnan, 2008; Gupta *et al*., 2003; Kandra, 2003). Glucose is the key to fuel bacterial growth and division, providing a constant supply of a readily metabolized sugar. Jones *et al*. (2008) suggested that glycogen and maltose utilisation provides a competitive advantage to *Escherichia coli O157:H7*, as these polymers play an important role in colonisation and a ready source of energy.

The decrease in the *PilT* gene suggests a decrease in the need for motility, since *PilT* is cited for being responsible for twitching motility in *Pseudomonas aeruginosa* and social gliding motility in *Myxococcus xanthus* (Okamoto and Ohmori, 2002). *TrkG* encodes sodium transporters and it is involved in cation transmembrane transporter activity. This gene's disappearance suggests that osmotolerance offers no advantage in this specific environment (Schlösser *et al*., 1991). Similarly, *HdeB*, once required for acid stress protection, decreases to 0 at day 2, showing that the environment in which these isolates live in, pH wise, is not unusual for any *E. coli* bacterium (Kern *et al*., 2007)

		Number of copies present						
Gene	Function	D0	D2-B control	D2-A	D6-B control	D6-A	D10-B control	D10-A
AidA-I	Potent bacterial adhesin	2	1	1	1	0	0	7
amyA	Alpha-amylase; Hydrolysis of internal α-1,4-glycosidic linkages in starch (Mikawirawnz, 2016)	2	2	4	1	3	1	2
FimB	Type 1 fimbriae regulatory protein	1	1	1	1	1	1	2
HdeB	Acid stress chaperone	1	1	0	1	0	1	0
waaL	O-antigen ligase; Attachment of a variety of poly-saccharides to core lipid A (Abeyrathne et al., 2005)	1	1	1	2	2	2	3
PilT	Twitching motility	1	0	0	0	0	0	0
TrkG	Potassium uptake	2	1	1	1	1	1	1

Legend: Virulence Survival Both

Table 1. Some interesting genes and their functions, showing significant divergence in copy number between the A (treated) isolates and control (B).

In *Enterobacteriaceae*, the O-antigen ligase has been shown to be responsible for attachment of a variety of polysaccharides to core lipid A. This is an important step in the lipopolysaccharide (LPS) biosynthetic pathway, which contributes extensively to bacterial pathogenicity (Han *et al*., 2012) The biological activity of endotoxins is associated with lipopolysaccharides (LPS), where their toxicity is related to the lipid component (Lipid A) The cell wall antigens (O antigens) of *Escherichia* coli are components of LPS, which may be a part of the pathology of most Gram-negative bacterial infections (Wang and Quinn, 2010). *FimB*, a type 1 fimbriae regulatory protein, is also raised in the day-10 isolate. According to Connell *et al*. (1996), type 1 fimbrial expression enhances *Escherichia coli* virulence by promoting bacterial persistence and increasing the inflammatory response to infection (Connell *et al*., 1996). On the other hand, *AidA-I* is a potent bacterial adhesin that mediates bacterial attachment to a broad variety of human and other mammalian cells. *AidA-I* possesses additional virulence properties, as it is capable of mediating bacterial autoaggregation via intercellular self-recognition and it is a highly efficient initiator of biofilm formation (Sherlock *et al*., 2004).

4.2 The role of the CRISPR-cas system

Recent findings about the CRISPR–cas system have explored a variety of functions beyond adaptive immunity (Koonin and Makarova, 2013; Westra *et al.*, 2014), where an increase in the number of CRISPR

repeats might not only be indicative of accumulated bacteria-phage interactions but also of increased bacterial pathogenicity. Indeed, a recent study indicates that CRISPR–Cas regulates endogenous genes involved in driving the virulence of the pathogen *Francisella novicida* (Sashital *et al*., 2012; Jiang and Marraffini, 2015). Similarly, Louwen and Oost (2014) suggest that the control of gene expression by CRISPR-Cas may enhance bacterial pathogenicity by promoting host colonization. Thus there is a role of CRISPR-Cas in processes other than defence, more specifically, pathogenesis.

4.3 The significance of genome rearrangement

Hotspots for recombination are typically found near the origin of replication, thus its relocation is very significant (Mackiewickz *et al*., 2001). Bi and Liu (1996) observed that DNA rearrangements in prokaryotes are typically mediated by inverted repeats (*i.e. wxyz…….zyxw*). Two points of inversion in the D10-A isolate (day 10 isolate incubated in the presence of phage) (marked by the red circles) were found to be flanked by a unique pair of inverted repeats:

AGCTCGCCCATA …… ATACCCGCTCGA

(3,476,630-3,478,176) (4,259,756-4,261,302)

Genome arrangement likely impacts gene expression. For example, genes placed in proximity to the origin of replication (*oriC*) are systematically expressed at higher levels (Miller and Simons, 1993). This could be due to a transient increase in copy number during growth, which creates a huge difference when the organism is adapting to a harsh or foreign environment. For instance, O-antigen polymerase *wzy* is shown to become proximal to *oriC* in D10-A, signalling the significance of this virulence gene, which is involved in the synthesis of the lipopolysaccharide (LPS) that is a key component of the bacterial outer membrane (Zuo *et al*., 2019). While the *wzy* gene and the lipopolysaccharide are linked to virulence because the cells stick more easily to an infected host, it is also possible that an altered or enhanced outer membrane could be useful in the harsher soil environment.

In a broader scope, Sousa *et al*. (1997) propose chromosomal positioning as a mechanism of adaptation to changing environments because it regulates the patterns of gene expression to an optimal balance of activity. This is consistent with the research by Merrikh and Merrikh (2018), which finds that chromosomal structure changes in prokaryotes are associated with rapid naturalisation and even pathogenicity. In addition, Cui *et al*. (2012), working with *Staphylococcus* aureus, found that the bacterium generates a reversible, large-scale inversion of its chromosome (about half of its total genome) at high frequencies of up to once every four generations, which rapidly produces two distinct phenotypes within a population. Since the relative success of each phenotype depends on the prevailing conditions, they suggest that the reversible switching is a 'bet-hedging' strategy to ensure that a suitably-adapted phenotype is always available. Given that gut *E. coli* in the intestine is almost certainly destined to become faeces in the external environment, it would not be surprising if the bacterium employs a similar strategy.

It is interesting to note that the reverse complement of one inverted repeat sequence was found in the day 0, 2, and 6 isolates of both phage-treated and control samples. What prompted the bacterium to change its order of its genome in the D10-A isolate, however, is particularly significant and should be investigated in further detail. Because this process occurred equally in treated and untreated isolates, the signals inducing the inversion and the mechanism are probably not phage dependent.

5. Conclusion

Isolates treated with a soil-phage preparation show more insertions, deletions and rearrangements than the control as a way of survival. Virulence and survival genes are clearly associated as part of environmental adaptation, since they often appear right next to newly emerged prophages. In addition, large-scale genomic rearrangements were also found to help this enteric strain naturalise to the soil environment. What was unexpected, however, is that the identical rearrangement takes place in both treated and untreated isolates. To what extent do genomic inversions, as well as insertions, duplications and deletions occur in the absence of additional phages; is there a role for pre-existing prophages; is there any other signal from the soil that stimulates a reaction in the isolates?

Despite these outstanding questions, it is clear that the *Escherichia coli* has a very dynamic genome that responds rapidly to adapt to the unfamiliar soil environment.

6. Future Research and Improvements

Further work will help to confirm the role endogenous phages play in genome inversion and gene insertions/deletions. The results above will also need to be demonstrated to be repeatable. It would also be interesting to investigate the signals and mechanisms

prompting these responses – for example, would there be similar results when an enteric *E. coli* is transferred to other environments, such as seawater or is subject to a change in pH or an extreme temperature?

In addition, the results have unexpectedly demonstrated that genome inversions in *E. coli* may be quite common. Further investigation could help find out just how frequent inversions are, the rate at which they occur, and their diversity in terms of position and the range of sizes of the inverted segments.

References

Abeyrathne, P. D., Daniels, C., Poon, K. K., Matewish, M. J., Lam, J. S. (2005). Functional Characterization of WaaL, a Ligase Associated with Linking O-Antigen Polysaccharide to the Core of *Pseudomonas aeruginosa* Lipopolysaccharide. *Journal of Bacteriology*, 187(9), 3002-3012. doi: 10.1128/jb.187.9.3002-3012.2005

Anderson, K. L., Whitlock, J. E., Harwood, V. J. (2005). Persistence and Differential Survival of Fecal Indicator Bacteria in Subtropical Waters and Sediments. *Applied and Environmental Microbiology*, 71(6), 3041-3048. doi: 10.1128/aem.71.6.3041-3048.2005

Aslam, B., Wang, W., Arshad, M. I., Khurshid, M., Muzammil, S., Rasool, M. H., . . . Baloch, Z. (2018). Antibiotic resistance: A rundown of a global crisis. *Infection and Drug Resistance*, 11, 1645-1658. doi: 10.2147/idr.s173867

Barrangou, R., Fremaux, C., Deveau, H., Richards, M., Boyaval, P., Moineau, S., . . . Horvath, P. (2007). CRISPR Provides Acquired Resistance Against Viruses in Prokaryotes. *Science*, 315(5819), 1709-1712. doi: 10.1126/science.1138140

Bi, X., Liu, L. F. (1996). DNA rearrangement mediated by inverted repeats. *Proceedings of the National Academy of Sciences*, 93(2), 819-823. doi: 10.1073/pnas.93.2.819

Broeze, R. J., Solomon, C. J., Pope, D. H. (1978). Effects of low temperature on in vivo and in vitro protein synthesis in *Escherichia coli* and *Pseudomonas fluorescens*. *Journal of Bacteriology,* 134(3), 861-874. doi: 10.1128/jb.134.3.861-874.1978

Clokie, M. R., Millard, A. D., Letarov, A. V., Heaphy, S. (2011). Phages in nature. *Bacteriophage*, 1(1), 31-45. doi:10.4161/bact.1.1.14942

Connell, I., Agace, W., Klemm, P., Schembri, M., Marild, S., Svanborg, C. (1996). Type 1 fimbrial expression enhances Escherichia coli virulence for the urinary tract. *Proceedings of the National Academy of Sciences*, 93(18), 9827-9832. doi:10.1073/pnas.93.18.9827

Cui, L., Neoh, H., Iwamoto, A., & Hiramatsu, K. (2012). Coordinated phenotype switching with large-scale chromosome flip-flop inversion observed in bacteria. *Proceedings of the National Academy of Sciences*, 109(25). doi:10.1073/pnas.1204307109

Effantin, G., Boulanger, P., Neumann, E., Letellier, L., Conway, J. F. (2006). Bacteriophage T5 structure reveals similarities with HK97 and T4 suggesting evolutionary relationships. Journal of Molecular Biology, 361(5), 993–1002. doi: 10.1016/j.jmb.2006.06.081

Garretto, A., Miller-Ensminger, T., Wolfe, A. J., Putonti, C. (2019). Bacteriophages of the lower urinary tract. Nature Reviews. Urology, 16(7), 422–432. doi: 10.1038/s41585-019-0192-4

Gupta, R., Gigras, P., Mohapatra, H., Goswami, V. K., Chauhan, B. (2003). Microbial α-amylases: A biotechnological perspective. *Process Biochemistry*, 38(11), 1599-1616. doi:10.1016/s0032-9592(03)00053-0

Han, W., Wu, B., Li, L., Zhao, G., Woodward, R., Pettit, N., . . . Wang, P. G. (2011). Defining Function of Lipopolysaccharide O-antigen Ligase WaaL Using Chemoenzymatically Synthesized Substrates. *Journal of Biological Chemistry*, 287(8), 5357-5365. doi:10.1074/jbc.m111.308486

Harrison, E., Brockhurst, M. A. (2012). Plasmid-mediated horizontal gene transfer is a coevolutionary process. *Trends in Microbiology*, 20(6), 262-267. doi: 10.1016/j.tim.2012.04.003

Haug-Baltzell, A., Stephens, S. A., Davey, S., Scheidegger, C. E., Lyons, E. (2017). SynMap2 and SynMap3D: Web-based whole-genome synteny browsers. *Bioinformatics*, 33(14), 2197-2198. doi:10.1093/bioinformatics/btx144

Hille, F., Richter, H., Wong, S. P., Bratovič, M., Ressel, S., Charpentier, E. (2018). The Biology of CRISPR-Cas: Backward and Forward. *Cell*, 172(6), 1239-1259. doi:10.1016/j.cell.2017.11.032

Ishino, Y., Krupovic, M., Forterre, P. (2018). History of CRISPR-Cas from Encounter with a Mysterious Repeated Sequence to Genome Editing Technology. *Journal of Bacteriology*, 200(7). doi: 10.1128/jb.00580-17

Ishii, S., Sadowsky, M. J. (2008). *Escherichia coli* in the Environment: Implications for Water Quality and Human Health. *Microbes and Environments*, 23(2), 101-108. doi: 10.1264/jsme2.23.101

Jang, J., Hur, H. G., Sadowsky, M. J., Byappanahalli, M. N. (2017). Environmental *Escherichia coli:* Ecology and public health implications—a review. *Journal of Applied Microbiology*, 123, 570-581. doi: 10.1111/jam.13468

Jiang, W., Marraffini, L. A. (2015) CRISPR-Cas: New Tools for Genetic Manipulations from Bacterial Immunity Systems. *Annual Review of Microbiology,* 69(1), 209-228. doi: 10.1146/annurev-micro-091014-104441

Jones, S. A., Jorgensen, M., Chowdhury, F. Z., Rodgers, R., Hartline, J., Leatham, M. P., Struve, C., Krogfelt, K. A., Cohen, P. S., Conway, T. (2008). Glycogen and Maltose Utilization by *Escherichia coli* O157:H7 in the Mouse Intestine. *Infection and Immunity*, 76(6), 2531-2540. doi: 10.1128/iai.00096-08

Kandra, L. (2003). α-Amylases of medical and industrial importance. Journal of Molecular Structure: *THEOCHEM*, 666-667, 487-498. doi: 10.1016/j.theochem.2003.08.073

Kern, R., Malki, A., Abdallah, J., Tagourti, J., Richarme, G. (2006). *Escherichia coli* HdeB Is an Acid Stress Chaperone. *Journal of Bacteriology*, 189(2), 603-610. doi: 10.1128/jb.01522-06

Koonin, E., Makarova, K. (2013). The basic building blocks and evolution of CRISPR–Cas systems. *Biochemical Society Transactions*, 41(6), 1392-1400. doi: 10.1042/bst20130038

Lim, J. Y., Yoon, J., Hovde, C. J. (2010). A Brief Overview of *Escherichia coli* O157:H7 and Its Plasmid O157. *Journal of Microbiology and Biotechnology*, 20(1), 5-14. doi:10.4014/jmb.0908.08007

Louwen, R., Oost, J. V. (2014). The Role of CRISPR-Cas Systems in Virulence of Pathogenic Bacteria. *Microbiology and Molecular Biology Reviews*, 78(1), 74-88. doi: 10.1128/mmbr.00039-13

Luo, C., Walk, S. T., Gordon, D. M., Feldgarden, M., Tiedje, J. M., Konstantinidis, K. T. (2011). Genome sequencing of environmental *Escherichia coli* expands understanding of the ecology and speciation of the model bacterial species. *Proceedings of the National Academy of Sciences*, 108(17), 7200-7205. doi: 10.1073/pnas.1015622108

Luo, H., Zhang, C., Gao, F. (2014). Ori-Finder 2, an integrated tool to predict replication origins in the archaeal genomes. *Frontiers in Microbiology*, 5: 482. doi: 10.3389/fmicb.2014.00482

Mackiewicz, P., Mackiewicz, D., Kowalczuk, M., Cebrat, S. (2001). Flip-flop around the origin and terminus of replication in prokaryotic genomes. *Genome Biology*, 2, 1-4. doi:10.1186/gb-2001-2-12-interactions1004

Madeira, F., Park, Y. M., Lee, J., Buso, N., Gur, T., Madhusoodanan, N., Basutkar, P., Tivey, A., Potter, S. C., Finn, R. D., Lopez, R. (2019). The EMBL-EBI search and sequence analysis tools APIs in 2019. *Nucleic Acids Research*, 47(W1), W636–W641. doi:10.1093/nar/gkz268

Merrikh, C. N., & Merrikh, H. (2018). Gene inversion potentiates bacterial evolvability and virulence. *Nature Communications*, 9(1). doi:10.1038/s41467-018-07110-3

Mikawlrawng, K. (2016). Aspergillus in Biomedical Research. In Gupta, V.K. (Ed.) *New and Future Developments in Microbial Biotechnology and Bioengineering: Aspergillus System Properties and Applications*, Elsevier B.V. (Delhi); Ch. 18, pp. 229-242. doi:10.1016/B978-0-444-63505-1.00019-1

Miller, W. G., & Simons, R. W. (1993). Chromosomal supercoiling in *Escherichia coli*. *Molecular Microbiology*, 10(3), 675-684. doi:10.1111/j.1365-2958.1993.tb00939.x

Mosig, G., Calendar, R. (2002). Horizontal Gene Transfer in Bacteriophages. In Syvanen, M., Kado, C.I. (eds.*), Horizontal Gene Transfer* (2nd. Edn.), Academic Press; Ch. 13, pp. 141-146, VII-VIII. doi: 10.1016/b978-012680126-2/50017-7

Odonkor, S. T., Ampofo, J. K. (2013). Escherichia coli as an indicator of bacteriological quality of water: An overview. *Microbiology Research*, 4(1), 2. doi: 10.4081/mr.2013.e2

Okamoto, S., Ohmori, M. (2002). The Cyanobacterial PilT Protein Responsible for Cell Motility and Transformation Hydrolyzes ATP. *Plant and Cell Physiology*, 43(10), 1127-1136. doi:10.1093/pcp/pcf128

Rajagopalan, G., Krishnan, C. (2008). α-Amylase production from catabolite derepressed *Bacillus subtilis* KCC103 utilizing sugarcane bagasse hydrolysate. *Bioresource Technology*, 99(8), 3044-3050. doi: 10.1016/j.biortech.2007.06.001

Rath, D., Amlinger, L., Rath, A., Lundgren, M. (2015). The CRISPR-Cas immune system: Biology, mechanisms and applications. *Biochimie*, 117, 119-128. doi:10.1016/j.biochi.2015.03.025

Sashital, D., Wiedenheft, B., Doudna, J. (2012). Mechanism of Foreign DNA Selection in a Bacterial Adaptive Immune System. *Molecular Cell*, 46(5), 606-615. doi: 10.1016/j.molcel.2012.03.020

Schlösser, A., Kluttig, S., Hamann, A., Bakker, E. P. (1991). Subcloning, nucleotide sequence, and expression of *trkG*, a gene that encodes an integral membrane protein involved in potassium uptake via the Trk system *of Escherichia coli. Journal of Bacteriology*, 173(10), 3170-3176. doi: 10.1128/jb.173.10.3170-3176.1991

Sherlock, O., Schembri, M. A., Reisner, A., Klemm, P. (2004). Novel Roles for the AIDA Adhesin from Diarrheagenic *Escherichia coli*: Cell Aggregation and Biofilm Formation. *Journal of Bacteriology*, 186(23), 8058-8065. doi: 10.1128/jb.186.23.8058-8065.2004

Song, W., Sun, H., Zhang, C., Cheng, L., Peng, Y., Deng, Z., . . . Xiao, M. (2019). Prophage Hunter: An integrative hunting tool for active prophages. *Nucleic Acids Research*, 47(W1), W74–W80. doi: 10.1093/nar/gkz380

Sousa, C., Lorenzo, V. D., & Cebolla, A. (1997). Modulation of gene expression through chromosomal positioning in *Escherichia coli*. *Microbiology*, 143(6), 2071-2078. doi:10.1099/00221287-143-6-2071

Sun, D. (2018). Pull in and Push Out: Mechanisms of Horizontal Gene Transfer in Bacteria. *Frontiers in Microbiology*, 9: 2154. doi: 10.3389/fmicb.2018.02154

Wagner, P. L., Waldor, M. K. (2002). Bacteriophage Control of Bacterial Virulence. *Infection and Immunity*, 70(8), 3985-3993. doi: 10.1128/iai.70.8.3985-3993.2002

Wallrodt, I., Jelsbak, L., Thomsen, L. E., Brix, L., Lemire, S., Gautier, L., . . . Olsen, J. E. (2014). Removal of the phage-shock protein PspB causes reduction of virulence in serovar *Typhimurium* independently of NRAMP1. *Journal of Medical Microbiology*, 63(6), 788-795. doi:10.1099/jmm.0.072223-0

Wang, X., Quinn, P. J. (2010). Endotoxins: Lipopolysaccharides of Gram-Negative Bacteria. *Subcellular Biochemistry*, 53, 3-25. doi:10.1007/978-90-481-9078-2_1

Wattam, A. R., Davis, J. J., Assaf, R., Boisvert, S., Brettin, T., Bun, C., . . . Stevens, R. L. (2016). Improvements to PATRIC, the all-bacterial Bioinformatics Database and Analysis Resource Center. *Nucleic Acids Research*, 45(D1): D535-D542. doi:10.1093/nar/gkw1017

Westra, E. R., Buckling, A., Fineran, P. C. (2014). CRISPR–Cas systems: Beyond adaptive immunity. *Nature Reviews Microbiology*, 12(5), 317-326. doi: 10.1038/nrmicro3241

Zuo, J., Tu, C., Wang, Y., Qi, K., Hu, J., Wang, Z., Mi, R., Yan Huang, Chen, Z., Han, X. (2019). The role of the wzy gene in lipopolysaccharide biosynthesis and pathogenesis of avian pathogenic *Escherichia coli*. *Microbial Pathogenesis*, 127, 296–303. doi:10.1016/j.micpath.2018.12.021

To what extent can Mathematics be used to Evaluate and Change the Degree of Musical Harmony of an Equal Tempered Music Scale?

Yifan Tony Tu

1. Introduction

Despite a lot of discussion about musical scales in western music and how their harmony compares, there isn't much mathematical literature that has been done on the topic of the degree of musical harmony of the equal tempered music scale. The reason that 12-note, instead of 10-note or 20-note, equal temperament scale is the most widely used scale in modern music, might be justifiable through mathematical analysis of musical harmony. This analysis thus intend to explore a new way of calculating the degree of musical harmony of an equal temperament music scale, which might help justify the use of 12-note-equal temperament scale, thus the research question:

"To what extent can mathematics be used to evaluate and change the degree of musical harmony of an equal tempered music scale?"

1. This exploration can be broken further into three stages:
2. How to mathematically compare harmony level between notes
3. How to mathematically calculate harmony level of a scale
4. How to mathematically evaluate the degree of harmony of different scales

There are some music background knowledge that is important for this discussion:

Music notes can be defined by a frequency. For example, a note can be defined by stating that it has frequency of 432 Hz. A scale is a set of multiple different notes.

Widely attributed to Pythagoras, the discovery that music notes "repeat" is also important to my discussion. According to Durfee and Colton, two notes with a frequency ratio of 2:1 (for example, 440 Hz and 880 Hz) sound very similar and are therefore to used to define an octave. (835) So we believe that when the frequency is doubled, it is the same note, but on a higher octave. This is the foundation of constructing any scale.

According to them, another important property of sound is that when two notes with frequency ratio of 3:2. For example, when 440 Hz and 660 Hz are played together, they sound nice, or harmonious in musical term. (835)

The important assumption is that two note will sound harmonious when their ratio difference can be expressed using simple fractions like $\frac{3}{2}$ or $\frac{5}{4}$.

Pythagorean tuning is one of the most ancient and well known ways to generate a music scale which is the basis for the development of equal temperament scale. Pythagoras utilised the fact that when two notes of different frequency are played together, if the two frequencies differ by the ratio $\frac{3}{2}$, they will sound very pleasing to human. Based on this fact, to construct a scale starting with base frequency f, the following notes should have frequencies of

$(\frac{3}{2})^2 f$, $(\frac{3}{2})^2 f = \frac{9}{4} f$, $(\frac{3}{2})^3 f = \frac{27}{8} f$, *etc.*

Combining with the idea of octave (notes with frequency ratio of 1:2 sounds very similar), the notes can be then be put in-between the ratio of 1:1 to 1:2. For example, $\frac{9}{4}$ x $\frac{1}{2} = \frac{9}{8}$, is the frequency of a note within the ratio range of 1 and 2. Similarly, $\frac{27}{8}$ becomes $\frac{27}{16}$.

Another way to generate harmonious note base on the base frequency is to divide the ratio $\frac{3}{2}$. This can also be considered as multiply by $\frac{2}{3}$, when placed in-between 1 and 2 by multiplying the ratio by 2, the ratio is 43. This ratio of $\frac{4}{3}$ is therefore a new note.

This article was written in partial fulfillment of the Extended Essay for the IB Diploma, May 2019.

However, there are two problems to the harmonious ratio of $\frac{3}{2}$, one being the Pythagorean comma according to Durfee and Colton (837). This is seen as multiplying $\frac{3}{2}$ six times result in $(\frac{3}{2})^6 = \frac{729}{64}$, which is $\frac{729}{512} \approx 1.42$ when place in between 1 and 2, while dividing $\frac{3}{2}$ six times result in $(\frac{2}{3})^6 = \frac{64}{729}$. When placed in between 1 and 2, this ratio is $\frac{1024}{729} \approx 1.40$. This means the when these two notes are played together, their frequency ratio in simplified form is $\frac{531441}{524288}$ and thus not harmonious sounding. This implies that the generated notes from Pythagorean scale sounds different if it is extending towards two different directions, greatly limit its utility and application.

Another big problem of the scale is that multiplying the ratio of $\frac{3}{2}$ will never end up with a multiple of 2. This means that practically, this scale can only work within one octave, thus not useful for musical performances which almost always requires more than one octave in modern music.

The modern scale used for tuning a piano nowadays is equal temperament (Page 2418). Instead of multiplying the ratio of 32, multiply the scale by an irrational number:

$$2^{\frac{1}{12}}$$

This way, when multiplied 12 times, the ratio becomes

$$2^{\left(\frac{1}{12} \times 12\right)} = 2^1 = 2$$

Dividing this ratio 12 times, the ratio becomes

$$2^{\left(-\frac{1}{12} \times 12\right)} = 2^{-1} = \frac{1}{2}$$

This means that either way, this tuning will sound harmonious after playing 12 notes. This is how a 12-note equal temperament scale works: the 12th note is an octave away from the first note.

n general, a n-note equal temperament scale consists of *n-1* unique frequencies. Let f_0 be the base frequency, then frequency of ath note, f_a is defined as:

$$f_a = f_0 \times 2^{\frac{a}{n}}$$

Another important knowledge is that musical notes has their frequencies increase logarithmically, and the analysis of whether two notes sound harmonious depends on the ratios of their frequencies rather than their absolute difference. Alexander J. Ellis invented a unit called cent based on the fact above. Cent is a unit that is helpful for my mathematical analysis of musical harmony, since this unit is a fraction, where increase the frequency by 1200 cents is same as multiplying it by 2. The following figures from Wikipedia illustrates the comparison between the two units.

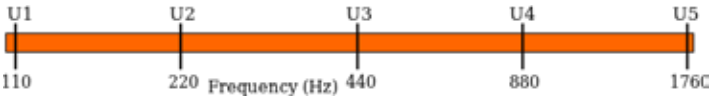


Figure 1.SEQ Figure* ARABIC 16. Octaves increase exponentially when measured on a linear frequency scale (Hz). From Wikipedia

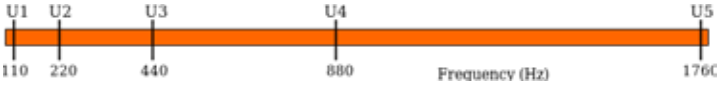


Figure 2. Octaves are equally spaced when measured on a logarithmic scale (cents). From Wikipedia

Schechter has done a mathematical analysis using continued fractions and reached the conclusion that increasing dividing the scale into 29 parts can increase the harmonious level than the modern scale using (42). This research is using knowledge beyond the scope of my Mathematics level. However, the result might be a helpful data for me to check my own results.

2. Discussion

2.1 Mathematical comparison of harmony between notes

Summarising the research, a good way to calculate harmony level is by using relative distance instead of absolute distance since the distance of notes increase logarithmically. The unit of cent will be used since it can represent the relative distance to be more readable than simply a ratio. This calculation also makes more sense from a musical perspective since human ear are less sensitive to a change from 1200 Hz to 1250 Hz comparing to from 400 Hz to 450 Hz despite their absolute distances are both 50 Hz.

The equation calculation for cents c between two frequencies f_0 and f_1 is:

$$c = 1200 \times \log_2\left(\frac{f_1}{f_2}\right)$$

A lower cents means that the two notes sounds close to one another.

This method can be used to calculate how harmonious sounding a note in a scale is. For example, the first note in a 12-note equal temperament scale has a frequency of $2^{\frac{1}{12}}$ times base frequency. One harmonious frequency that is close to this number is $\frac{16}{15}$times base frequency. The frequency this number represents is also based on the base frequency.

$$c = 1200 \times \log_2\left(\frac{\frac{16}{15} \times f_0}{2^{\left(\frac{1}{12}\right)} \times f_0}\right) = 1200 \times \log_2\left(\frac{\frac{16}{15}}{2^{\left(\frac{1}{12}\right)}}\right) = 1200 \times 0.009776071\ldots = 11.731285270\ldots$$

≈ 11.73 cents

Their distance is therefore:

It is more conventional to round to 1 decimal places for musical cents, but I would like to make the calculation more accurate, so I added one more decimal places.

This number represents the harmonious level of this note, which itself is arbitrary. We can't simply use that 11.73 to determine whether these two notes played together is harmonious or not without real-life experiments, but it can be a good basis for comparison. For example, the first note of 13-note equal temperament scale is 19.42 cents if calculated using the same

$$c = 1200 \times \log_2\left(\frac{f_1}{f_2}\right) = 1200 \times \log_2\left(\frac{r_1 \times f_0}{r_2 \times f_0}\right) = 1200 \times \log_2\left(\frac{r_1}{r_2}\right)$$

method.

19.42 >11.73 informs that the first note of a 13-note equal temperament scale is less harmonious than the first note from 12-note equal temperament scale.

As we can see from the calculation above, the calculation can be simplified by comparing the ratio to base frequency instead of the actual frequency. This leads to a new equation, where cents *c* between two frequencies with ratio r_1 and r_2 from base frequency is:

2.1.1 Calculating of the Degree of Harmony of a Scale

One method of comparing scale is taking the cents calculated above for each note and take the average of them to use as a number that represent the harmonious level of a scale.

For 5-note equal temperament scale, taking harmonious ratio $\frac{8}{7}, \frac{4}{3}, \frac{3}{2}$, and $\frac{7}{4}$,, the smallest cents for each notes can be calculated by calculating the cents

$$c = 1200 \times \log_2\left(\frac{f_1}{f_2}\right) = 1200 \times \log_2\left(\frac{r_1 \times f_0}{r_2 \times f_0}\right) = 1200 \times \log_2\left(\frac{r_1}{r_2}\right)$$

of a note compared to all the harm$2^{\frac{1}{5}} = 1.148698\ldots \approx 1.15$.1 then take the smallest.

For instance, the first note has ratio

This is closest to $\frac{8}{7} = 1.1428571\ldots \approx 1.14$ So for the first note the smallest cents is calculated by substituting these two values into the cents equation above.

The table below shows the result.

Notes:	First	Second	Third	Fourth
Smallest Cents	8.83	18.04	18.04	8.83

Table 1. The cents of different notes in 5-note equal temperament scale

The average cents for 5-note equal temperament scale is therefore

(8.83 + 18.04 +18.04 + 8.83) 4 = 13.435 ≈ 13.44.

Looking at another example, we have 12-note equal temperament scale. Taking ratios $\frac{16}{15}, \frac{9}{8}, \frac{6}{5}, \frac{5}{4}, \frac{4}{3}, \frac{7}{5}, \frac{3}{2}, \frac{8}{5}, \frac{3}{2}, \frac{16}{9}, \frac{15}{8}$, we can calculate that its average cents is around 10.12 using the same method above.

Ideally this allows the conclusion that 12-note equal temperament scale will sound more harmonious than the 5-note equal temperament scale, because 10.12 < 13.44. However there is a problem with the harmonious ratios considered. (The harmonious ratios considered are the visually closest ratio taken from Figure. 3.)

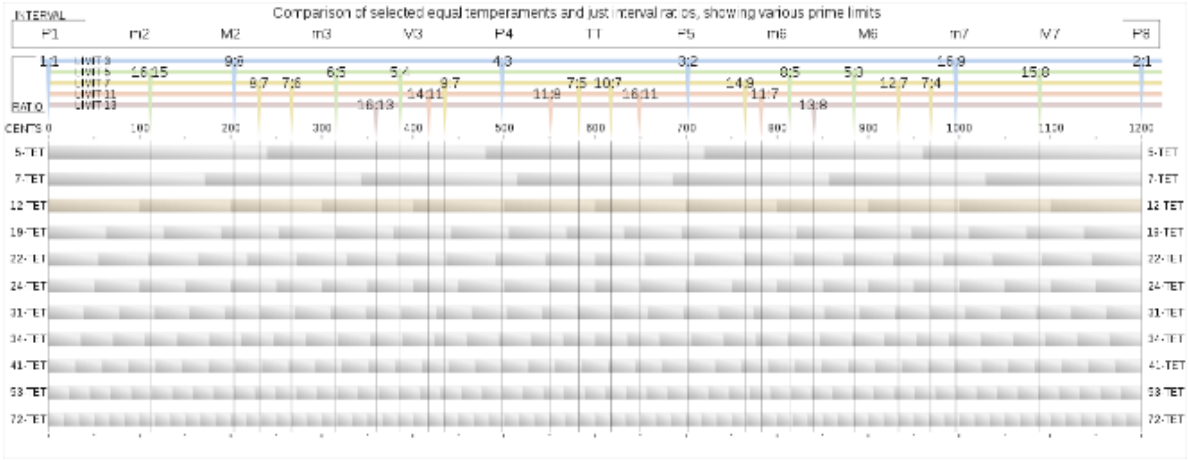


Figure 3. Comparison of selected equal temperaments and just interval ratios, showing various prime limits. From Wikipedia (Appendix - Figure 3)

Theoretically, without any limitation, the chosen ratios can be constructed using two very large numbers, such as $\frac{127}{120}$ instead of $\frac{16}{15}$. The cents difference between ratio of first note of a 12-note-equal temperament scale $2^{\frac{1}{12}}$ and $\frac{16}{15}$ is calculated previously to be 11.73, and using the same method to calculate the cents difference between $\frac{127}{120}$ and $2^{\frac{1}{12}}$ would be 1.85. This shows that by picking a ratio with larger base will allows us to affect the harmonious calculation significantly. The reason behind that is because using ratio $\frac{127}{120}$ would assume that human ear is much more sensitive to changes in pitches. We can assume even higher sensitivity by increasing the denominator in most reduced form, like $\frac{509}{480}$. In general, we can interpret harmonious ratios as assumed human sensitivity to pitches, and without any limit to how the harmonious ratios are chosen, the cents calculated is meaningless.

2.1.2 Using Limit to Compare Scales

One of the ways to determine limit is through prime limit. This is the method used in Figure.3. This method has limits set to prime numbers. The ratios are allow only when their numerator and denominator can both be factored using primes no larger than the prime limit. For example, the number $\frac{9}{8}$, is allow when the limit is 3 because 9 can be factored using only 3 while 8 can be factored using only 2 and 3. The number $\frac{16}{15}$ is not allowed because 15 requires the prime number 5. This method seems to have the benefit of distributing the ratios across the graph more evenly than the other methods. The problem is that this method has outcomes much harder because it is hard to list out all the allowed harmonious ratios concerning prime numbers.

Another method is called odd limit, which allows only "those ratios with identities not larger than [the limit], in which [the limit] is present". (Partch, 74) For example, limit 3 allows $\frac{3}{2}$ and $\frac{4}{3}$, while limit 5 allows numbers from limit 3 as well as $\frac{5}{4}$, $\frac{8}{5}$, $\frac{5}{3}$, and $\frac{6}{5}$.

Considering the two methods, another method that adds more predictability to the odd limit method is by limiting the denominator only.

In general for limit n, the harmonious ratios allowed are $\frac{a}{n}$ where both a and n are integers. The value of $\frac{a}{n}$ must be in between 1 and 2 (exclusive).

For example, limiting the harmonious ratio to have denominator smaller or equal to 2 means only 1, $\frac{3}{2}$ and 2 are allowed. While limiting the harmonious ratio to have denominator smaller or equal to 3 allows a wider numbers of harmonious notes to choose from, including 5 numbers: 1, $\frac{4}{3}$, $\frac{3}{2}$, $\frac{5}{3}$ and 2. This method has a large benefit being the predictability of cents decrease.

Different than the two existing methods for defining limit, this new method allows a calculation of average cents that can then be used as a base for evaluating whether or not the new limit is optimised for the number of notes or not. This method allows us to determine if a certain number of notes would be the most harmonious considering a certain level of harmonious ratio (human ear sensitivity to pitches).

2.2 Method for Evaluate and Compare Different

2.2.1 Calculating Average Cents of a Limit

When limited to limit a, we can calculate its average cents relatively easily. This number also has the trend of decreasing as the number of limit increase. Therefore, this average cents of limit can be then compared to the average cents of the scale for evaluating how well the scale is fitting to that particular limit.

In general, for limit a, we can take all the harmonious ratios and find their cents comparing to the closest lower ratio, up to ratio 2.

For example, the 5-note-equal temperament scale seen from the previous example used ratios $\frac{8}{7}$, $\frac{4}{3}$, $\frac{3}{2}$ and $\frac{7}{4}$, this can be considered limit 7 because the largest denominator used is the 7 in $\frac{8}{7}$.

Starting from 1 as the first frequency, the next frequency will be $\frac{8}{7}$. The cents in between those two notes are around 231.17. Then the next frequency is $\frac{8}{7}$, comparing it to $\frac{9}{7}$ gives 203.91 cents. Doing this for all the notes, except for the last ratio $\frac{13}{7}$ which will be compared to ratio 2, will produce the following table.

From there we can then calculate an average cents for the limit to be around 171.43.

Looking at other limits, we can observe a trend that as the limit increase, this average cents decrease. The average cents for limit 5 is 240.00 which is smaller than the 171.42 of limit 7 which is then smaller than the 102.39 of limit 12. This phenomenon can be therefore used to calculate a optimization value based on percentage.

The average cents of the 5-note equal temperament scale from calculation above was 13.44. That calculation used limit 7 as stated above which gives an average cents of 171.42. This means that the average cent of the scale is 7.84% of the average cents of the limit. This percentage will be referred as optimization percentage of the limit and the scale, and the lower this value is, the better optimized the scale is to that limit.

2.2.2 Calculating Scale Optimization

In the previous example, we find that 5-note equal temperament scale is 7.84% to limit 7. This number itself is still arbitrary and we need another optimization percentage to compare. The following example uses the 12-note equal temperament scale from above. That calculation can be considered limit 15 because the ratio $\frac{16}{15}$ has the largest denominator among all the harmonious ratios used. The average cents of limit 15 is 80.00.

In above sections, it is stated that the average cents for a 12-note equal temperament scale is 10.12. But this is not the average cents of this scale. The average cents of the 5-note one worked because its limit was small enough and the introduction of the limit didn't affect its value. This 12-note equal temperament scale is greatly affected because we can now use closer ratios. The final calculation follows the exact same method as before but added the new ratios of limit 15. The result is that the average cent of the scale is 5.70. This means that the scale is 7.13% of the average cents of the limit. Again, this percentage will be referred as optimization percentage of the limit and the scale.

It is easy to see that 7.13% < 7.84% but we can only say that 12-note equal temperament scale to limit 15 is more optimised than 5-note equal temperament scale to limit 7, which doesn't lead to a meaningful conclusion, because that two optimization percentages are based on two different scales each with a different limit. The following section will apply this optimization percentage to evaluate different limits for 12-note equal temperament scales and different notes for limit 15.

3. Application

Is limit 15 the most optimised limit for a 12-note equal temperament scale?

To answer the question of the subtitle, I decided to use statistical analysis by plotting the optimization percentages of different limits to the 12-note equal temperament scale. This calculation is mostly discrete and thus it is more efficient to write a computer function that takes the input of the number of notes and the limit and produce the final percentage that shows how fitting the limit and the scale are.

Ratios	$\frac{7}{7}$	$\frac{8}{7}$	$\frac{9}{7}$	$\frac{10}{7}$	$\frac{11}{7}$	$\frac{12}{7}$	$\frac{13}{7}$
Cents	231.17	203.91	182.40	165.00	150.64	138.57	128.30

Table 2. The cents of limit 7.

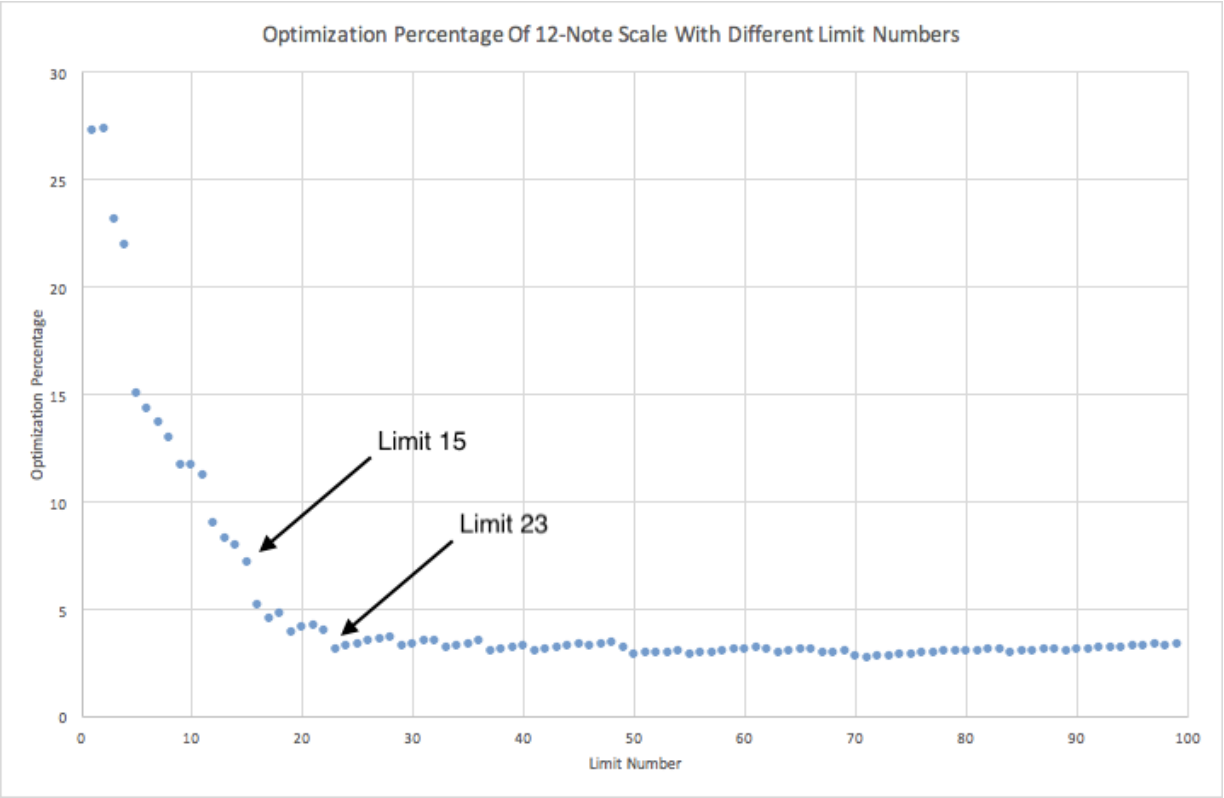


Figure 4. Optimization Percentage of 12-Note Scale With Different Limit Numbers

Figure 4 shows a trend where the Optimization Percentage continue to decrease until around limit 20, where the Optimization Percentage stabilizes around 4%. According to the graph, at Limit 15 (Limit Number = 15), the Optimization Percentage being around 7.5% can be reduced by increasing the limit number. This graph doesn't really show the most optimal number of limit because it seems that the Optimization Percentage will approach a certain number as the limit number keeps increasing. It is therefore not logical to conclude any limit shown to have the least optimization number. However, since we observe the trend that it stabilizes around 4%, the first limit, limit 23, to reach that number can be viewed as the most optimized limit to be picked to analyse 12-note equal temperament scale. Answering the question of subtitle, limit 15 is not the most optimized limit.

Figure 5 shows a trend where the Optimization Percentage converges as the number of notes increase. The limit is around 9% according to the graph. From the graph, we can see that 12 note is not the most optimized scale for limit 15. From the graph, it seems that 9-note equal temperament scale is most optimized for limit 15. It is true that the 1 note scale has a lower optimization percentage, but having two notes in a scale is not at all functional for music production, and therefore it should be excluded. But seeing 2-note having the lowest optimization percentage and 5-note having the highest optimization percentage, it seems that lower number of notes are more likely to be 'lucky' or 'unlucky', where the limit ratios and the equal-temperament scales happen to match very well or not match at all. So in general, we can find graphically that 9-note equal temperament scale is most optimized for limit 15.

3.1 Real-life Application considering aggregated limit

Taking the concept one step further towards real life, we should consider not just one limit, but the average for multiple limits combined. This reflects how individuals have different hearing abilities and sensitivities.

The calculation of the data below simply takes the average of the optimization percentage of the note with different limit from 2 to 20

Is the 12-note equal temperament scale the most optimised one for limit 15?

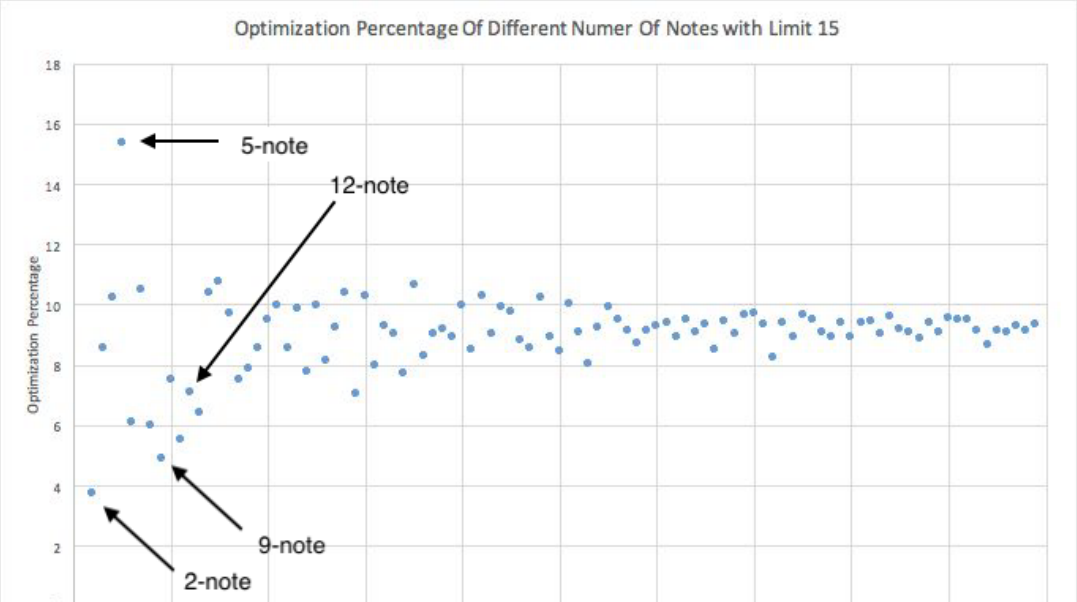


Figure 6. Average Optimization Percentage of Different Number of Notes with Limits from 2 to 20.

Figure 5. Optimization Percentage of Different Number of Notes with Limit 15.

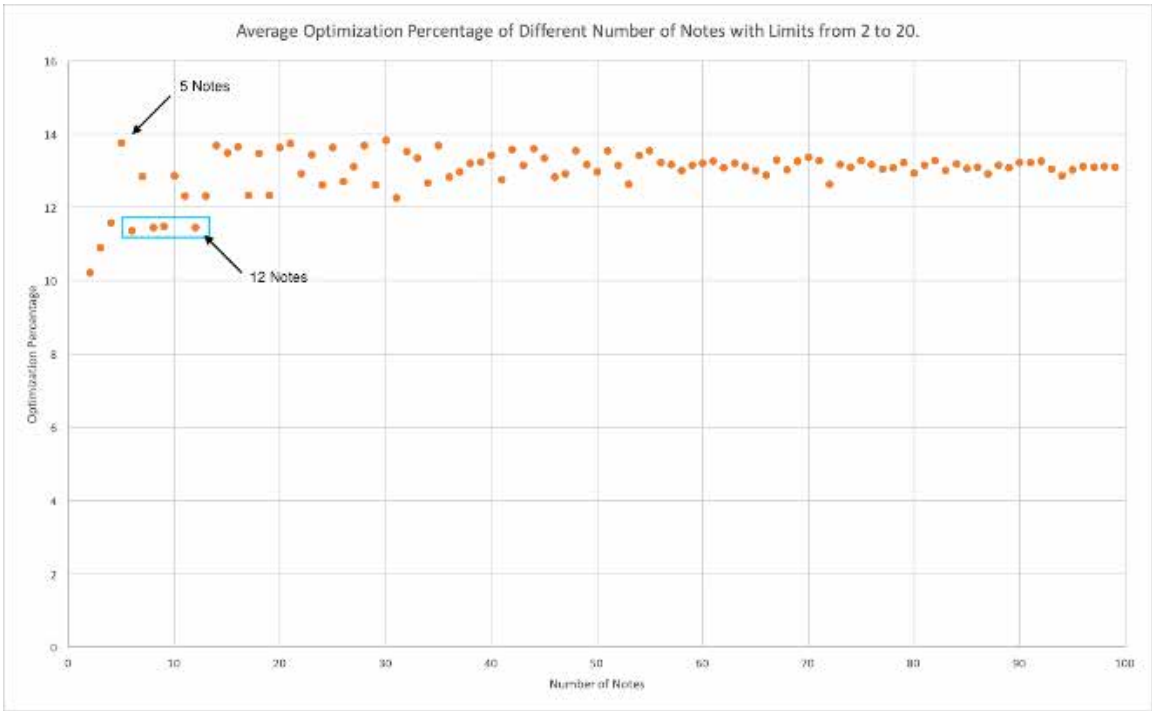


Figure 6. Average Optimization Percentage of Different Number of Notes with Limits from 2 to 20.

Firstly, this graph validates the trend that more notes in the number doesn't decrease the optimization percentage, so the most optimal choice of scale lies in the region where the number of notes is small. The worse scale to pick is probably the 5-note-equal temperament scale because it has the highest optimization percentage. The blue box contains the best scales because the only scales that produce better harmony are 2 and 3 notes scales which are not very useful. The less notes there are in a scale, the less combination there are, and thus less utility for that scale during music production. Among the four inside the blue box, 12-note-equal temperament scale provides the best utility. This finding makes sense because the most widely used scale in music production nowadays is the 12-note-equal-temperament scale. However there is an assumption that limit 20 reflects the limit of human sensitivity and all the limits have the same weighting. This is a crude but generally correct result that can be improved a by real-life testing of human pitch sensitivities.

4. Conclusion

I have explained a new method for comparing the degree of harmony, built on top of the many foundational knowledge including the way to construct a scale as well as a good way to evaluate the logarithmic relation of notes. As explained, I was inspired by the idea of using limit to evaluate a scale, but modified the way of choosing the limit to make it more predictable.

My method seems to be a theoretically promising way to evaluate and generate new scales. It reached the result that 12-note-equal-temperament-scale is one of the best options considering utility of a scale. However, this result doesn't necessarily mean 12-note-equal-temperament-scale is the best because this piece only considered harmonies between two notes. When making music in real life, the harmony are more complex than just two notes. A chord usually contains three or more notes and making those notes harmonious sounding is essential for a piece of music to sound good.

So in general, this new method can inform a way of evaluating the level of harmony of different scales. It does show the potential of using mathematics to analyse the degree of musical harmony. Musicians might use this result to try different scales for music production. For example, a musician intending to make the music sound harsh might intentionally use the 5-note-equal-temperament scale.

In the future, this research can be further improved by integrating real-life experiments maybe in the form of survey or blind test, so that the way a scale is identified as harmonious can be more realistic and applicable to music production.

References

"Comparison of selected equal temperaments and just interval ratios, showing various prime limits" *Equal temperament*, Wikipedia, 18 July. 2008, en.wikipedia.org/wiki/Equal_temperament

Durfee, Dallin S., and John S. Colton. "The Physics of Musical Scales: Theory and Experiment." *American Journal of Physics*, vol. 83, no. 10, July 2015, pp. 835–842., aapt.scitation.org/doi/abs/10.1119/1.4926956.

"Octaves are equally spaced when measured on a logarithmic scale (cents)." *Cent (music),* Wikipedia, 16 June. 2011, en.wikipedia.org/wiki/Cent_(music)

"Octaves increase exponentially when measured on a linear frequency scale (Hz)" *Cent (music*), Wikipedia, 16 June. 2011, en.wikipedia.org/wiki/Cent_(music)

Partch, Harry. *Genesis of a Music*. University of Wisconsin Press, 1949.

Page, Michael F. "Perfect Harmony: A Mathematical Analysis of Four Historical Tunings." *The Journal of the Acoustical Society of America*, vol. 116, no. 4, 2004, pp. 2416–2426., asa.scitation.org/doi/abs/10.1121/1.1788732.

Loosen, Franz. "The Effect of Musical Experience on the Conception of Accurate Tuning." *Music Perception: An Interdisciplinary Journal*, vol. 12, no. 3, 1995, pp. 291–306. JSTOR, JSTOR, www.jstor.org/stable/40286185.

Schechter, Murray. "Tempered Scales and Continued Fractions." *The American Mathematical Monthly*, vol. 87, no. 1, 1980, pp. 40–42. JSTOR, JSTOR, www.jstor.org/stable/2320380.

Artist: Katrina Chan, G9
Title: Reflection in Qing
Medium: Digital painting
Description: This digital artwork was created in response to the Statement of Inquiry "Signs and symbols can express cultural identity". In this unit, students explored both traditional and digital painting to explore signs, symbols and metaphors that express their personal cultural identity.