Single-Cell Phenotyping for the Detection of Antimicrobial Resistance

Alison Farrar

Keble College
University of Oxford

A thesis submitted for the degree of Doctor of Philosophy

Trinity 2024

Abstract

Ultra-rapid antibiotic susceptibility tests (ASTs) that can return results within a few hours are increasingly important for properly treating antibiotic-resistant infections and for minimising unnecessary antibiotic use. Antibiotic susceptibility phenotyping on a single cell basis, through the visualisation of intracellular changes associated with the antibiotic response, is a promising technology for rapid testing because it could eliminate lengthy sample culturing steps and provide a rich image dataset to study resistance mechanisms and bacterial heterogeneity. These data can be used to train convolutional neural networks to rapidly identify phenotypes associated with antibiotic resistance or susceptibility. In this work, we extend these methods by showing that the spatial distribution of fluorescently labelled ribosomes, or the ribosome phenotype, can be used as an AST. We show that ribosomes are a useful target for single-cell phenotyping because they are abundant, outline the cell periphery for image segmentation, and act as an inverse DNA stain. Ribosomes also contain RNA sequences that can be used for species identification, if species-specific probes are used.

This thesis investigates three questions related to the single-cell phenotyping for the detection of antimicrobial resistance. First, can the ribosome phenotype be used to detect antibiotic susceptibility for different classes of antibiotics? Using *E. coli* MG1655 as a model organism, we show that a convolutional neural network can identify the antibiotic response for ciprofloxacin, gentamicin, chloramphenicol, and carbenicillin with high accuracy (99%, 98%, 95% and 99%). Second, can the ribosome phenotype be used to identify antibiotic-susceptible and resistant pathogenic *E. coli*? We find that ribosome phenotypes can be used to classify clinical isolates with 91% accuracy for susceptible cells and 99% accuracy for resistant cells. Third, what image features are associated with misclassifications? We investigate this using data from the citizen science project Infection Inspection, which ran from February to April 2023.

Single-Cell Phenotyping for the Detection of Antimicrobial Resistance



Alison Farrar

Keble College

University of Oxford

A thesis submitted for the degree of

Doctor of Philosophy

Trinity 2024

Statement of Originality

The work presented in this thesis was conducted by me under the supervision of Achillefs Kapanidis and Hafez El Sayyed at the University of Oxford from April 2021 to September 2024. This work was supported by the Clarendon Scholarships and by the Oxford Martin School Programme on Antimicrobial Resistance Testing. No part of this thesis has been submitted in support of another degree, diploma, or other qualification at the University of Oxford or any other university. Any contribution made to this research by collaborators is explicitly acknowledged in the text and detailed below.

- Chapters 2-4 are extended versions of my publication "Ribosome phenotypes for rapid classification of antibiotic-susceptible and resistant strains of *Escherichia coli*," including text and figures. Achillefs Kapanidis and Nicole Stoesser edited the text of this manuscript. Piers Turner designed the single-cell ribosome image classification pipeline.
- Sammi Ta conducted the experiments to determine carbenicillin MIC in Chapter 2.
- The *E. coli* clinical isolates were obtained through a collaboration with the John Radcliffe Hospital microbiology laboratory with help from Conor Feehily and Nicole Stoesser as part of the Oxford Martin School Programme on Antimicrobial Resistance.
- Chapter 5 is an extended version of my publication on the work of the Infection Inspection project, "Infection Inspection: using the power of citizen science for image-based prediction of antibiotic resistance in *Escherichia coli*," including text and figures. Conor Feehily contributed to the writing and editing of the manuscript, and led the work on accuracy and participation analysis.
- The Infection Inspection project was conceptualised by myself and the other members of the Oxford Martin School AMR Testing team. Conor Feehily and I developed the project workflow, tutorial, and field guide. The bacterial micrographs utilised for the project were collected by Alexander Zagajewski and curated by Piers Turner.

Alison Farrar
September 2024

Publications

1. **Farrar, A.**, Turner, P., El Sayyed, H., Feehily, C., Chatzimichail, S., Ta, S., Crook, D., Andersson, M., Oakley, S., Barrett, L., Nellåker, C., Stoesser, N., Kapanidis, A.N. (2025). Ribosome phenotypes for rapid classification of antibiotic-susceptible and resistant strains of *Escherichia coli*. *Commun Biol* **8**, 319 (2025). DOI:10.1038/s42003-025-07740-6.

2. **Farrar, A.***, Feehily, C.*, Turner, P., Zagajewski, A., Chatzimichail, S., Crook, D., Andersson, M., Oakley, S., Barrett, L., El Sayyed, H., Fowler, P., Nellåker, C., Kapanidis, A., Stoesser, N. Infection Inspection: using the power of citizen science for image-based prediction of antibiotic resistance in *Escherichia coli* treated with ciprofloxacin. *Sci Rep* **14**, 19543 (2024). DOI:10.1038/s41598-024-69341-3.

3. Chatzimichail, S., Turner, P., Feehily, C., **Farrar, A.**, Crook, D., Andersson, M., Oakley, S., Barrett, L., Sayyed, H.E., Kyropoulos, J., Nellåker, C., Nicole Stoesser, N., Kapanidis, A.N. Rapid Identification of Bacterial isolates Using Microfluidic Adaptive Channels and Multiplexed Fluorescence Microscopy. *Lab Chip* **24**, 4843-4858 (2024). DOI:10.1039/D4LC00325J.

Abstract

Ultra-rapid antibiotic susceptibility tests (ASTs) that can return results within a few hours are increasingly important for properly treating antibiotic-resistant infections and for minimising unnecessary antibiotic use. Antibiotic susceptibility phenotyping on a single cell basis, through the visualisation of intracellular changes associated with the antibiotic response, is a promising technology for rapid testing because it could eliminate lengthy sample culturing steps and provide a rich image dataset to study resistance mechanisms and bacterial heterogeneity. These data can be used to train convolutional neural networks to rapidly identify phenotypes associated with antibiotic resistance or susceptibility. In this work, we extend these methods by showing that the spatial distribution of fluorescently labelled ribosomes, or the ribosome phenotype, can be used as an AST. We show that ribosomes are a useful target for single-cell phenotyping because they are abundant, outline the cell periphery for image segmentation, and act as an inverse DNA stain. Ribosomes also contain RNA sequences that can be used for species identification, if species-specific probes are used.

This thesis investigates three questions related to the single-cell phenotyping for the detection of antimicrobial resistance. First, can the ribosome phenotype be used to detect antibiotic susceptibility for different classes of antibiotics? Using *E. coli* MG1655 as a model organism, we show that a convolutional neural network can identify the antibiotic response for ciprofloxacin, gentamicin, chloramphenicol, and carbenicillin with high accuracy (99%, 98%, 95% and 99%). Second, can the ribosome phenotype be used to identify antibiotic-susceptible and resistant pathogenic *E. coli*? We find that ribosome phenotypes can be used to classify clinical isolates with 91% accuracy for susceptible cells and 99% accuracy for resistant cells. Third, what image features are associated with misclassifications? We investigate this using data from the citizen science project Infection Inspection, which ran from February to April 2023.

Acknowledgements

I could not have asked for a happier experience than I had in my 3+ years in the Kapanidis group, learning and growing as a scientist. I want to thank every current and former member of the Kapanidis group who made going to work so exciting and fun. Lots of labs might call themselves interdisciplinary, but we live and breathe it! You're extraordinary scientists, and above all, you're my bros — all of you. Lunch doesn't feel the same without your banter. Through every rollercoaster or whitewater rapid, I knew I could count on you. I feel honoured to call all of you my colleagues and friends. You inspire me.

I want to especially thank my supervisors, Prof. Achillefs Kapanidis and Dr. Hafez El Sayyed, for being excellent scientists and mentors. Achillefs: thank you for all of your wisdom and mentorship, your scientific guidance, and for teaching me how to be a calm and confident leader by your example. Hafez: your friendship (and tabbouleh) is boundless and beyond comparison; thank you for the kind guidance you've always given me in work and in life.

Thank you to the researchers who made the AMR project possible: Dr. Piers Turner, Dr. Conor Feehily, and Dr. Stelios Chatzimichail. I could not have dreamed of a better team to work with for my DPhil. Thanks for always being up for a brainstorming or troubleshooting session, and for all you have taught me. We should be proud of what we accomplished, and I can't wait to see what comes next.

To Dr. Nicole Stoesser and Dr. Christoffer Nellåker: I never knew when I joined this team that I would get two "bonus" supervisors. Your expertise in medicine and in machine learning has been enormously helpful and valued.

To Dr. Claudia Mewes: working with you gave me a passion for science and for physics. Thank you for being a fabulous mentor. As hard as it might be for you to believe, this DPhil was done in the Condensed Matter Physics department. I guess I never left CMP after all.

To Kathleen Summerland: you have been my teacher and friend for as long as I can remember. Even when it was hard for me to believe it, you always knew I could do things beautifully and wonderfully. Thank you for teaching me the French Suite, and so much more.

To the many friends outside of lab who made my time in Oxford so special, I want to thank you from the bottom of my heart. Even when our government-permitted

activities were limited to fajita nights, kitchen pilates, and Keble brunch, we made the best of it. Whether on Hurst Street or Walton Well Road, I was lucky to always have treasured friends to come home to. I will never forget our garden chats and coffee walks. You will always have a friend to visit, wherever I am.

I want to thank a few of the organisations that allowed me to do this DPhil, including the BBSRC Interdisciplinary Bioscience Doctoral Training Programme, the Oxford Martin School Programme on Antimicrobial Resistance, the Clarendon Scholarships, and the Randall Research Scholars Program at The University of Alabama. Roll Tide! I also want to thank Oxford Nightline and the Parkrun organisation for giving my evenings and Saturday mornings extra purpose and meaning.

I want to thank my parents for supporting me in every dream I've ever had and for passing down some fraction of their talents in graphic design. Mom and Dad, nobody could teach the values of hard work, creativity, and love better than you.

Char, we proved it — there's nobody I'd rather be stuck on an island with. Thank you for taking me to see all the bands I've never heard of and all the cool countries nobody goes to. You don't get to pick your siblings, but I'd always choose you.

This thesis is dedicated to Aunt Vicky, who taught me that everyone can be a friend, and that nobody cares how long it took you, they just see the result. In honour of your memory, I hope this work helps make antibiotic-resistant infections a little less terrible.

Contents

List of Figures **x**

List of Abbreviations **xii**

1 Introduction **1**
1.1 The rise of antimicrobial resistance . . . 1
1.2 Mechanisms of antimicrobial resistance . . . 9
1.3 Detection of antimicrobial resistance . . . 12
1.4 Bacterial antibiotic response phenotyping . . . 18
1.5 Machine learning for antimicrobial resistance detection . . . 19
1.6 Towards a rapid phenotypic test . . . 21
1.7 Thesis outline and aims . . . 22

2 Development of a single-cell ribosome phenotype assay **23**
2.1 The ribosome phenotype . . . 24
2.2 *Escherichia coli* as a model organism . . . 26
2.3 Experimental strains . . . 27
2.4 Antibiotics . . . 27
2.5 Sample preparation . . . 29
2.5.1 Antibiotic treatment and fixation . . . 29
2.5.2 FISH . . . 31
2.6 Image acquisition . . . 31
2.7 Image processing and segmentation . . . 32
2.8 Neural network training . . . 32
2.9 Accuracy metrics . . . 33
2.10 Cell phenotype statistics . . . 33
2.11 Sample size simulation . . . 34
2.12 Discussion . . . 35
2.13 Conclusions . . . 36

3 Ribosome phenotypes for detecting antibiotic response 38
3.1 Introduction . . . 39
3.2 Characterisation of the *E. coli* antibiotic response by ribosome distribution . . . 40
3.3 Strategies to maximise transferable learning . . . 46
3.4 Identification of antibiotic-susceptible ribosome phenotypes by a neural network . . . 50
3.5 Discussion . . . 54
3.5.1 Strains and antibiotics . . . 55
3.5.2 Neural network training and augmentations . . . 57
3.6 Conclusions . . . 59

4 Detecting antimicrobial resistance in clinical isolates by ribosome phenotype 60
4.1 Introduction . . . 61
4.2 Model trained on lab strain has variable accuracy on clinical isolates 63
4.3 Developing a susceptible-resistant clinical isolate classifier . . . 65
4.4 Necessary sample size for testing . . . 69
4.5 Model performance on unseen strains and antibiotic concentrations 71
4.6 Discussion . . . 74
4.7 Conclusions . . . 76

5 Citizen science for image feature analysis of antibiotic phenotype misclassifications 77
5.1 Introduction . . . 78
5.2 Methods . . . 81
5.2.1 Image dataset . . . 81
5.2.2 Accuracy analysis . . . 84
5.2.3 Image feature analysis . . . 84
5.3 Development of the Infection Inspection project . . . 86
5.3.1 User participation . . . 87
5.3.2 Accuracy analysis . . . 89
5.4 Results of feature analysis of sensitive and resistant phenotypes . . 90
5.4.1 Differences in DNA morphology drive confusion . . . 90
5.4.2 Incorrectly classified images have more ambiguous phenotypes 94
5.4.3 Feature importance in ciprofloxacin-sensitive and resistant cells 97
5.4.4 Phenotypic differences in resistant *E. coli* treated with high ciprofloxacin concentrations . . . 97
5.5 Discussion . . . 99
5.5.1 Feature analysis . . . 99

5.5.2 Volunteer participation 100
5.5.3 Limitations 101
5.5.4 Future directions 102
5.6 Conclusions 103

6 Conclusion **104**

References **110**

Appendices

A Data Availability and Ethics **127**
A.1 Software and Data 127
A.2 Ethics approval 127

List of Figures

1.1 Timeline of AMR . . . 3
1.2 Antibiotic consumption by class (England, 2023) . . . 5
1.3 Deaths attributable to AMR by 2050 . . . 6
1.4 Mechanisms of resistance acquisition . . . 9
1.5 Molecular mechanisms of resistance . . . 11
1.6 Conventional AST methods . . . 13
1.7 AST time to result . . . 14

2.1 *E. coli* bacterial morphology . . . 27
2.2 Protocol for bacterial sample preparation . . . 30
2.3 Protocol for cell staining . . . 31

3.1 Proposed ribosome phenotyping protocol . . . 40
3.2 Antibiotic-susceptible ribosome phenotypes . . . 41
3.3 Phenotypic changes with gentamicin at various concentrations . . . 42
3.4 Phenotypic changes associated with carbenicillin treatment at various concentrations . . . 43
3.5 Fluorescence properties by antibiotic . . . 45
3.6 Cross validation by rotating holdout . . . 46
3.7 Image resizing . . . 47
3.8 Image augmentations with and without shear and blur . . . 49
3.9 Effects of shear and blur . . . 49
3.10 Ribosome phenotype classifier accuracy by antibiotic . . . 50
3.11 Representative classifications for each antibiotic . . . 51
3.12 Confusion matrices for the ribosome phenotype classifier . . . 53
3.13 Cells classified as ciprofloxacin-treated . . . 54
3.14 Cells incorrectly classified by the chloramphenicol model . . . 54
3.15 Model accuracy with and without shear and blur . . . 59

4.1 Classifier trained on *E. coli* MG1655 applied to clinical isolates . . . 64
4.2 Susceptible-resistant classifier differentiates clinical isolates by MIC 66
4.3 Confusion matrix for the susceptible-resistant classifier . . . 67
4.4 Susceptible-resistant classifier accuracy by clinical isolate . . . 68

4.5 Accuracy of susceptible-resistant classifier compared to MG1655 classifier by strain 69
4.6 Sample size simulation for the susceptible-resistant classifier 70
4.7 1× EUCAST classifier accuracy on cells treated at 20× EUCAST . 72
4.8 20× EUCAST classifier accuracy on cells treated at 1× EUCAST . 73

5.1 Infection Inspection workflow 80
5.2 Distribution of the number of classifications per user 88
5.3 User engagement with Infection Inspection 89
5.4 Image classification accuracy 91
5.5 User accuracy varied by image for both resistant and sensitive cells 92
5.6 Feature measurements associated with correct and incorrect classifications 93
5.7 Phenotype analysis of correctly and incorrectly classified cells by PCA 96
5.8 Feature significance measured by SHAP values 98
5.9 Features of resistant *E. coli* treated with high concentrations of ciprofloxacin 99

List of Abbreviations

AMR . . . . . Antimicrobial resistance.

AST . . . . . Antimicrobial susceptibility test.

Cam . . . . . Chloramphenicol.

Carb . . . . . Carbenicillin.

CFU . . . . . Colony-forming unit.

Cip . . . . . Ciprofloxacin.

CNN . . . . . Convolutional neural network.

CT . . . . . Computed tomography.

Cy3 . . . . . Cyanine-3.

DAPI . . . . . 4',6-diamidino-2-phenylindole.

DNA . . . . . Deoxyribonucleic acid.

E. coli . . . . . *Escherichia coli*.

EUCAST . . . European Committee on Antimicrobial Susceptibility Testing.

FISH . . . . . Fluorescence *in situ* hybridisation.

Gent . . . . . Gentamicin.

ID . . . . . Identification.

IQR . . . . . Inter-quartile range.

MIC . . . . . Minimum inhibitory concentration.

MRSA . . . . . Methicillin-resistant *Staphylococcus aureus*.

NPV . . . . . Negative predictive value.

PBS . . . . . Phosphate-buffered saline.

PCA . . . . . Principal component analysis.

PCR . . . . . Polymerase chain reaction.

PPV . . . . . Positive predictive value.

RGB . . . . . Red-green-blue.

RNA . . . . . Ribonucleic acid.

SHAP . . . . . SHapley Additive exPlanations.

TTR . . . . . Time to result.

1 Introduction

Contents

1.1 The rise of antimicrobial resistance 1
1.2 Mechanisms of antimicrobial resistance 9
1.3 Detection of antimicrobial resistance 12
1.4 Bacterial antibiotic response phenotyping 18
1.5 Machine learning for antimicrobial resistance detection 19
1.6 Towards a rapid phenotypic test 21
1.7 Thesis outline and aims 22

1.1 The rise of antimicrobial resistance

Effective and affordable antimicrobial compounds made possible the rising quality of life and lifespan seen around the world over the past century; without them, modern surgery, chemotherapy, maternal and neonatal care, and agriculture could not exist [1]. In 1900 in the United States, 1 in 2 people died from pneumonia, influenza, tuberculosis, enteritis, or diphtheria [2]. After the introduction of penicillin, the first widely available modern antibiotic drug, deaths caused by infectious diseases plummeted by more than 70% [3]. In developed nations, their impact has changed cultural expectations of disease and death so completely that it is hard for many to imagine the anxieties of the pre-antibiotic world, when infections were so prevalent

and so fatal. Perhaps only the early days of the COVID-19 pandemic, before the vaccine, come close [4].

Despite being true "miracle drugs," antibiotics have tempered rather than eliminated the threat of bacterial infections. Although pathogenic species of bacteria are far out-numbered by commensal, harmless, and beneficial species (∼0.15% vs. 99.85%), bacterial infections were associated with 14% of all global deaths and the majority of sepsis-related deaths in 2019 [5, 6, 7]. Five leading pathogens (*Staphylococcus aureus*, *Escherichia coli*, *Streptococcus pneumoniae*, *Klebsiella pneumoniae*, and *Pseudomonas aeruginosa*) were responsible for more than half of those infections [7]. Among these, *Escherichia coli* has been the leading cause of bloodstream infections since 2005 [8].

The widespread use of antibiotics in the treatment and prevention of bacterial infections is complicated by the strong evolutionary pressure for microbes resistant to these compounds [9]. However, the bacterial fight for survival is also the source of most clinically relevant classes of antibiotics. Before antimicrobial chemicals were ever isolated or invented, Louis Pasteur and others proposed that microbes secreted compounds that they used to compete with and kill other species [10]. It has since been shown that naturally occurring antibiotic compounds are synthesised by microbes for use as intra-community signalling molecules and as inter-community bacterial warfare [11]. The antimicrobial properties of *Penicillium* moulds and soil-dwelling bacteria may be why mouldy bread and soil were named as remedies on an Egyptian papyrus from 1550 B.C.E. [12]. Even today, most clinically used antibiotics are derived from natural products, from *Actinomycetes* (e.g. aminoglycosides, tetracyclines, amphenicols, macrolides); fungi (e.g. penicillins, cephalosporins); or other bacteria (e.g. polypeptides, bacitracin, polymyxins) [13]. For this reason, genes that confer resistance to antibiotic compounds may exist prior to the use of an antibiotic compound by humans. In many cases, resistance is discovered shortly after the discovery of the antibiotic as an effective drug—or, in the case of penicillin, even before its use in humans (Figure 1.1).

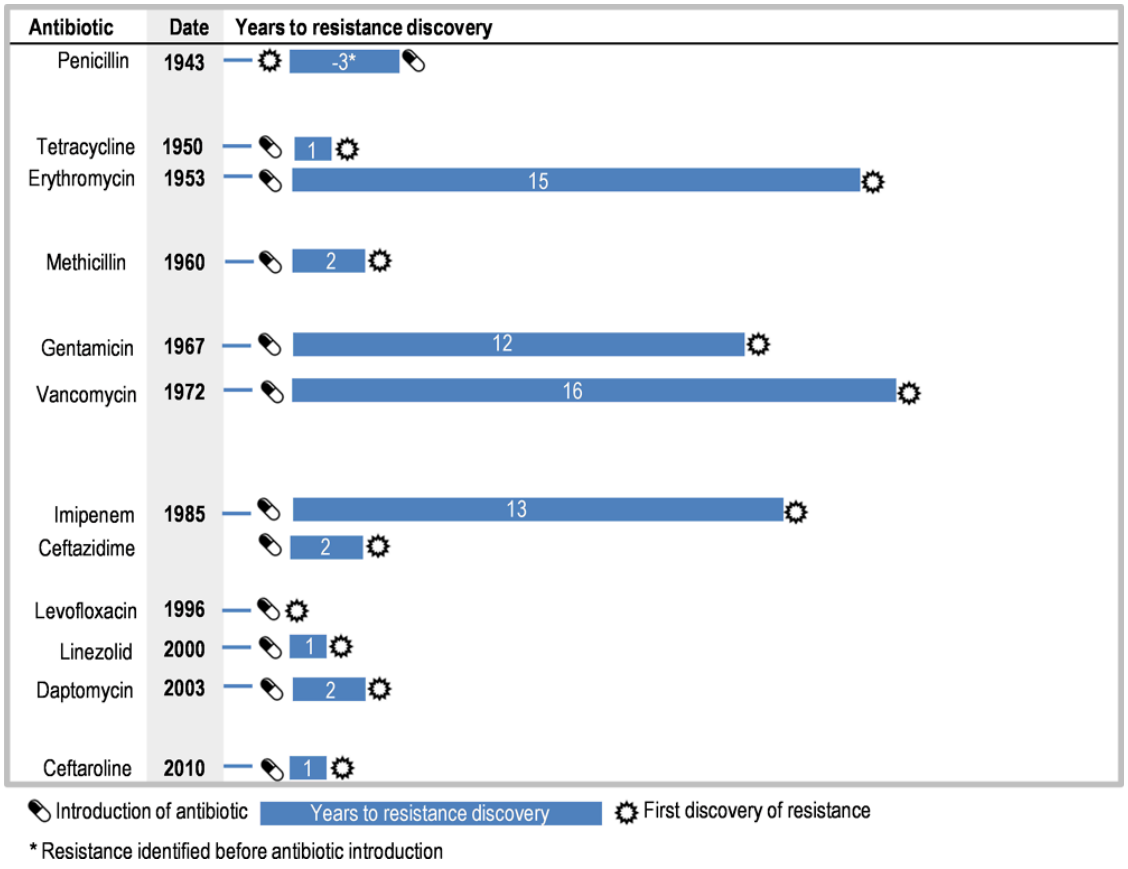


Figure 1.1: Timeline of antibiotic discovery and resistance. The time between antibiotic discovery and the discovery of resistance is shown for many antibiotic compounds. Figure from the Organisation for Economic Co-operation and Development report "Stemming the Superbug Tide." [3]

The ideal antibiotic should kill or inhibit bacterial growth at the sight of infection, while causing minimal side effects to the patient and being selective to the targeted pathogen [14]. Most small molecule antimicrobial compounds accomplish this by selectively interfering with or inhibiting processes that are essential for the bacterial life cycle. Broadly, the antibiotics commonly in use today target bacterial cell wall synthesis, nucleic acid synthesis, and protein synthesis (Table 1.1) [15, 16]. Antimicrobial compounds can also be categorised as bactericidal or bacteriostatic based on whether they kill bacteria or inhibit bacterial growth by slowing metabolic processes [17].

The beta-lactam drugs, including penicillins and cephalosporins, are currently the most used class of antibiotics, accounting for approximately half of all global use [18]. The beta-lactams are also the most used drugs in England, making up 40%

Class	Target	Example(s)	Primary effects
Fluoroquinolones	Topoisomerase II, topoisomerase IV	Ciprofloxacin*	DNA replication
Trimethoprim sulfamethoxazole	Tetrahydrofolic acid synthesis	Co-trimoxazole	Nucleotide biosynthesis
Rifamycins	DNA-dependent RNA polymerase	Rifampin	RNA transcription
Beta-lactams	Penicillin-binding proteins	Penicillins (Carbenicillin*), Cephalosporins, Carbapenems	Cell wall synthesis
Glyco (lipo) peptides	Peptidoglycan D-Ala-D-Ala units	Vancomycin	Cell wall synthesis
Lipopeptides	Cell membrane	Daptomycin	Cell wall synthesis
Aminoglycosides	30S ribosome	Gentamicin*	Translation: tRNA mismatching
Tetracyclines	30S ribosome	Tetracycline	Translation: aminoacyl tRNA binding
Macrolides	50S ribosome	Erythromycin	Translation: elongation and translocation
Streptogramins	50S ribosome	Dalfopristin	Translation: initiation, elongation, and translocation
Phenicols	50S ribosome	Chloramphenicol*	Translation: elongation

Table 1.1: Antibiotic classes and their primary effects on bacterial physiology. Major antibiotic classes are listed with their primary target, examples of drugs in the class, and the primary pathways affected by the antibiotic mechanism of action. Antibiotics used in this thesis are labelled with (*). Content derived from Kohanski *et al.*, 2010 [16].

of all antibiotic doses in 2023, followed by tetracyclines (26%) and macrolides (14%) (Figure 1.2).

In less than a century, our use and misuse of modern antibiotics has created a crisis. In 2019, antibiotic or antimicrobial resistance (AMR) caused 1.27 million deaths and was associated with 4.95 million deaths worldwide [20]. If no action is taken, global mortality is projected to rise as high as 10 million deaths per year by 2050 [21]. If this happens, it will kill more people every year than cancer

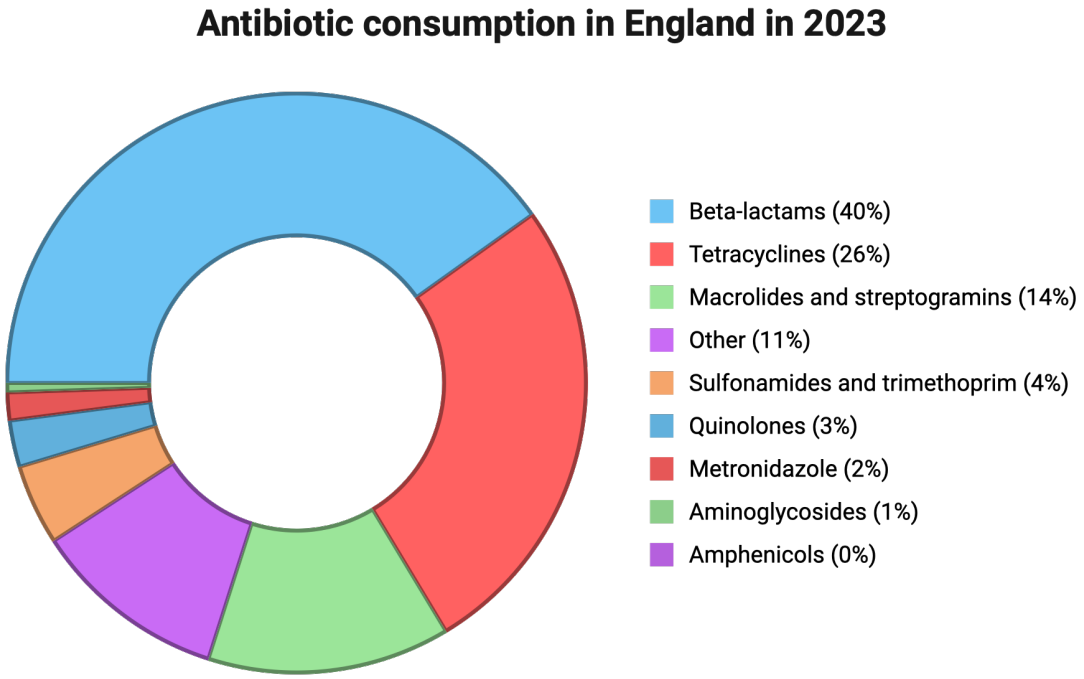


Figure 1.2: Antibiotic consumption in England in 2023, by class. Total antibiotic consumption in England in 2023, divided by class, measured in Defined Daily Doses per 1,000 individuals. Data derived from Public Health England [19]. Figure made with Biorender.com.

— a return to the pre-antibiotic era.

Tragically, the burden of antibiotic-resistant infections is expected to disproportionately affect low- and middle-income countries, overburdening weak healthcare infrastructures and increasing global economic inequality (Figure 1.3) [22]. Regionally, levels of antibiotic use and AMR prevalence are closely associated, and are both influenced by a variety of systemic, cultural, and environmental factors [23]. Within the healthcare system, antibiotic prescription rates are influenced by clinician education, fear of treatment failure, antibiotic prescription guidelines, and patient expectations [24, 25]. Slow diagnostic methods mean that clinical teams often have to choose antibiotic treatments empirically, at least for the hours or days until the antibiotic susceptibility results come back. In his landmark report commissioned by the government of the United Kingdom, economist Jim O'Neill wrote, "I find it incredible that doctors must still prescribe antibiotics based only

on their immediate assessment of a patient's symptoms, just like they used to when antibiotics first entered common use in the 1950s" [26]. In many parts of the world, limited access to formal healthcare means that non-prescription antibiotic consumption accounts for 19-100% of antibiotic use outside northern Europe and North America [27]. This, in addition to environmental health factors like nutrition and hygiene, leads to the unequal burden of AMR and infectious disease [28]. For these reasons, the strategy for antibiotic conservation must be tailored to the local context. The World Health Organisation has supported 178 countries in crafting national action plans aligned with their global strategy [29].

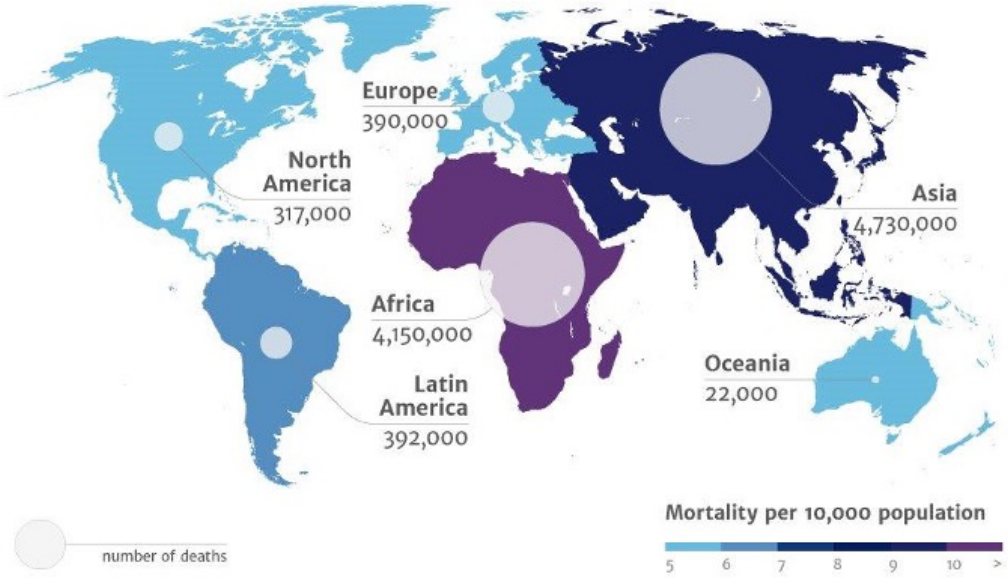


Figure 1.3: Global projection of deaths attributable to AMR by 2050. The O'Neill report estimated that by, 2050, 10 million deaths per year will be caused by AMR infections, with disproportionate impacts on low- and middle-income countries. Figure from "The Review on Antimicrobial Resistance" [21].

Biologically, the rise of AMR is an unavoidable reality, but we have the power to prevent the worst-case scenario from coming true. Modelling suggests that increasing funding for public health interventions in low-income and middle-income countries could reduce the global AMR burden by 10% by the year 2030, preventing as many as 1 million deaths per year by improving access to clean water, sanitation, and vaccines [30]. One AMR action plan presents a cost-effectiveness analysis in which for every US$ 1 invested in new antimicrobial drugs, environmental hygiene,

diagnostics, and improved access to latrines, a net return between US$ 7.2 - 13.1 is expected in the form of saved productivity and lifespan [31]. In this plan, about 1/4 of funding was directed to novel antimicrobial drug discovery, 1/4 to improved access to latrines, 1/5 to diagnostics and antimicrobial stewardship programmes, and the rest to other interventions. For relatively low costs, these programmes would save lives and improve quality of life in the places they are implemented, and would decrease the burden of AMR worldwide.

In regions where public knowledge about AMR is limited, educational interventions can also serve to improve awareness and change behaviours related to AMR, such as the sharing of antibiotics, using antibiotics without proper diagnosis, and not finishing treatment courses [32]. Research conducted on 12,000 people in Germany, India, Japan, Kenya, the UK, the USA, and Thailand showed that by communicating the threat of AMR to modern medicine, explaining the fundamental science, emphasising the universal threat to all of us, and framing AMR as a solvable issue, people were more motivated to take action [33].

For healthcare providers, digital health tools have shown promise. Treatment algorithms and antibiotic stewardship plans are already widespread in well-resourced healthcare settings, but most antibiotics in low- and middle-income countries are obtained from pharmacies or informal providers, where there is little regulation or decision-making support. In Tanzania, a randomised-controlled trial showed that a digital clinical decision support algorithm combined with point-of-care tests for C-reactive protein, haemoglobin, and blood oxygenation reduced antibiotic prescription rates from 70.1% to 23.2% with no difference in treatment failure or safety outcomes [34]. The lessons learned from trials like this one can aid the development of novel diagnostics, designed to fit within clinical workflows and to be accessible where they are most needed. Rapid, low-cost, antigen-based tests can be targeted to key clinical questions, such as detecting *Neisseria gonorrhoeae* in primary care facilities [35] or detecting specific resistance-conferring enzymes in clinical samples [36]. Other surveys have also shown point-of-care tests to be cost-effective and to reduce antibiotic prescriptions [37]. These results suggest

that diagnostic algorithms can encourage rational use of antibiotics, and reduce prescriptions based on the fear of a bad outcome.

Together with antibiotic conservation, the discovery of new antimicrobial compounds is essential. New drugs will be necessary as existing compounds become increasingly ineffective against pathogenic bacteria. For example, in an intensive care unit in the United States, around half of *Pseudomonas aeruginosa*, *Klebsiella pneumoniae*, and *Methicillin-Resistant Staphylococcus aureus* (*MRSA*) infections and almost 100% of *Acinetobacter baumannii* infections were found to be multi-drug resistant, meaning that they were resistant to at least one drug in three or more antimicrobial classes [38]. The number of drugs in the pre-clinical antibacterial pipeline is small but increasing [39]. It is difficult for novel antimicrobial drugs to be developed through traditional, profit-driven pipelines, especially when the process from discovery to regulatory approval can cost US$ 1 billion and drugs are most needed by low-income patients and governments which can't afford to cover the cost [40]. Collaborative strategies, like the TB Drug Accelerator, that bring multiple companies together to share compound libraries, datasets, risk, and credit, have accelerated development. The TB accelerator has already produced 8 candidate drugs for tuberculosis [41]. From a policy perspective, governments are also moving from "push" incentives, which aim to reduce the costs of antimicrobial research but have been insufficient, and adding "pull" incentives, which aim to improve long-term profitability and are often tied to drug efficacy and global availability [42]. As of June 2024, there were 113 novel antimicrobials in development, and 6 of the 12 new marketed drugs were small molecules targeting pathogens on both the World Health Organisation critical and the Centers for Disease Control and Prevention urgent lists [43]. Non-traditional antimicrobial agents such as antibodies, bacteriophages, immunomodulators, and microbiome therapies represented 38% (43/113) of the treatments in development, which is promising for the future of these agents for treating bacterial infections. Bacteriophage therapy may work synergistically with other antimicrobial compounds and can reduce the spread and acquisition of antimicrobial resistance in bacterial populations [44].

1.2 Mechanisms of antimicrobial resistance

On a population level, the mechanisms by which bacteria survive exposure to antimicrobial compounds are what we call antimicrobial resistance. These mechanisms can arise *de novo* from errors in DNA replication, which occur more often in stressed states and enable drug-insensitive progeny to survive [45]. More commonly, within microbiomes, strains can acquire mobile resistance genes by horizontal transfer [46]. Resistance-conferring genes can be acquired through phage-mediated transduction, transformation of naked DNA, membrane vesicles, or cell-to-cell conjugation (Figure 1.4). Compared to the spontaneous development of resistance as a result of antibiotic exposure, plasmid-associated horizontal gene transfer is particularly responsible for the spread of multi-drug resistance among Gram-negative bacteria, as well as resistance to antibiotics of last resort, such as colistin and carbapenems [47].

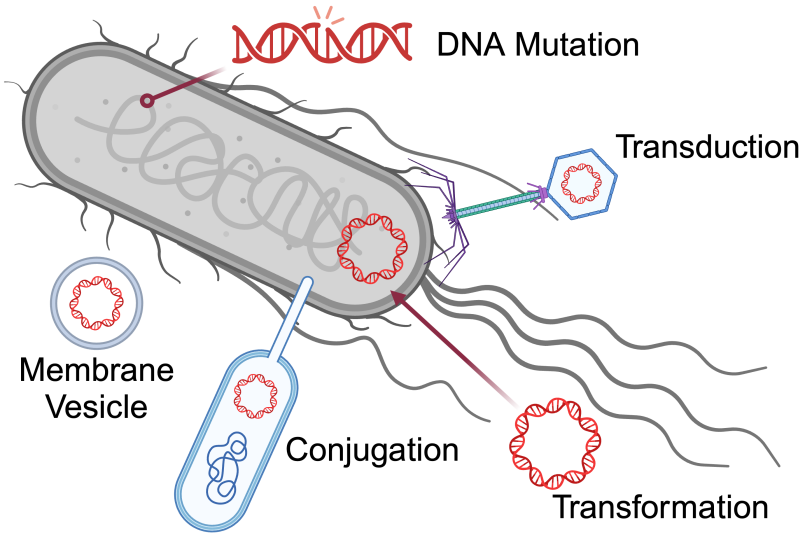


Figure 1.4: Mechanisms of acquisition of antimicrobial resistance genes. A resistance-conferring gene (red) is shown within an *E. coli* cell. The gene can be acquired through DNA mutation or can be acquired from the environment by phage-mediated transduction, transformation of naked DNA, cell-to-cell conjugation, or membrane vesicles. Figure elements not drawn to scale. Figure made with Biorender.com.

When transcribed, resistance-associated genes make antibiotics less effective through a variety of mechanisms (Figure 1.5). Ultimately, resistance-associated genes and molecular mechanisms of resistance lead to an increase in the Minimum

Inhibitory Concentration (MIC, mg/L) of the drug required to inhibit growth or kill the bacteria [48]. Bacteria can become less permeable to the molecule, through changes in their cell wall, membrane composition, periplasm, or capsule coat, resulting in lower intracellular concentrations [49, 50, 51]. In Gram-negative bacteria, β-barrel protein channels called porins allow passive diffusion of small hydrophilic compounds. Mutations that alter the structure of these porins or down-regulate their expression reduce influx of drugs into cells [52]. Changes in cell permeability can also change with cell cycle stage or growth rate, with slow-growing or non-growing cells being less permeable to antimicrobial compounds [53].

Within the cell, mutations modifying the structure of the target site can lead to lower binding efficiency. This mechanism is most important in *Mycobacteria*, where plasmid-mediated resistance is more rare, and point mutations to first-line drugs accumulate over long treatment periods [54]. Non-structural alterations to the drug target are another mechanism; for example, methylation of the ribosomal RNA has been shown to increase resistance to macrolides and aminoglycosides [55].

Bacteria can also produce compounds that inactivate the antimicrobial drug. Even before the human discovery of penicillin, β-lactamase enzymes have been produced to inactivate β-lactam compounds released by other micro-organisms. Similar inactivation enzymes have been found for tetracyclines, carbapenems, cephalosporins, and monobactams. Extended-spectrum β-lactamase enzymes are found on multiple widespread multi-drug resistance-associated plasmids, and can confer resistance to all β-lactam drugs except aztreonam [56].

Rather than modifying the drug, target bypass makes the drug ineffective by creating an alternative pathway that bypasses the original target. The canonical example is the *mecA* gene in MRSA, which produces an alternative penicillin-binding protein (PBP2a) with lower affinity for β-lactam antibiotics that maintains function for cell wall synthesis [57].

Efflux of the drug by active transport pumps is often the most effective mechanism of resistance. Over-expression or enhanced activity of efflux pumps can increase the MIC of a strain by several orders of magnitude. In the two-level

hypothesis, efflux pumps are considered the primary non-specific defence against antibiotics, and other molecular mechanisms such as target modification and drug inactivation are a second, more specific line of defence [58]. Asymmetric division of efflux pumps during cell division contributes to varying resistance within a cell population, which could contribute to the survival of sub-populations during antibiotic treatment. Multi-drug efflux pumps, especially the AcrAB-TolC and Mex pumps, contribute to intrinsic and acquired multi-drug resistance as well as the excretion of pathogenic virulence factors [59].

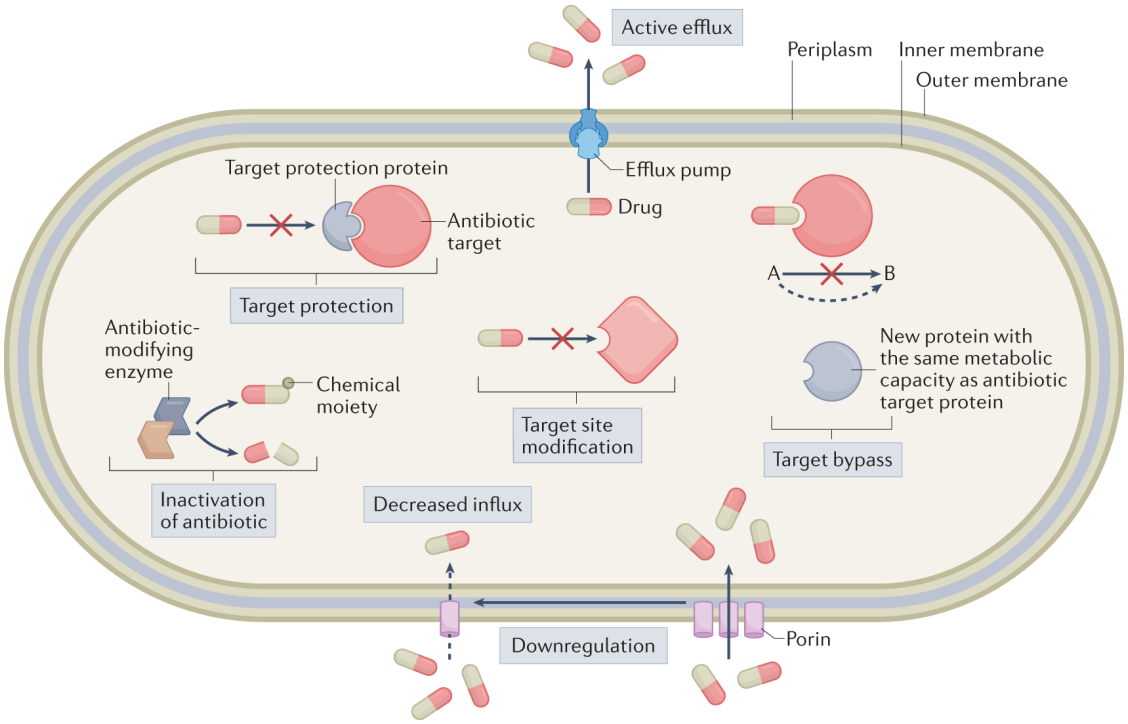


Figure 1.5: Molecular mechanisms of drug resistance. Various mechanisms of resistance to a small molecule antibiotic are represented in a Gram-negative bacterium, including inactivation of the antibiotic, target protection, down-regulation of porins, target site modification, efflux pumps, and target bypass. Figure adapted from [60].

These molecular resistance mechanisms and bulk MIC measurements are the most well understood forms of AMR, and are the most easily suited to conventional laboratory tests. However, bacteria can also adapt to antibiotic treatment without genetic changes. Aggregated populations of bacteria on surfaces, called biofilms, decrease antibiotic diffusion, increase AMR gene transfer, and can be induced by antibiotic exposure [61, 62]. Individual bacteria can also be persisters or viable

but non-culturable cells, both subsets of cells that are tolerant to stresses including high doses of antibiotics and macrophage phagocytosis, and have been linked to resistance evolution and chronic infections [63, 64, 65]. Such responses are difficult, if not impossible, to detect without single-cell diagnostic assays.

1.3 Detection of antimicrobial resistance

Antibiotic susceptibility tests (ASTs), which determine a bacterial specimen's susceptibility to a panel of drugs, are an essential tool for refining antibiotic treatment and minimising inappropriate antibiotic use. First, a biological sample is collected from the patient with concern for infection. This could be blood, urine, cerebro-spinal fluid, joint fluid, or other specimens depending on the clinical picture. Then, the specimen usually undergoes a primary culture period, between 1 and 7 days, during which it is assessed for changes in turbidity associated with bacterial growth. For slow-growing species such as *Mycobacterium tuberculosis*, this period can extend as long as 6 to 8 weeks before the lab can report a negative result [66]. It is only after growth is detected in this primary culturing period that the process of species identification and susceptibility testing can begin.

The current gold-standard tests for antibiotic susceptibility are broth micro-dilution and disk diffusion (Figure 1.6) [67]. Broth micro-dilution is performed in a well plate, allowing several antimicrobial compounds and a range of dilutions to be tested simultaneously. After each well is inoculated, the plate is incubated for a period of 4 to 24 hours and monitored periodically for bacterial growth by optical density or fluorescent indicators. This procedure can be performed manually or can be done with an automated culture system, such as the VITEK 2 (bioMérieux), MicroScan Walkaway (Dade-Behring MicroScan [68]), and the Phoenix system (BD Diagnostic Systems [69]).

In disk diffusion, antibiotic disks are placed on agar plates inoculated with a bacterial culture. After 24 hours, the susceptibility of the bacterial culture to the antibiotic is measured as the diameter of the inhibition zone where there is no bacterial growth. Gradient tests are a variation of the disk diffusion assay,

using strips with a continuous gradient of antibiotic concentrations. A commercial automated system is available for reading the results of gradient tests, the ADAGIO Automated System (Bio-Rad Laboratories). Recently, the European Committee on Antimicrobial Susceptibility Testing (EUCAST) defined rapid disk diffusion protocols for incubation periods as short as 4, 6, and 8 hours, directly from positive blood culture bottles [70]. This enables the possibility of reporting AST results in less than 24 hours, around half as long as conventional methods.

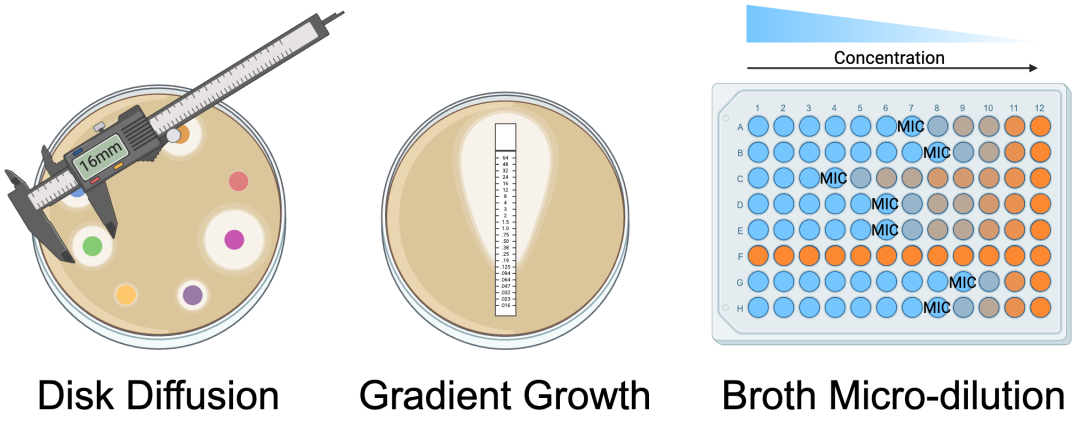


Figure 1.6: Conventional methods for antimicrobial susceptibility testing. Disk diffusion, gradient growth, and broth micro-dilution are the current gold-standard methods for ASTs. In disk diffusion and gradient growth, cultures are incubated on agar plates with variable antimicrobial concentrations as a function of drug diffusion. Results are reported as an inhibition zone diameter. In broth micro-dilution, standard inoculations of bacterial cultures are added to a range of antibiotic dilutions. After the incubation period, the minimum inhibitory concentration (MIC) is the lowest-concentration well with no growth. Figure made with Biorender.com.

Even with this acceleration, 24 hours is often too long to wait in outpatient clinics or in cases of life-threatening infection, leading clinicians to prescribe empirically and to use combinations of broad-spectrum antibiotics [71]. In clinical trials, the use of a rapid AST with a time-to-result (TTR) of less than 8 hours has been shown to decrease the use of broad-spectrum antibiotics [72, 73]. For example, in a study of 96 hospitalised patients with COVID-19 and bacteraemia, the use of a rapid AST that delivered results within 6 hours improved the mean time to optimal targeted antibiotic treatment (22.7 hours vs. 92.3 hours, $p = 0.041$) and lowered the percentage of patients receiving broad spectrum antibiotics at 72 hours after

blood collection (60.0% vs. 76.0%) [74]. In another multi-centre trial of 448 patients with Gram-negative bloodstream infections, patients in the rapid AST group with resistant pathogens were more quickly escalated to effective antibiotics (18.4 hours vs. 61.7 hours) [75]. The fast selection of appropriate antibiotic therapy is increasingly important for good clinical outcomes in a landscape of high prevalence of AMR infections. In patients with sepsis, inappropriate empiric antibiotic therapy has been associated with higher mortality for regimens that are inadequate (odds ratio = 1.19, $P = 0.02$) and those that are unnecessarily broad (odds ratio = 1.22, $P = 0.007$) [76].

Conventional methods like broth microdilution, disk diffusion, and gradient strips, which attempt to grow the bacteria in the presence of different antibiotic concentrations, are limited by the bacterial growth rate [77, 67]. Even high-throughput automated culture systems require a minimum of 4 to 24 hours for susceptibility testing after initial culture isolation (Figure 1.7) [78].

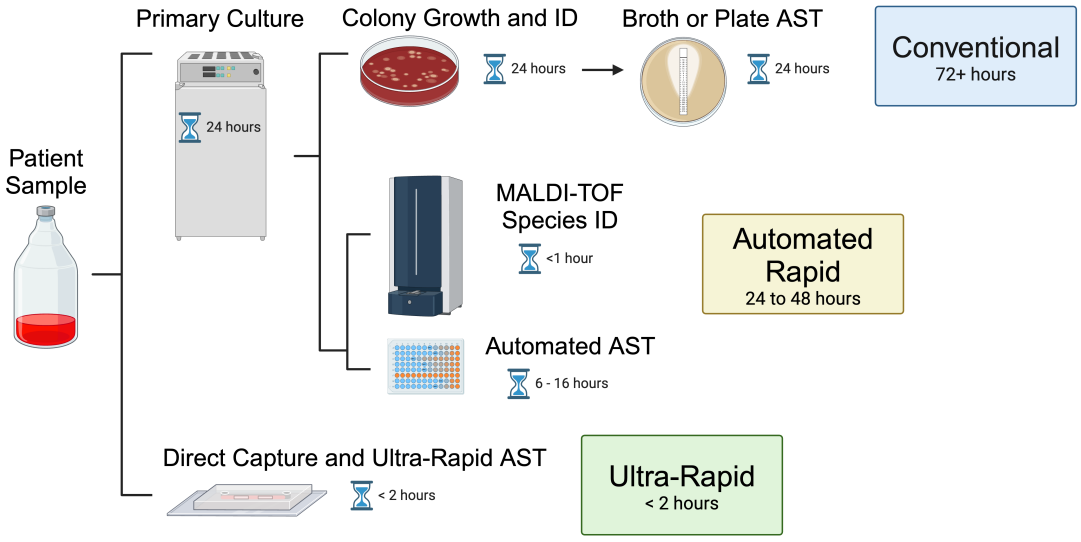


Figure 1.7: Time to result for conventional, automated rapid, and ultra-rapid ASTs. AST pathways are shown for conventional, automated rapid, and ultra-rapid methods, with time estimates for each stage of the process. Figure made with Biorender.com.

The need for faster ASTs has motivated the development of new diagnostic technologies that cut down the time-to-result [79]. Molecular methods that can detect the presence of AMR-associated proteins or genes are one way to eliminate the

need for lengthy culture steps. Matrix-assisted laser desorption/ionization time-of-flight mass spectrometry (MALDI-TOF MS) proteomic profiling is already frequently used for species identification, and methods to use it for detecting resistance-related proteins are in development and could decrease the TTR to 3 hours or less [80, 81, 82]. In DNA-based diagnostics, one commercially available platform is the bioMérieux BioFire FilmArray system that utilises multiplex polymerase chain reaction with probe kits to detect species-specific and resistance-associated genes within 1 hour for patients with infectious syndromes (e.g. respiratory, bloodstream, or joint infections). However, although genotype assays are rapid and accurate for identifying known resistance genes, the absence of AMR-associated genes does not always equate to antibiotic sensitivity and PCR kits cannot detect AMR genes outside their PCR probe set [83, 84, 85].

A rapid phenotypic test that directly measures the bacterial response to antibiotics offers advantages over genotypic assays, especially in Gram-negative species which are more likely to have polygenic and combinatorial mechanisms of resistance [86]. Phenotypic changes are still considered to be the most direct assessment of a bacterial strain's susceptibility to the effects of the antimicrobial compound. Many methods have been proposed, including using microscopy to measure growth rate [87, 88, 89, 90], structural changes [91, 92, 93], cell death [94], antibiotic accumulation [95], flow cytometry [96], Raman spectroscopy [97, 98, 99], cell impedance [100], elastic light scattering [101], and nano-motion analysis [102]. The diversity of innovation in this area is promising, and likely some of these methodologies will be more suited to clinical applications, whereas others will be useful for screening antimicrobial compounds or other research. When evaluating these systems in comparison to conventional methods like disk diffusion and broth micro-dilution, it is important to consider the complexity, supply costs, equipment costs, and laboratory expertise required for reliable testing.

The first commercially available, rapid phenotypic system was the Accelerate Diagnostics Pheno System, which was authorized by the US Food and Drug Administration (FDA) in 2017. The Pheno system combines fluorescence *in*

situ hybridisation (FISH) for species identification with monitoring of single-cell growth rates to report antibiotic susceptibility directly from positive blood cultures within 7 hours [103].

In 2024, four new commercial systems related to rapid, phenotypic ASTs were given clearance by the FDA. Selux Diagnostics received FDA clearance for the Selux PBC Separator, an automated system for processing blood samples that have been identified as positive on initial culture. In less than an hour, the Separator isolates the pathogens from the blood culture and suspends the bacteria in saline to the standard of 0.5 McFarland turbidity required for susceptibility testing. The Separator is designed to work with their FDA-cleared micro-broth dilution panels for Gram-positive species (13 antimicrobial combinations) and Gram-negative species (17 antimicrobial combinations) [104]. The average time-to-result of these panels from the positive blood cultures is 5.65 hours.

Two of the other systems are also advances in automated micro-dilution systems. The LifeScale AST from Affinity Biosensors combines micro-dilution well plates with micro-fluidic devices to measure cell mass in small volumes and predict susceptibility from positive blood cultures in 4.5 hours from the positive blood culture [105]. The ASTar system from Q-Linea is another variation on micro-dilution testing, using automated sample preparation and culture chambers on disks to test 23 antimicrobial compounds on Gram-negative samples, directly from positive blood cultures, in less than 6 hours [106].

The most novel of the FDA-cleared systems is the VITEK REVEAL, which uses small molecule, fluorescent detectors to profile metabolic gases emitted by bacterial cultures on micro-dilution well plates. This is unique compared to the other systems, which rely on cell growth rate or mass data rather than metabolic phenotypes. The REVEAL system is currently limited to Gram-negative bloodstream infections and claims to return results in around 5.5 hours, but its clinical trial data is currently proprietary [107].

These four systems represent the start of a new era in rapid AST technology, as increased investment in research and development for AMR begins to yield results

for clinical microbiology laboratories. The FDA clearance of these systems should also encourage other companies who are pursuing AST development. However, although the systems shorten the time-to-result compared to standard micro-dilution assays, and have an extended panel of antimicrobial compounds and concentrations, they still rely on the initial positive blood culture, which introduces a significant time delay of around 24 hours. For this reason, the average time-to-result from the time the sample was obtained was 28 hours and 49 minutes for the ASTar system (Q-Linea) compared to 44 hours and 18 minutes for the standard procedure [106]. The authors of the ASTar clinical trial noted that it is unclear whether patients will benefit from a change in antibiotic therapy 29 hours after blood culture sampling.

A more significant decrease in time-to-result will come when susceptibility testing can be conducted directly on bacteria captured from the patient's sample, without needing to wait 12 to 24 hours for the primary culture. The challenge of bacterial capture depends on the biological specimen. In blood, the presence of $\leq$100 colony-forming units (CFUs) per mL is sufficient to cause septicaemia in the adult patient, whereas $\geq$ 1,000 CFU/mL of urine is common for urinary tract infections [108, 109]. In other specimens such as cerebro-spinal fluid or joint fluid, bacterial concentrations can vary significantly depending on the pathogen and mechanism of infection [110]. The high prevalence of urinary tract infections (around 400 million cases per year), and the high CFU/mL and clean sample matrix (low concentrations of human cells and proteins), make urine samples a natural candidate for trying to eliminate the primary culture step.

Indeed, in June 2024 the PA-100 system by Astrego won the Longitude Prize on AMR for their "real-time" urine AST. This system is built on their rapid test for antimicrobial resistance [87], which uses a microfluidic device to capture bacteria from urine in cell traps. With these microfluidic traps, they can positively or negatively identify the presence of infection within 15 minutes, by the fraction of traps occupied by bacteria. Then, the trapped cells are exposed to various antibiotic treatments while their growth rates are measured. This enables susceptible and resistant pathogens to be differentiated within 45 minutes of loading the urine

sample. Because this system relies on growth rate, it can be used without additional labelling reagents, which significantly decreases the complexity of the assay, an important advantage for its use in outpatient practice. However, the phase contrast microscopy system does not identify the bacterial species causing the infection.

For bloodstream infections, another group has used magnetic nanoparticles coated with synthetic β - 2 - glycoprotein I peptides to capture *E. coli* and *S. aureus* with >90% efficiency at concentrations of 1,000 CFU/mL in blood samples, without primary cultures [90]. From these nanoparticle-captured bacteria, around 500 CFU per well are "micro-cultured" with or without antibiotics for 30 to 40 minutes, during which the colony area is monitored for growth on a μm^2 scale. This micro-scale growth agrees with standard blood culture AST measurements, and this technology could become an ultra-rapid test for bloodstream infections in the future.

ASTs that can return a result on this time scale, within 1 to 2 hours, have the power to change clinical practice significantly because, in many cases, clinicians could wait to prescribe tailored antibiotic therapy with knowledge of the susceptibility profile. These ultra-rapid tests will enable rational and targeted antibiotic conservation.

1.4 Bacterial antibiotic response phenotyping

Visually apparent changes to the intracellular structure of bacterial cells can also be used to measure the bacterial antibiotic response. As antibiotics disrupt normal cellular physiology, long-recognised and characteristic antibiotic response phenotypes develop, which have recently been profiled at scale with high-content imaging [111, 112]. These changes in cell size, shape, and DNA morphology have been observed for a range of antibiotics in a number of Gram-positive and Gram-negative species, including *E. coli* [113], *Klebsiella pneumoniae* [111], *Salmonella enterica* serovar Typhimurium [111], *Staphylococcus aureus* [111], *Acinetobacter baumanii* [112], and *Bacillus subtilis* [114].

The characteristic single-cell phenotypes can be used to identify the cellular pathway affected by the antimicrobial, for example synthesis of DNA, protein, RNA, peptidoglycan, or lipid [113]. This property makes phenotype identification a

promising tool for characterising novel antimicrobial compounds. In fact, it has been used with machine learning to identify promising antimicrobial drug candidates [115], to screen compounds to potentiate tuberculosis therapy [116], and to identify antibiotic-resistant bacterial strains isolated from clinical samples [93, 117].

1.5 Machine learning for antimicrobial resistance detection

Machine learning is well-suited for challenges like predicting antimicrobial resistance or screening antimicrobial compounds, due to its ability to make predictions on datasets with complex patterns and nuanced features, without prior knowledge of the exact underlying mechanisms [118]. Compared to statistics, machine learning methods are designed to automatically improve through experience [119]. Their capacity for "data-guided pattern learning" makes them effective for making predictions or classifications about poorly understood phenomena, especially when there are hundreds or thousands of features per observation [120]. For this reason, machine learning has been used on large genome sequence datasets to tackle the genotye-to-phenotype prediction challenge in *Salmonella*, *Mycobacterium tuberculosis*, and other species [121, 122, 123] and machine learning models outperformed conventional methods when predicting antibiotic susceptibility from whole-genome sequencing [124]. Deep learning has also been applied to predict gene regulation and RNA abundance from genomic data, a key problem in the utilisation of genotypic ASTs [125].

For most of these numerical datasets, conventional machine learning methods such as logistic regression, support vector machines, and random forest methods are appropriate. However, for these methods to be used on imaging data, quantifiable features must be extracted, suitable for the task. With deep learning methods, the model learns its own layered representations of the image features. In a supervised learning task, such as learning to classify a strain's antimicrobial susceptibility, this is done by an iterative process called training. The model is shown an image and it returns a vector of scores, predicting the category of the image. The error in

these predictions is calculated and fed back to the machine, which modifies the weights of the parameters it uses to make classifications by computing a gradient vector. The gradient vector is a partial derivative of the loss with respect to each parameter, which is updated as:

$$w := w - \alpha * \frac{\partial L}{\partial w}$$

where w is the weight of the parameter, α is the learning rate, and L is the loss function. This process is repeated, showing the model thousands of images while it adjusts all of its weights, until the error in the model's predictions stops decreasing each time.

A convolutional neural network (CNN) is an interconnected architecture of many of these layers, organised into feature maps, which allow the model to detect local motifs within arrayed data, such as images. This enables the model to learn important, recurring motifs that form structures within the image, while learning to ignore features irrelevant to the task [126]. The power of these models to learn nuanced features and find unexpected correlations within datasets means that care must be taken to avoid over-fitting to the training data and to test the model on unseen data.

When combined with large datasets and graphical processors, many applications of CNNs have been revolutionary in their fields. In one of the ground-breaking examples of the power of a CNN for image recognition, a model with 60 million parameters and 5 convolutional layers learned to classify 1.3 million images in the ImageNet dataset into 1,000 different classes like "container ship," "leopard," and "mushroom" despite variations in lighting, position, and surrounding objects [127]. Increasing the depth of the model has been found to increase performance, although this should be balanced with the size of the training dataset and the complexity of the task [128].

CNNs have been widely applied in other areas of diagnostics, such as detecting lesions in low-dose computed tomography (CT) scans [129] and in mammograms [130], or from a large dataset of medical images [131]. In single-cell data, a

multi-scale CNN was used to classify cancer cells from pleural fluid with improved performance over classical methods [132]. High-content imaging of eukaryotic cell phenotypes has also been used to train a CNN to predict the mechanism of action of compounds [133]. Dozens of publicly available datasets are available for training and optimising these CNN-aided diagnostic techniques. These successes in biomedical image classification suggest that deep learning, combined with a rich dataset of bacterial antibiotic treatment phenotypes, could be used to train a powerful predictor of antimicrobial susceptibility.

1.6 Towards a rapid phenotypic test

A rapid, point-of-care AST is needed more than ever to conserve antibiotics and treat infections effectively. A single-cell phenotypic assay has several advantages, especially in testing directly from patient samples, detecting intra-sample heterogeneity, and avoiding genotype-phenotype mismatch. A mobile imaging platform could be combined with a CNN to classify bacteria from clinical samples based on their antibiotic response. Given the expected diversity of bacterial strains and antibiotic response phenotypes, a CNN-based assay has the unique advantage that its performance would be expected to improve when it is able to learn from additional, real-world data [134].

This concept was demonstrated on *E. coli* isolated from clinical samples, showing that trained CNNs could recognise antibiotic effects on the nucleoid and cell membrane within 30 minutes, and that these phenotypic changes correlate with MICs as measured by conventional ASTs [93]. This approach had high single-cell accuracy (>84%), which would lead to very high prediction confidence in classifying a whole sample. Another advantage of this method is that by classifying individual bacteria, such a test could be performed directly on cells isolated from a clinical sample, eliminating the need for pre-culturing, and potentially reducing the time-to-result to less than 1 hour.

1.7 Thesis outline and aims

Building on the promising developments in single-cell imaging and machine learning predictions of antimicrobial susceptibility, the aim of this thesis is to investigate methods for single-cell phenotyping for the detection of antimicrobial resistance. In Chapter 1, we introduce the challenge of antimicrobial resistance and the current state of antibiotic susceptibility testing, and demonstrate why machine learning is useful for antibiotic susceptibility testing. In Chapter 2, we detail the methods of our single-cell ribosome phenotype assay. In Chapter 3, we show that fluorescent ribosome probes can be used to visualise the intracellular antibiotic response, that these can be detected by a CNN, and consider the benefits of using ribosome phenotypes over other methods of phenotype detection. In Chapter 4, we explore the accuracy of a CNN-based ribosome phenotype classifier for detecting AMR in *E. coli* isolated from clinical samples. In Chapter 5, we investigate which image features are associated with misclassification of *E. coli* from clinical samples using data from the citizen science project Infection Inspection. Finally, in Chapter 6, we interpret our findings and discuss the potential for future developments in rapid, phenotypic ASTs.

2

Development of a single-cell ribosome phenotype assay

Contents

2.1 The ribosome phenotype 24
2.2 *Escherichia coli* as a model organism 26
2.3 Experimental strains 27
2.4 Antibiotics 27
2.5 Sample preparation 29
2.5.1 Antibiotic treatment and fixation 29
2.5.2 FISH 31
2.6 Image acquisition 31
2.7 Image processing and segmentation 32
2.8 Neural network training 32
2.9 Accuracy metrics 33
2.10 Cell phenotype statistics 33
2.11 Sample size simulation 34
2.12 Discussion 35
2.13 Conclusions 36

In this chapter, we detail the methodology used in our single-cell ribosome phenotype assay. We also discuss why the ribosome phenotype is useful as a diagnostic tool, and why *E. coli* was chosen as our model organism. Our goal was to show how antibiotic treatment results in characteristic perturbations of the internal structure of *E. coli*, using standard microbiology protocols so that

our results could be compared to previous work. We grew bacterial cultures to exponential phase in rich media, treated them with antibiotic concentrations defined by the EUCAST breakpoint, and then fixed and permeabilised the cells so that they could be stored for later use. From those samples, we labelled the DNA with a fluorescent dye and the ribosomes with fluorescent *in situ* hybridisation (FISH) probes that bind to the ribosomal RNA. These labels allowed us to visualise the DNA and ribosome phenotypes of single cells with fluorescence microscopy. These images were segmented to extract single-cell images, which were used to train convolutional neural networks and extract statistics about the antibiotic phenotypes. From our accuracy data, we were able to estimate the sample size necessary for a diagnostic test using our assay.

2.1 The ribosome phenotype

Escherichia coli have been described as a "finely tuned machine for building ribosomes" [135]. Ribosomes, the protein-manufacturing machinery of the cell, are composed of three ribosomal ribonucleotide molecules (rRNA) and 54 ribosomal proteins [136, 137, 138, 139]. Within the cell, ribosomes exist as 30S and 50S ribosomal sub-units, 70S monosomes, and polysomes of multiple ribosomes bound to and translating mRNA into proteins. The prevalence of each of these species depends on nutritional availability, antibiotic treatment, and growth rate [140]. In typical growth conditions, about 80% of ribosomes exist as polysomes [141]. The role of ribosomes in actualising gene expression makes their function essential to nearly every cellular process. In foundational microbiology experiments, bacterial growth rates in different media were shown to be directly related to the proportion of ribosomes to cellular protein content [142]. *E. coli* cultures can double their population as frequently as every 22 minutes in ideal growth conditions [143]. In these conditions, there are as many as 72,000 ribosomes per cell [141]. This makes ribosomes a promising labelling target for visualising intracellular changes.

In addition to their large numbers, ribosomes also have several biophysical properties that make them useful targets. The spatial distribution of ribosomes

within the cell results from their biological role and from entropy and excluded volume effects. Super-resolution imaging and simulations have shown that 70S polysomes are largely excluded from the nucleoid, whereas the smaller populations of 30S and 50S sub-units diffuse into DNA-rich regions [144, 145, 146]. Therefore, ribosome and DNA density anti-correlate and the ribosome distribution can report changes in DNA morphology because ribosome intensity is lowest in regions with greater DNA density. This has been established in bacterial strains with fluorescently tagged ribosomal protein S2 [145, 147]. Entropic space-filling effects also push ribosomes, especially polysomes, to the cell periphery [148]. This means that the ribosome distribution can be used to outline or segment the cell without needing a membrane stain or label.

For these reasons, we hypothesised that the spatial distribution of ribosomes within the cell could function as an inverse DNA stain for the purpose of single-cell phenotyping of the antibiotic response. We call this distribution the ribosome phenotype.

We use fluorescence *in situ* hybridisation (FISH) to label and visualise the ribosome phenotype. FISH utilises the property of a fluoresecently labelled single-stranded DNA probe to anneal to a complementary strand of DNA or RNA [149, 150]. This oligonucleotide FISH probe sequence is designed to optimise its binding energy and specificity [151]. FISH protocols have been used for decades in applications such as human chromosomal diagnostics, transcriptomics, and bacterial species ID [152, 153, 154, 155, 156]. While protocols vary depending on the target sample, usually FISH protocols involve a fixation step to preserve the cellular morphology and a permeabilisation step to allow the probes to access their intracellular targets [151].

FISH probes can be used for single-colour, multi-colour, or spectral imaging, and for bulk fluorescence or single-molecule fluorescence studies [157]. The probes can be synthesised with fluorescent dyes conjugated to their 3' end, 5' end, or both, or they can be designed as a multiplexed probe set with primary and secondary probes. With multiplexed FISH, barcoding and combinatorial methods have enabled the detection of $\mathcal{O}(10^3)$ targets with $\mathcal{O}(10^1)$ probes [158, 159].

In the context of antibiotic-susceptibility testing, combinatorial FISH has been used to identify species-specific ribosomal RNA sequences for the purpose of species ID from clinical samples captured in a micro-fluidic device [160, 161]. Because species-specific FISH probes can be used, the ribosome image can serve for both species ID and AST phenotyping. The ability to provide clinicians with the bacterial species and phenotypic antibiotic susceptibility, within a single rapid assay, is one advantage of the ribosome phenotyping method.

2.2 *Escherichia coli* as a model organism

Escherichia coli is a widely prevalent, facultative anaerobic bacillus, commonly found in the mammalian gut microbiome. The ability to culture hardy, non-pathogenic strains from human samples has made it a popular model organism, ever since its use in early investigations of bacterial physiology and genetics. Over the past hundred years, it has become "the best understood organism on the planet" [162]. *E. coli* was used in several fundamental studies of intracellular ribosome dynamics and the development of a ribosome phenotype during antibiotic exposure, to which we can compare our results [144, 146]. *E. coli* is also a clinically important pathogen. In addition to its role as a commensal species, several pathotypes of *E. coli* cause disease in humans, including at least six that cause enteric disease, and others that cause meningitis, bloodstream infections, and urinary tract infections [163] (Figure 2.1). *E. coli* is a leading cause of bloodstream infections and sepsis-related deaths [7, 8]. Therefore, *E. coli* was chosen as the model organism for this study because of its clinical and scientific relevance.

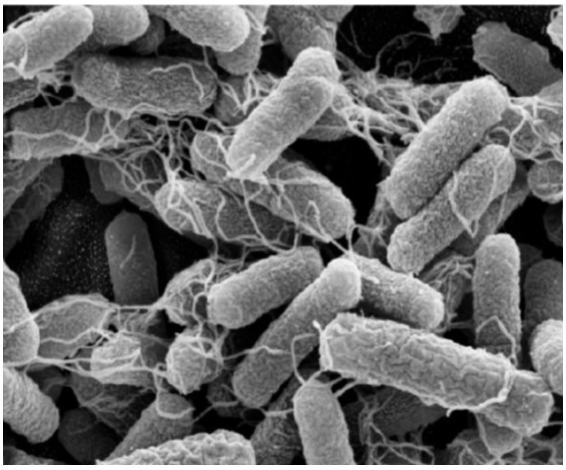


Figure 2.1: A sample of a pathogenic *E. coli* biofilm, up close. Scanning electron micrograph of the enteropathogenic *E. coli* O119:HND, strain A111, grown on a glass coverslip. Image credit: [164].

2.3 Experimental strains

Escherichia coli MG1655, a lab-adapted non-pathogenic K-12 derivative, served as the reference strain for methodological development and for characterising antibiotic-susceptible ribosome phenotypes in Chapter 3.

The *E. coli* clinical isolates used in this work were obtained from positive blood culture specimens obtained by the Microbiology Laboratory of the Oxford University Hospitals NHS Foundation Trust for diagnostic purposes. These isolates were anonymised and cleared for use in this research. For a previous study, each of the clinical isolates had been sequenced on the Illumina platform and AMR genotypes were assigned using the ResFinder [165] database with Abricate v0.9.8 [166] (–min-id 95 –min-cov 95) [167].

2.4 Antibiotics

The antibiotics were chosen to represent antibiotic classes with different mechanisms of action: ciprofloxacin, a fluoroquinolone, targeting DNA gyrase/topoisomerases and DNA replication; gentamicin, an aminoglycoside, targeting the 30S ribosomal subunit and translation; chloramphenicol, an amphenicol, targeting the 50S ribosomal subunit and translation elongation; and carbenicillin, a penicillin,

targeting peptidoglycan synthesis (Table 1.1). For each of the antibiotics, the minimum inhibitory concentration (MIC) of each strain was tested in biological duplicate according to the broth micro-dilution method (Table 2.1) [168]. The MIC was defined as the lowest antibiotic concentration inhibiting growth when the cultures were incubated overnight in Mueller-Hinton broth at 37°C, as is standard in clinical microbiology labs.

	MIC (mg/L)			
	Cip	**Gent**	**Cam**	**Carb***
***E. coli* MG1655**	0.016	0.125	3.75	20
EUCAST Breakout Point	>0.5	>2	>8	>8
EUCAST:MIC Ratio	31.25	16	2.13	0.4

Table 2.1: MG1655 MICs vs. EUCAST breakout points for ciprofloxacin, gentamicin, chloramphenicol, and carbenicillin. The minimum inhibitory concentration (mg/L) for *E. coli* MG1655 is shown for ciprofloxacin, gentamicin, chloramphenicol, and carbenicillin (Cip, Gent, Cam, Carb). These concentrations were used to determine antibiotic treatment concentrations, e.g. 1× EUCAST = 8 mg/L chloramphenicol, 1× EUCAST = 0.5 mg/L ciprofloxacin, 20× EUCAST = 40 mg/L gentamicin, 3× EUCAST = 24 mg/L carbenicillin. MICs were determined by broth micro-dilution. The ratio refers to the EUCAST Breakout Point divided by the MG1655 MIC for each antibiotic. *NB: Carbenicillin is a β-lactam without a EUCAST breakpoint. Its mechanism of action is closely related to ampicillin and other penicillins, with the main benefit of increased stability, but higher cost. Here we reference the EUCAST breakout point for penicillins without β-lactamase inhibitors. Sammi Ta provided the carbenicillin MIC data.

The antibiotic treatment concentrations were chosen as a multiple of the European Committee on Antimicrobial Susceptibility Testing (EUCAST) breakpoint [169] for *Enterobacterales* including *E. coli*, so that an empirical benchmark could be applied to other strains. The EUCAST breakpoint is a defined concentration of antibiotic used to classify a microorganism as antibiotic susceptible and resistant, accounting for clinical factors including antibiotic dosage, target infections, pharmacokinetics, and resistance mechanisms. The MIC of a bacterial isolate can be compared to the EUCAST breakpoint for a given antibiotic and bacterial species to classify it as clinically susceptible (S) or resistant (R). The antibiotic concentration should be optimised depending on the antibiotic's mechanism of action to induce differentiable phenotypes in susceptible and resistant strains within the treatment period.

To demonstrate our methods on clinical isolates, we narrowed our focus to ciprofloxacin-susceptible and resistant strains as a proof-of-concept. To represent a sub-section of pathogenic *E. coli*, we chose three susceptible strains (S1, S2, S3) and four resistant strains (R1, R2, R3, R4), each with a different mutation in ciprofloxacin resistance-associated genes (Table 2.2). We called an isolate "resistant" if its MIC was above the EUCAST breakpoint [169]. Isolates with MICs below the EUCAST breakpoint were called "susceptible."

		Relevant Genotype Information			
Clinical Isolate	**MIC (mg/L)**	***gyrA***	***marR***	***parC***	***parE***
Susceptible 1 (S1)	0.015				I355T
Susceptible 2 (S2)	0.25		S3N		
Susceptible 3 (S3)	0.25				I529L
Resistant 1 (R1)	2	D87Y, S83L		S80I	
Resistant 2 (R2)	16	D87N, S83L		E84V, S80I	I529L
Resistant 3 (R3)	64	D87N, S83L		S80I	S458A
Resistant 4 (R4)	128	D87N, S83L		S80I	S458A

Table 2.2: *E. coli* clinical isolates with their ciprofloxacin MICs and AMR genotypes. Each clinical isolate used in Chapter 4 is listed with its ciprofloxacin MIC and ciprofloxacin-relevant genotype information. All strains are *Escherichia coli* isolated from bloodstream infections in the United Kingdom, obtained and whole-genome sequenced for a previous study [170]. MICs were determined by broth micro-dilution as described in [168].

2.5 Sample preparation

2.5.1 Antibiotic treatment and fixation

From the start of antibiotic treatment to the start of imaging, this protocol can be completed in less than 2 hours (approximately 1 hour and 45 minutes). This time is well within the standard for rapid ASTs, allowing a time-to-result that could influence clinical practice within an 8-hour working shift [171]. The protocol for sample preparation was the same for *E. coli* MG1655 and clinical isolates, with the exception that MG1655 was cultured in lysogeny broth and clinical isolates were cultured in Mueller-Hinton broth for consistency with standard clinical laboratory protocols.

Bacterial strains were streaked from frozen stocks on lysogeny broth (MG1655) or Mueller-Hinton (clinical isolates) agar plates and incubated at 37°C until visible

colonies appeared. A single colony from the plate was used to inoculate 5 mL lysogeny broth, which was cultured overnight at 37°C in a shaking incubator (Figure 2.2, Panel 1). The next day, the overnight culture was diluted 1:1000 in 5 mL lysogeny broth (MG1655) or Mueller-Hinton (clinical isolates) and incubated at 37°C, shaking, until reaching early exponential growth, or OD_{600nm} of 0.2, approximately 1.5 hours (Figure 2.2, Panel 2).

After the fresh culture reached the exponential growth phase, antibiotics were added to reach the specified concentration as a multiple of the EUCAST breakout point (Table 2.1) and the samples were returned to the incubator for the 30-minute treatment period (Figure 2.2, Panel 3). Samples were then fixed in 2% paraformaldehyde for 20 minutes (Figure 2.2, Panel 4). After fixation, the samples were split into 1 mL aliquots and centrifuged for 3 minutes (Figure 2.2, Panel 5). The cell pellets were washed once with 1 mL phosphate-buffered saline (PBS), then re-centrifuged (3 minutes) and re-suspended in 1 mL PBS before being centrifuged once more (3 minutes) and re-suspended in 500 μL absolute ethanol (Figure 2.2, Panel 6). The samples were mixed on the nutator for 20 minutes in the ethanol for permeabilisation. The samples were stored at -20°C until use.

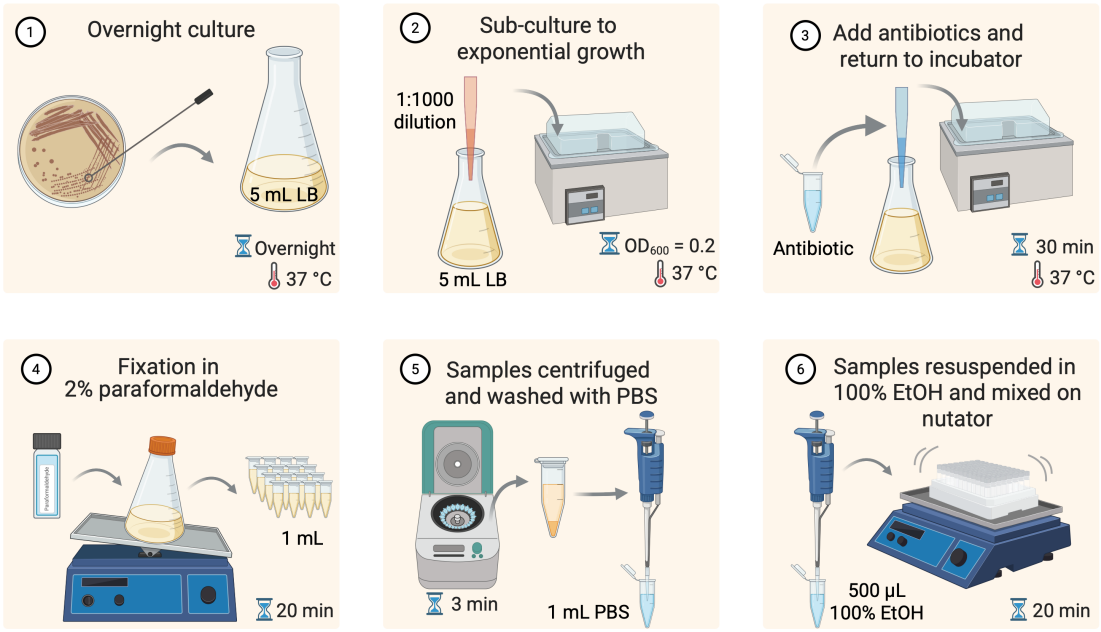


Figure 2.2: Illustrated protocol for bacterial culture, antibiotic treatment, fixation, and permeabilisation. Figure made with Biorender.com.

2.5.2 FISH

Before imaging, the cells were centrifuged to remove the ethanol supernatant (3 minutes), washed with 500 μL PBS, and re-suspended in hybridisation buffer (20% v/v formamide, 0.9 M NaCl, 20 mM Tris pH 7.5, 0.01% w/v sodium dodecyl sulfate). For labelling, 4',6-diamidino-2-phenylindole (DAPI, 1 μg/mL) and 25 nM EUB338-Cyanine3 were added to the solution and the sample was incubated for 20 minutes at room temperature. The ssDNA EUB338-Cyanine3 FISH probe has the sequence Cyanine3 — 5' — gct gcc tcc cgt agg agt — 3' (Sigma Aldrich). The clinical isolate samples were labelled with EUB338-Cy3 FISH probes but not with DAPI because only the ribosome phenotype was required for the work in Chapter 4. Following 20 minutes of incubation, the samples were centrifuged (3 minutes), washed, and re-suspended in 150 μL PBS.

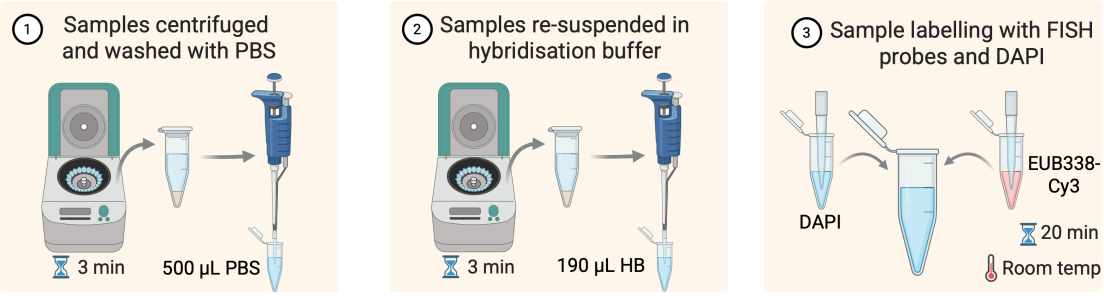


Figure 2.3: Illustrated protocol for cell staining and rRNA FISH. Figure made with Biorender.com.

2.6 Image acquisition

Samples were imaged on agarose pads prepared with 1.5% (w/v) high-purity agarose in distilled water. Images were collected on the Nanoimager-S microscope (ONI, Oxford, UK) with a 100× oil-immersion objective in multi-acquisition mode. Multi-acquisition mode allowed hundreds of fields of view to be collected from a single sample without selection bias. The DAPI stain was illuminated by a 405 nm laser in epifluorescence mode at a laser power of 5.1 kW/cm^2 for an acquisition time of 20 ms. The Cyanine3 fluorophore was illuminated with a 532 nm laser in epifluorescence mode at a laser power of 16.5 kW/cm^2 for 20 ms.

2.7 Image processing and segmentation

Each field of view was segmented using a custom-trained model in Napari-BacSeg [172]. This custom CellPose [173] model was trained on a sub-section of hand-segmented 532 nm ribosome fluorescence images for 100 epochs using the standard Napari-BacSeg hyperparameters to improve segmentation performance. After training on these hand-segmented images, the models showed robust performance and the segmentations were used without further editing. The standard CellPose model initially split cells with two well-defined, ribosome-rich regions, such as those treated with ciprofloxacin. However, after curated data from these samples was fed forward into the model, performance on these images became similar to those from other samples. Partial cells cut off by the edge of the image, overlapping cells, vertical cells, or cells outside of the focal plane were removed from the final dataset during the segmentation and curation process.

2.8 Neural network training

Images and segmentation maps were exported from Napari-BacSeg [172] to create standardised 64×64 images of each cell zero-filled outside the segmentation boundary. To account for different staining and illumination brightness across experiments, histogram normalisation was applied to each single-cell image. The images were randomly rotated, flipped, translated, and altered with geometric and noise transforms from the Albumentations package [174] before being loaded into the neural network. The augmentation strategy is detailed in *Strategies to maximise transferable learning* (3.3).

The convolutional neural network was built with PyTorch [175]. Each model was run for 100 epochs with a training-validation split of 0.80/0.20. Optuna [176] was used to optimise the learning rate and batch size for each model. The validation and testing datasets were balanced by class to reduce bias. Because the task was a single-label binary classification, EfficientNet B0 [177] with the Cross-Entropy loss function was used as a lean and high-performing model architecture. The Adam

function [178] was used for adaptive learning rates. For each model, we verified that the model converged by plotting the training and validation accuracies by epoch. We also validated the model's attention to different parts of the image by plotting the gradient of each pixel in the final layer of the CNN [179].

2.9 Accuracy metrics

All accuracies are reported with the 95% confidence interval, $\pm 2\sigma$. Here we define cells showing the antibiotic phenotype as "Positives" and untreated cells as "Negatives." For clinical isolates, we refer to antibiotic-resistant cells as "Positives" and antibiotic-susceptible cells as "Negatives." For a one-class binary classification task:

$$Accuracy = \frac{True\ Positives}{True\ Positives + False\ Negatives}$$

For a two-class classification task, given that:

$$Sensitivity = \frac{True\ Positives}{True\ Positives + False\ Negatives}$$

and

$$Specificity = \frac{True\ Negatives}{True\ Negatives + False\ Positives}$$

then

$$Balanced\ Accuracy = \frac{1}{2}(Sensitivity + Specificity)$$

The predictive value of such classifications can be calculated as:

$$Positive\ Predictive\ Value = \frac{True\ Positives}{Number\ of\ Cells\ Classified\ Positive}$$

$$Negative\ Predictive\ Value = \frac{True\ Negatives}{Number\ of\ Cells\ Classified\ Negative}$$

2.10 Cell phenotype statistics

From the curated single-cell segmentations, cell lengths, widths, and midlines were generated using the ColiCoords [180] plugin within Napari-BacSeg [172] with 10 midline vertices. For each cell, the intensity in each channel was normalised from 0 to 1 and the mean intensity was calculated for 100 bins along the midline. For a

population of cells, the mean intensity and the standard deviation of the intensity was calculated for each of those 100 bins. This provided a mean intensity for each bin along the long axis of the cell. Because the cell length and width distributions are composed of cells at different stages of the cell cycle, and thus could not be assumed to be normally distributed, hypothesis testing was conducted with the Mann-Whitney non-parametric test with a significance value of 0.05 in Origin Pro 2021.

2.11 Sample size simulation

The cell classification simulation to determine the minimum sample size for a diagnostic test was calculated using MATLAB R2022b. Given a sample of cells of size N and a certain proportion of susceptible and resistant cells ($S : R = 100 : 0, 50 : 50, 0 : 100$), we simulated the measured fraction of resistant cells ($Detected_{Resistant}/Detected_{Total}$) and the 95% confidence interval of this measurement. To do this, we simulated random samples of 10,000 susceptible and resistant cells at the defined proportions and transformed them into detected samples using the accuracy of our susceptible-resistant classifier. The proportion of resistant cells detected as resistant was defined by the sensitivity, such that:

$$P(R|R) = Sensitivity = 0.993$$

$$P(S|R) = 1 - Sensitivity = 0.007$$

The proportion of susceptible cells detected as susceptible was defined by the specificity, such that:

$$P(S|S) = Specificity = 0.905$$

$$P(R|S) = 1 - Specificity = 0.095$$

Random populations of between 1 and 40 cells were sampled using the `randsample` function in MATLAB R2022b and the Resistant Fraction was calculated for each population. After 1,000 trials, the mean Resistant Fraction for each population size and the 99% confidence interval (2.58σ) was plotted for each sample size. We defined the minimum sample size as the smallest sample for which the 99% confidence intervals did not overlap.

2.12 Discussion

The experimental methods chosen in the development of our single-cell phenotyping assay influence the phenotypes we observe.

We chose to segment single cells to separate the intracellular phenotype data from the imaging background. This forces the model to learn phenotype features with less potential confounding information from the background, and could be useful for imaging in dirtier environments in the future, such as directly in a biological sample medium or microfluidics device. However, it forces the segmentation model to classify a cell in the process of cell division as either one or two cells. This decision by the image segmenter can significantly influence the appearance of the final images. If they are not sufficiently represented in the training data, cells in the process of cell division have a different appearance that could be confusing to the classifier.

Because we use cells grown in bulk culture, our cells are at different growth stages when exposed to antibiotic treatment. Some of the cells may also be in non-viable or stationary phases. An overnight *E. coli* culture in lysogeny broth can contain approximately 50% cells in a non-viable phase [181], depending on culture conditions. After the approximately 4 doubling times, this results in a 16:1 viable:non-viable ratio, or approximately 6% non-viable cells, if none of the non-viable cells recover during the subculture phase. These cells are likely to respond differently to antibiotic treatment than rapidly growing cells and may develop different phenotypes.

In addition, our decision to treat with antibiotics during rapid growth in rich media makes our phenotype results more comparable to prior work that treated cells in similar conditions [146, 182, 161], but the growth conditions are less similar to those expected in biological media directly from clinical samples, such as blood or urine. Measured in the *ori*:*ter* ratio, the growth rate of *E. coli* in human urine culture samples has been shown to range from approximately 1.1 to 2.2, in the absence of antibiotics, whereas the growth rate for the same strains in lysogeny broth is around 3 [183]. In slower growth conditions, *E. coli* has been shown to have more diffuse DNA regions [184] and less ribosome-nucleoid segregation, as more of the ribosome particles are dissociated into 30S and 50S subunits, which diffuse into

the nucleoid-rich regions [146, 145]. Therefore, the antibiotic phenotypes developed with our protocol are likely to be more defined than if we had used cells growing at a slower rate. In future work, it would be interesting to explore the classification accuracy of single-cell phenotypes from cells treated at slower growth rates, for example in clinical sample media such as blood, urine, or cerebrospinal fluid. These results are likely to be more similar to what we would find in bacteria collected directly from a clinical sample. It would also be possible to allow cells collected from a clinical sample to recover in a rich medium, just as our overnight culture is allowed to reach exponential growth prior to antibiotic treatment. This additional step would add time to our protocol but could result in more well-defined phenotypes than treating slow-growing cells with antibiotics and increase the accuracy of our diagnostic test. Alternatively, a model could be trained on data from a variety of treatment conditions, which might improve its robustness to different growth rates.

We chose to use agarose pads to immobilise our cells, rather than immobilising them directly on glass slides or using a microfluidics device. There are a variety of ways to immobilise *E. coli* on glass slides for single-cell fluorescence imaging, including coating the glass with compounds, such as poly-L-lysine or chitosan, that form electrostatic interactions with the negatively-charged cell membranes, or on gel pads made from agarose. These methods all result in similar cell "flatness" measurements on a stationary microscope stage [185]. We tested imaging methods using chitosan-coated slides and microfluidics chambers coated with poly-L-lysine (µ-Slide VI 0.4, Ibidi). Ultimately, we chose to image our samples on agarose pads because the movement of our microscope stage during multi-acquisition sessions led to a higher proportion of blurred and floating cells when using chitosan-coated slides. A disadvantage of choosing agarose pads for this study is that it is less similar to what might ultimately be used in a microfluidics device in a diagnostic setting.

2.13 Conclusions

In this chapter, we discussed the suitability of ribosomes as a cell phenotyping target and why *Escherichia coli* makes an ideal model organism for the development of our

assay, due to its well understood biology and clinical relevance as one of the leading causes of sepsis-related deaths. We then showed how *E. coli* were cultured, treated with antibiotics, and labelled with dyes and FISH probes to visualise their DNA and ribosome morphologies on a fluorescence microscope. These data were used to train neural networks to segment single cells and classify *E. coli* by their ribosome phenotypes. The results of these experiments are discussed in Chapters 3 and 4.

3

Ribosome phenotypes for detecting antibiotic response

Contents

3.1 Introduction 39
3.2 Characterisation of the *E. coli* antibiotic response by ribosome distribution 40
3.3 Strategies to maximise transferable learning 46
3.4 Identification of antibiotic-susceptible ribosome phenotypes by a neural network 50
3.5 Discussion 54
3.5.1 Strains and antibiotics 55
3.5.2 Neural network training and augmentations 57
3.6 Conclusions 59

In the previous chapter, we discussed the protocol development of our single-cell ribosome phenotype assay. In this chapter, we show that an *E. coli* cell's ribosome distribution can be used to characterise its antibiotic response. For four antibiotics from clinically relevant classes with different mechanisms of action, we show that that the ribosome phenotype provides equivalent information to a DNA and membrane stain, and that the ribosome phenotypes can be identified by trained CNN models. This chapter is adapted from [186].

3.1 Introduction

The internal organisation of a bacterial cell is strongly influenced by physical and biological processes. The shape of the nucleoid has been shown to be driven by physical factors such as entropy, confinement, and cell morphology [187]. Simultaneously, biological factors such as cell growth rate, environmental stressors, and the processes of transcription and translation result in changes to the nucleoid structure [188, 189, 184]. Because of this delicate interplay between physical and biosynthetic factors, antibiotic treatments that disrupt essential bacterial function result in dramatic morphological changes. In quantitative studies of DNA shape, the transcription-blocking antibiotic rifampicin and the translation elongation inhibitor chloramphenicol have been shown to cause significant changes in DNA morphology within minutes of exposure [146, 190]. For this reason, changes in nucleoid shape are a sensitive and rapid indicator of a cell's susceptibility to an antibiotic compound.

To some extent, as the nucleoid shape changes, so does the distribution of ribosomes and ribosomal subunits within the cell. Due to entropy and space filling effects, the ribosome distribution should look like an inverse of the DNA distribution [144, 148]. However, due to the effects of translation on nucleoid structure, ribosomal function and dysfunction also actively contribute to the spatial organisation of the cell [146]. In this work, where the ribosome distribution is assessed through bulk fluorescence measurements, the ribosome phenotype functions as a negative contrast for the DNA morphology, while also outlining the cell shape and offering the possibility of labelling with species-specific probes. It is possible, though, that with higher resolution imaging or time-dependent measurements, the ribosome distribution could provide additional or earlier diagnostic data on the disruption of cellular function by antibiotic treatment.

In this chapter, we explore the use of FISH probes and the ribosome phenotype for characterising the intracellular changes associated with antibiotic susceptibility. First, we show that the ribosome phenotype functions as an inverse DNA stain. Then, we use antibiotics from four classes used to treat *E. coli* infections to show that they produce characteristic ribosome distributions in susceptible cells. Finally,

we show that CNN models can be trained to reliably and accurately differentiate antibiotic-treated from untreated *E. coli* by their ribosome phenotype (Figure 3.1).

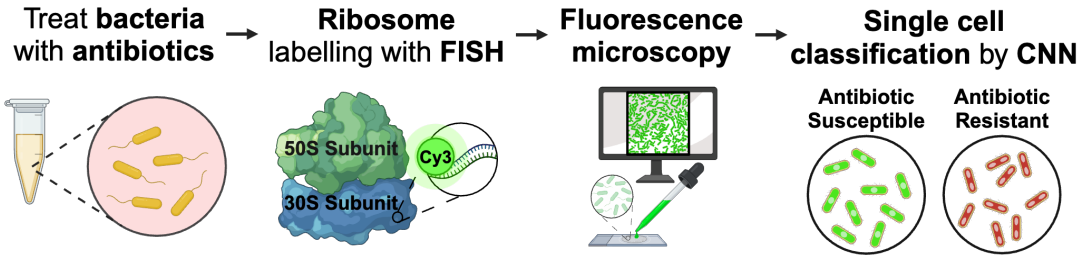


Figure 3.1: Proposed ribosome phenotyping workflow. First the bacteria are treated with antibiotics at a standard concentration for 30 minutes. Then, a standard FISH protocol is used to label the ribosomes with ssDNA fluorescent probes; in this case, EUB338-Cy3 binds a conserved region in the 30S subunit. The samples are imaged on a fluorescence microscope before neural networks use the ribosome signal to segment and then classify the cells as susceptible or resistant to the prescribed antibiotic treatment. If species-specific probes are used, the same images could be used for species ID. Figure made with Biorender.com.

3.2 Characterisation of the *E. coli* antibiotic response by ribosome distribution

We first characterised the sub-cellular ribosome phenotypes of the antibiotic-susceptible *E. coli* MG1655. This strain is a lab-adapted K-12 derivative that is susceptible to each of the tested antibiotics (Table 2.1). After treatment with each of the antibiotics for 30 minutes, the cells were stained with fluorescent FISH probes to visualise the effects of antibiotic treatment on their internal structure. For this, we used an 18-mer single-strand DNA probe with Cy3 dye conjugated to the EUB338 sequence, which targets a region in the 16S ribosomal RNA conserved in all members of the domain *Bacteria* [152]. Details of the sample preparation and imaging protocol can be found in Chapter 2.

The biological effect of antibiotic treatment on the susceptible MG1655 bacteria can be clearly visualised in its DNA and ribosome phenotypes within 30 minutes. Fluorescence images of the DNA and ribosome signals, when compared to those of the untreated sample, show the characteristic changes in cell spatial organisation

that occur as the cell responds to each antibiotic (Figure 3.2). In both untreated and antibiotic-treated conditions, comparing the DNA and ribosome fluorescence images clearly shows the spatial anti-correlation between DNA and ribosome density within the cell. We can detect changes in the nucleoid morphology indirectly because they result in changes in the ribosome fluorescence intensity profile. When the DNA is more dense, there is a greater decrease in the ribosome density in that region. The nucleoid compaction caused by chloramphenicol, ciprofloxacin, and gentamicin can be seen as clearly in the ribosome images as in the DNA images. For carbenicillin, the cell filamentation and multiple DNA regions with lower ribosome intensity can be seen. These images also show how the ribosomes fill the cell, allowing the ribosome signal to be used for both cell phenotyping and cell segmentation.

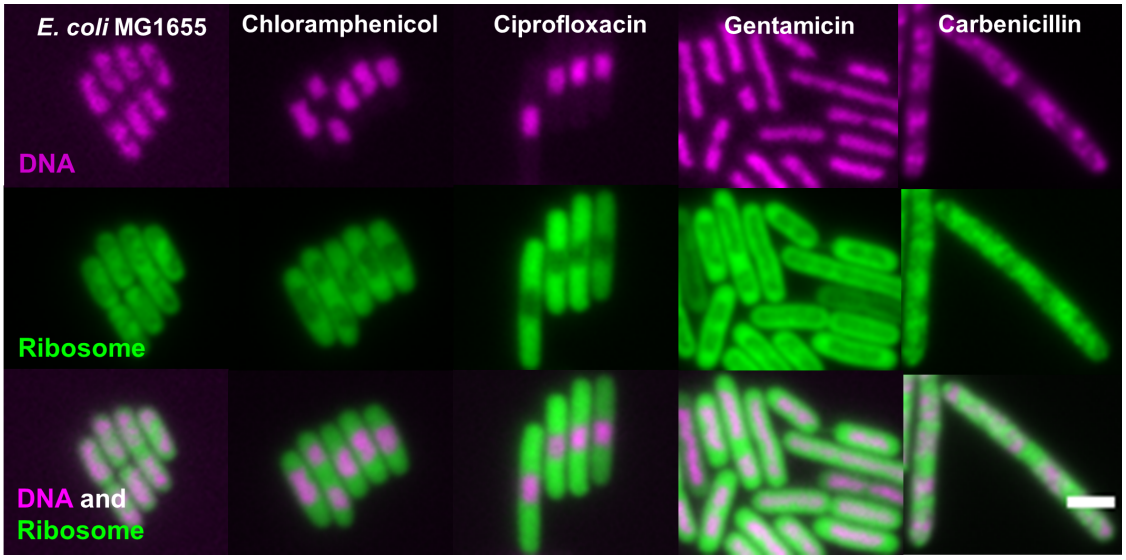


Figure 3.2: Antibiotic-susceptible ribosome phenotypes. Representative fluorescence images are shown of *E. coli* MG1655 with and without antibiotic treatment (magenta, DNA stained with DAPI; green, ribosomes labelled with EUB338-Cy3 probes; combined DNA and ribosome signal). The scale bar is 2 μm. The ribosome density can be seen to anti-correlate with the DNA-dense regions. The untreated panel shows fixed cells with no antibiotic treatment. The chloramphenicol panel shows cells treated with 8 mg/L chloramphenicol (1× EUCAST breakpoint) for 30 minutes before fixation. The ciprofloxacin panel shows the same, treated with 0.5 mg/L ciprofloxacin (1× EUCAST breakpoint). The gentamicin panel shows the same, treated with 40 mg/L gentamicin (20× EUCAST breakpoint). The carbenicillin panel shows the same, treated with 24 mg/L carbenicillin (3× EUCAST breakpoint).

We characterised the ribosome phenotypes from four biological replicates of *E.*

coli MG1655 totalling 5,286 untreated cells, 3,215 cells treated with ciprofloxacin (Cip) at 0.5 mg/L (1× EUCAST breakpoint), 5,935 cells treated with gentamicin (Gent) at 40 mg/L (20× EUCAST breakpoint), 6,439 cells treated with chloramphenicol (Cam) at 8 mg/L (1× EUCAST breakpoint), and 1,256 cells treated with carbenicillin (Carb) at 24 mg/L (3× EUCAST breakpoint). For *E.* coli MG1655, treatment with chloramphenicol or ciprofloxacin at 1× EUCAST induced phenotypic changes within 30 minutes, but gentamicin treatment concentrations lower than 20× EUCAST and carbenicillin treatment concentrations lower than 3× EUCAST did not induce phenotypic changes in most cells in this time frame (Figures 3.3, 3.4). These results for gentamicin and carbenicillin suggest that the antibiotic treatment concentration should be optimised for each antibiotic and its mechanism of action. Then, this established treatment concentration and phenotype can serve as a benchmark to characterise the magnitude of response in other strains.

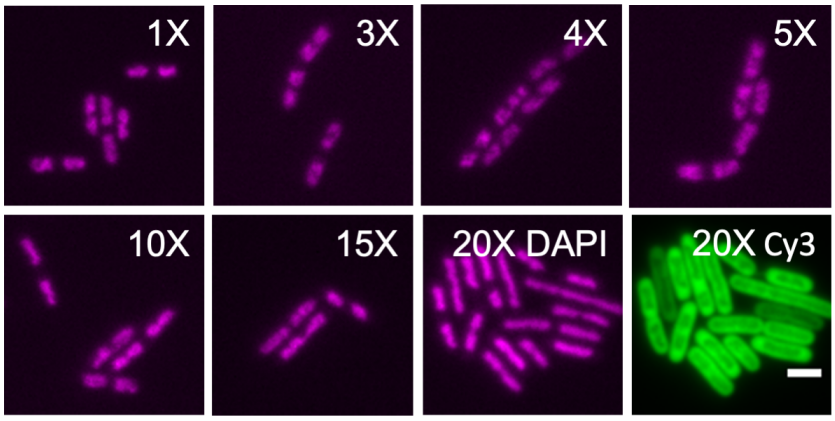


Figure 3.3: Phenotypic changes associated with gentamicin treatment for 30 minutes at various concentrations from 1× to 20× EUCAST. Representative images of *E. coli* MG1655 treated with gentamicin in LB media for 30 minutes, labelled with the gentamicin concentration in multiples of the EUCAST breakpoint (see Table 2.1). Images show the nucleoid (DAPI) unless otherwise labelled. Samples were labelled with DAPI and EUB338-Cy3 as described in the *Methods* (2.5.2). Scale bar, 2 µm.

We further characterised the treatment phenotypes by inspecting the mean DNA and ribosome fluorescence signals along the long axis of the cells, as described in 2.10. In untreated *E. coli* MG1655, the highest ribosome density was seen in the centre of the cell and in longer cells there were often two ribosome-poor nucleoid regions (Figures 3.2, 3.5a). Ciprofloxacin treatment caused a central,

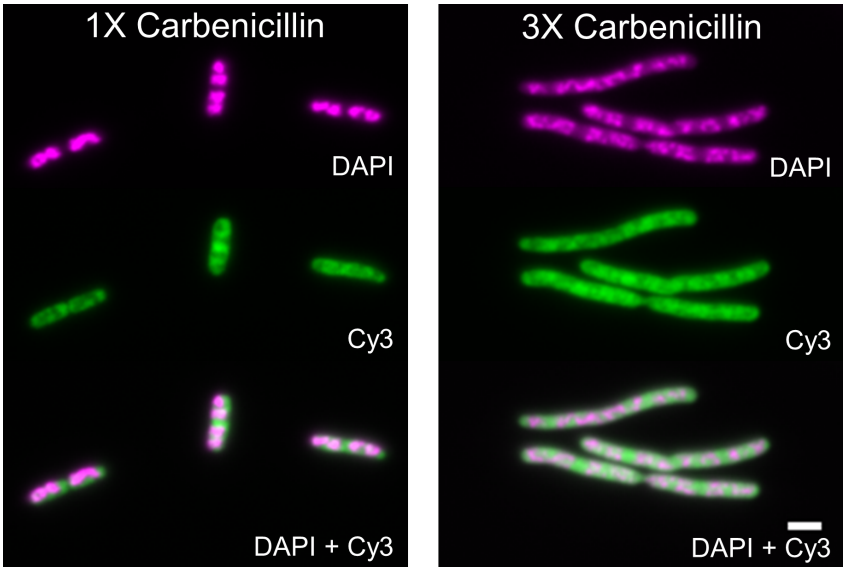


Figure 3.4: Phenotypic changes associated with carbenicillin treatment for 30 minutes at 1× and 3× EUCAST. Representative images of *E. coli* MG1655 treated with carbenicillin in LB media for 30 minutes, treated with 1× or 3× the EUCAST breakpoint (see Table 2.1). Images show the nucleoid (DAPI), the ribosomes (Cy3), and the super-imposed DAPI and Cy3 images. Samples were labelled with DAPI and EUB338-Cy3 as described in the *Methods* (2.5.2). Scale bar, 2 µm.

compact nucleoid region (Figures 3.2, 3.5a) and resulted in cells that were longer than untreated cells (Figure 3.5b). Gentamicin treatment led to a diffuse nucleoid region following the long axis of the cell that was often rod-shaped (Figures 3.2, 3.5a). Chloramphenicol treatment caused nucleoid compaction compared to the untreated phenotype, causing either a centralised DNA region or two dense DNA regions (Figures 3.2, 3.5a). Carbenicillin treatment caused filamentation with much longer cells (Figure 3.5b) and four copies of the DNA in a typical cell (Figures 3.2, 3.5a). All antibiotic treatments resulted in cells with significantly different average length and average width compared to the untreated phenotype (Mann-Whitney non-parametric hypothesis test $p<0.05$, Figure 3.5b-c).

The antibiotic treatment phenotypes found in our images aligned with those found in previous work [146, 191, 145] and with the mechanism of action of each antibiotic (Figures 3.2, 3.5). In the cases of ciprofloxacin and carbenicillin, the increased median cell length caused by the antibiotic halting cell division may be sufficient to judge an antibiotic response from this metric alone [192], but ribosome

phenotypes provide additional data on internal changes. Because these phenotypes are quantifiable by ribosome fluorescence intensity mapping and identifiable by the human eye, it follows that a neural network could be trained to associate them with an antibiotic treatment response.

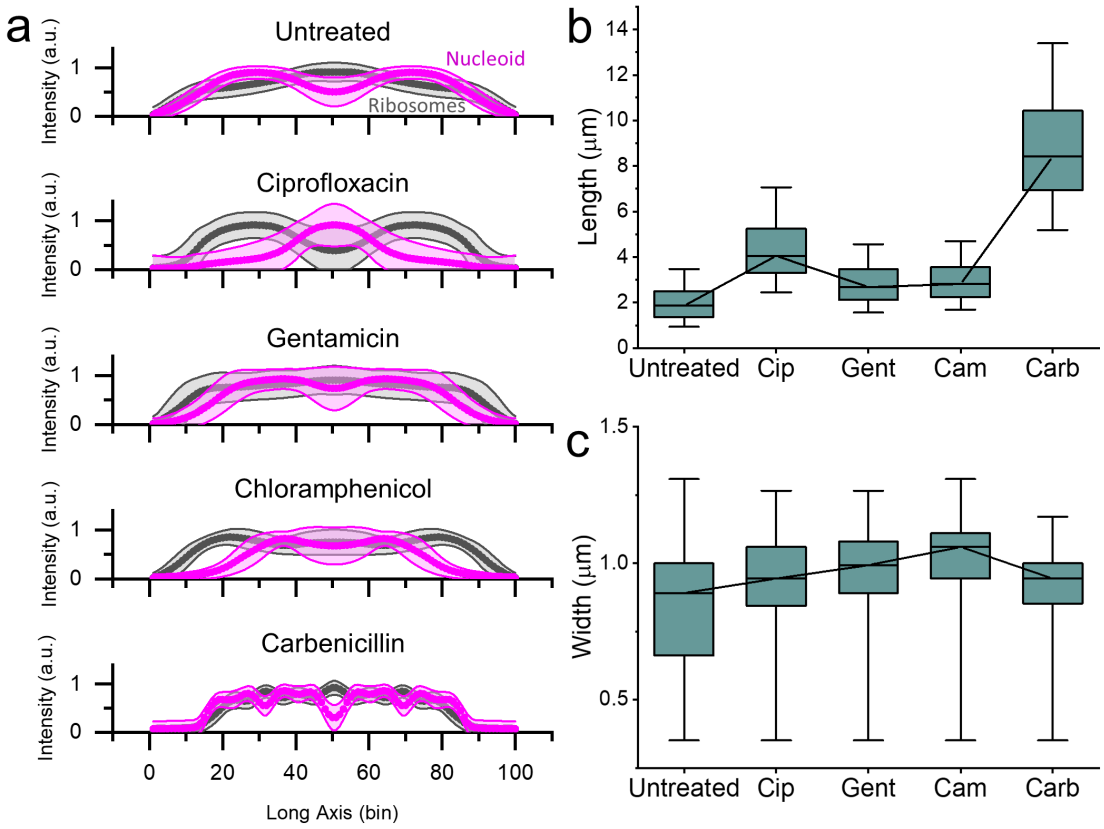


Figure 3.5: Fluorescence profiles reveal ribosome-nucleoid anti-correlation and characteristic phenotypes of antibiotic response.
(a) Along the long axis of the cell, the mean normalised ribosome (Cy3, grey) and nucleoid (DAPI, magenta) intensities are calculated for each of 100 bins. The shading shows ± 1 standard deviation of the mean. The untreated *E. coli* line profiles show two nucleoid-rich regions, correlated with decreased ribosome intensity. This figure is composed of profiles from 5,286 untreated *E. coli* MG1655. The ciprofloxacin panel shows the same, for 3,215 *E. coli* MG1655 treated with 1× EUCAST ciprofloxacin for 30 minutes. The line profile shows a central, compact nucleoid region with greater segregation from the ribosomes. The gentamicin panel shows the same, for 5,935 *E. coli* MG1655 treated with 20× EUCAST gentamicin for 30 minutes. This line profile shows a diffuse nucleoid region along the long axis of the cell with less ribosome-nucleoid segregation. The chloramphenicol panel shows the same, for 6,438 *E. coli* MG1655 treated with 1× chloramphenicol. The carbenicillin panel shows the same, for 1,256 *E. coli* MG1655 treated with 3× EUCAST carbenicillin for 30 minutes. This line profile shows multiple nucleoid regions along the long axis of the cell.
(b) The cell lengths (μm) are shown for untreated *E. coli* MG1655 and for each of the antibiotic treatments. The box shows the 25%∼75% percentile range, the bars show the 1%∼99% percentile range, and the line denotes the median. Each antibiotic results in a length distribution statistically different from the untreated population with $p<0.05$ by the Mann-Whitney non-parametric test.
(c) The cell widths μm are shown for untreated *E. coli* MG1655 and for each of the antibiotic treatments, shown as in (b). Each antibiotic results in a width distribution statistically different from the untreated population with $p<0.05$ by the Mann-Whitney non-parametric test.

3.3 Strategies to maximise transferable learning

To test the reliability and accuracy of the neural networks in differentiating the antibiotic-susceptible phenotypes from untreated *E. coli* MG1655, a cross validation test was performed for each antibiotic by rotating the holdout dataset (Figure 3.6). Cross validation, the process of training and testing a statistical model on separate datasets, is the simplest method for estimating the prediction error of a model [193]. For each experiment, a model was independently trained and validated on data from three of the biological replicates and tested on the fourth. The validation and testing datasets were balanced to include an equal number of untreated and antibiotic-treated cell images to minimise prediction biases. The accuracy across these cross validated samples should more closely approximate the accuracy of the model on unseen data.

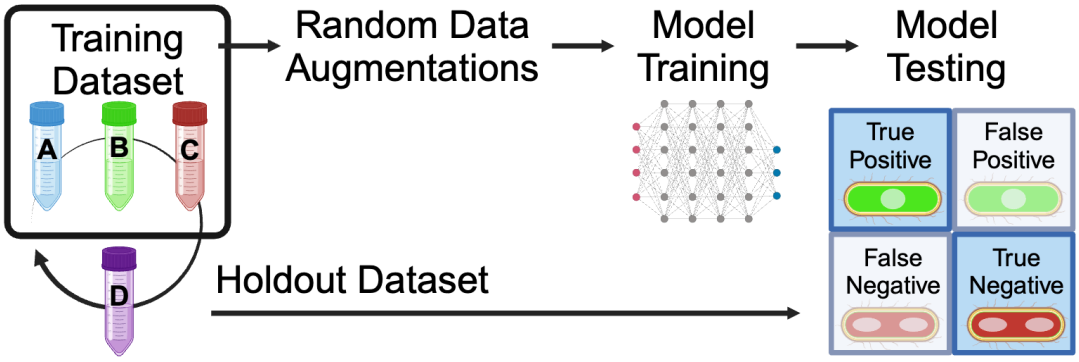


Figure 3.6: Cross validation procedure across biological replicates for model training and testing. Four biological replicates of *E. coli* MG1655 were tested for each antibiotic and for the untreated condition. To test phenotype robustness and repeatability, a holdout cross validation was performed in which each model was trained and validated on images from three of the biological replicates (e.g. A, B, C) and tested on images from the fourth (e.g. D). The training images received random data augmentations before being passed to the model for training, whereas the holdout dataset was passed directly to the model for testing.

Our fluorescence images were pre-processed prior to their use as training data. First, each single-channel image was segmented by a custom CellPose model [173] trained to segment *E. coli* cells by their ribosome fluorescence image. The segmentations were subsequently curated to refine the outlines and remove cells

that were outside of the field of view, overlapping, or outside of the focal plane [172]. These steps resulted in single-cell images removed from the image background.

Each single-cell segmentation was used to create a 64-by-64 zero-filled image with the ribosome fluorescence in the centre. The CNN requires every item in the dataset to have the same dimensions, and its feature learning is dependent on an object's scale. When a segmented cell exceeds these dimensions, the cell must be cropped or resized with a preserved aspect ratio. For our single-cell images, the cell length is a scale-independent feature because *E. coli* maintain a consistent width as they grow lengthwise. Resizing preserves data from the whole cell length, but when a cell is resized too much, the internal structure is compressed (Figure 3.7). The decision to crop or resize results in very different images to the human eye, but for our dataset caused almost no change in classification accuracy. Models trained on cropped or resized carbenicillin-treated cells had remarkably similar balanced accuracies (within ± 0.1%). We chose to use the cropping method to preserve the intracellular definition of the ribosome phenotype, but for other applications resizing might be more appropriate. Regardless of whether the cells are cropped or resized, it is likely that extremely long cells are the predominant feature of the carbenicillin-treated phenotype.

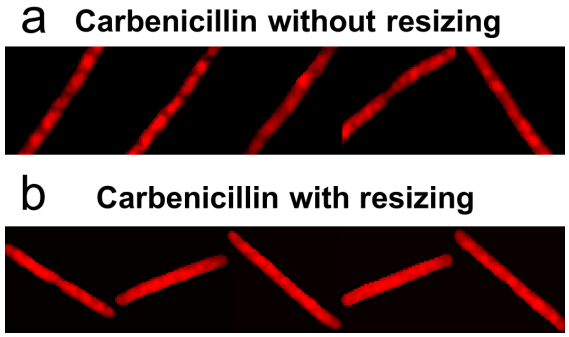


Figure 3.7: Effects of image resizing on carbenicillin ribosome phenotype. (a) Representative images of carbenicillin-treated cells without resizing. Each 64×64 single-cell image is 7.5×7.5 μm^2. Before the image is passed to the CNN, the segmented cell is centred, but the cell poles are cropped by the standardised size. However, the regions of decreased ribosome density (nucleoid regions) can be clearly seen throughout the cell.
(b) Representative images of carbenicillin-treated cells with resizing. These segmented cells are resized with preserved aspect ratio to fit in 64×64 images before being passed to the CNN. As a result, the intracellular definition is lost for the longest cells.

To regularise learning and minimise over-fitting, augmentations were applied before each image was loaded into the training dataset using the Albumentations package [174]. First, geometric transformations were used to remove the cell from its location in the image. Flips, 90-degree rotations, and translations within the 64-by-64 image were applied randomly. Changes in scale (scaling factor between 0.8 and 1.2) and shear (between -20 and 20 degrees) were used to further increase the diversity in the training dataset. Then, random amounts of Gaussian noise (variance range = 0.0005), pixel dropout (dropout probability = 0.05), and brightness and contrast adjustments (brightness limit = 0.1, contrast limit = 0.5) were randomly applied to limit the model's ability to learn spurious correlations in image qualities from a single dataset. These augmentations mean that many representations of a single cell can be shown to the model as a training image. Data augmentation is helpful and improves model training as long as the distortions do not remove key features that are essential for categorising the image [194].

We trained models with and without shear and blur (Figure 3.8). For ciprofloxacin and carbenicillin, classification accuracies were similar for both models; this might be because length is a strong predictive feature for these antibiotics (Figure 3.5). However, shear and blur distort the gentamicin and chloramphenicol ribosome phenotypes more significantly. Representative training images from gentamicin- and chloramphenicol-treated cells augmented with shear and blur show several examples where the internal morphology is significantly distorted (Figure 3.9). Looking at those distorted images, it would be impossible to say whether the nucleoid has been compacted from the antibiotic treatment. Augmentation to this degree can harm the model's ability to learn from the training data, especially when the augmentations are far beyond the variance expected in real-world datasets. For these reasons, removing shear and blur from our augmentations pipeline resulted in improved balanced accuracy for the gentamicin (+2%) and chloramphenicol (+4%) models, with less variance in cross-validation testing. For this project, we decided to remove shear and blur, while keeping the other augmentations such as translations, scaling, noise, brightness, and pixel dropout.

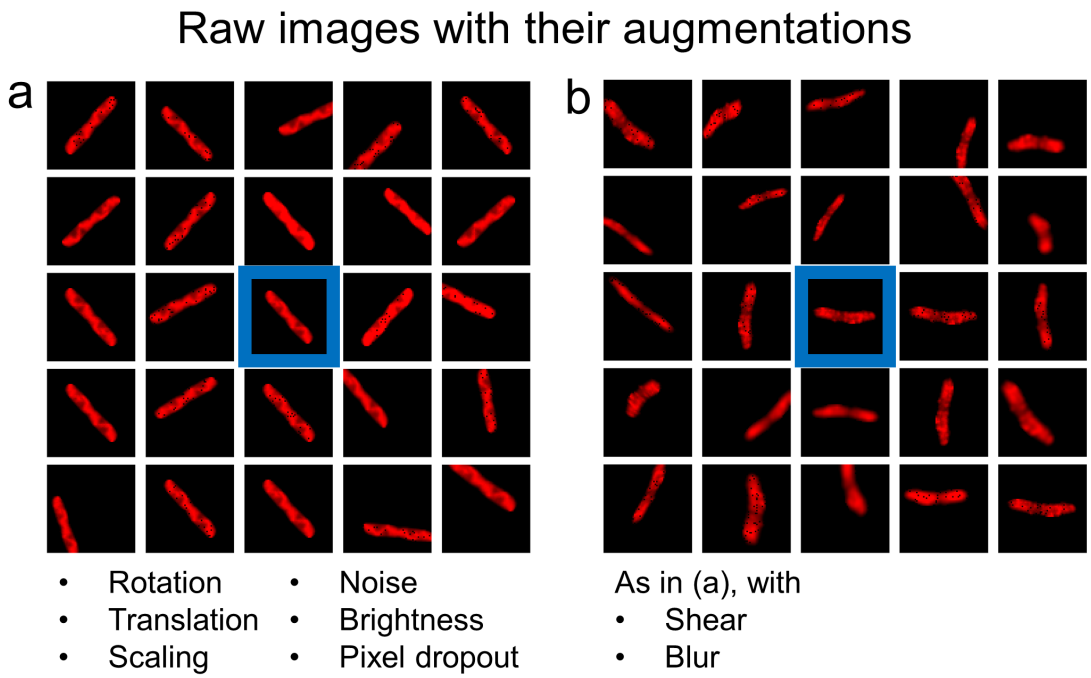


Figure 3.8: Image augmentations with and without shear and blur. Raw images (centre, highlighted in blue) and a selection of resulting augmented images are shown for representative ciprofloxacin-treated cells. Each 64×64 single-cell image is 7.5×7.5 μm^2.
(a) Some of the augmentations can be seen, such as random rotations, translations, scaling, Gaussian noise, brightness adjustments, and pixel dropout.
(b) As in (a), with additional random shear and blur.

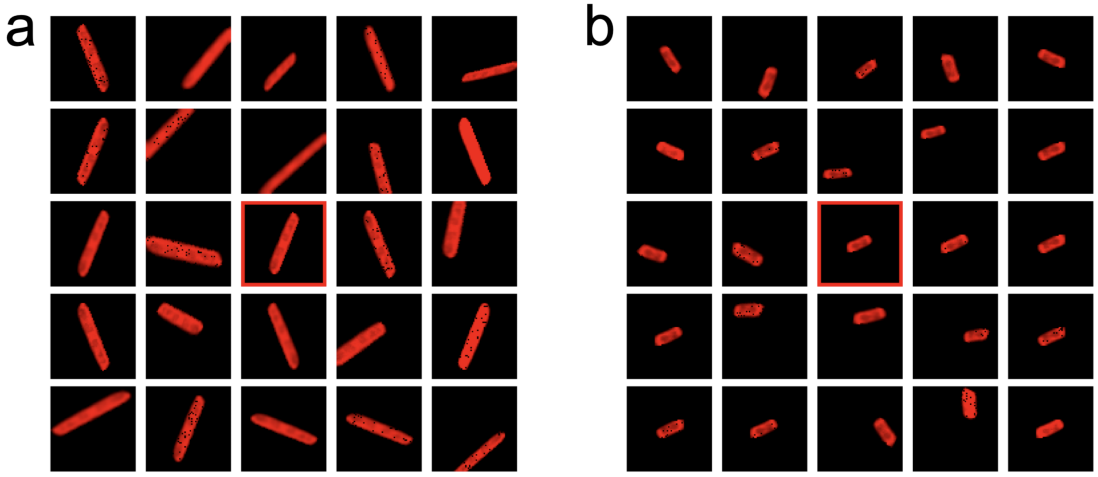


Figure 3.9: Gentamicin and chloramphenicol phenotypes are distorted by shear and blur. Raw images (centre, highlighted in red) and a selection of resulting augmented images are shown for representative (a) gentamicin-treated cells with blurring and shearing and (b) chloramphenicol-treated cells without blurring or shearing. Each 64×64 single-cell image is 7.5×7.5 μm^2. In (a), the ribosome phenotype is obscured in several of the augmented images. In (b), the ribosome phenotype can be seen in most images, despite added noise and brightness adjustments.

3.4 Identification of antibiotic-susceptible ribosome phenotypes by a neural network

The average balanced accuracy of the models on the four test datasets exceeded 95% for all antibiotics: ciprofloxacin ∼ 99 ± 1.5%, gentamicin ∼ 98 ± 2.5%, chloramphenicol ∼ 95 ± 8.8%, and carbenicillin ∼ 99 ± 1.6% (Figure 3.10). Looking at the models' predictions, the cells that were classified correctly with the highest confidence by the model (Figure 3.11) demonstrated the characteristic antibiotic phenotypes (Figure 3.2). The ciprofloxacin true positives show elongated or filamentous cells with a central, compact DNA region with decreased ribosome intensity. The gentamicin true positives show DNA regions that are compacted along the long axis of the cell and the chloramphenicol true positives have compact, central DNA regions without as much elongation as is seen in the ciprofloxacin-treated cells. The carbenicillin true positives are very elongated and have multiple DNA regions, showing multiple rounds of nucleoid replication without cell division.

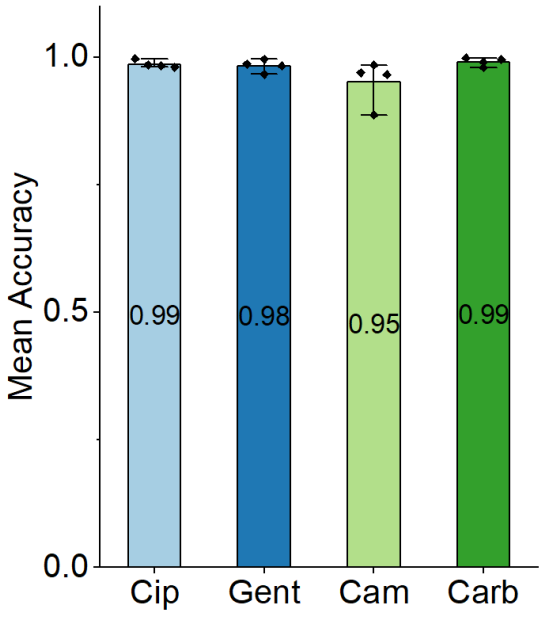


Figure 3.10: Ribosome phenotype classifier accuracy by antibiotic. The balanced accuracy of the ribosome phenotype classifier is shown for each antibiotic. Each point represents a cross-validated biological replicate. The mean balanced accuracy is shown on each column and the error bars indicate the 95% confidence interval of the mean on four biological replicates.

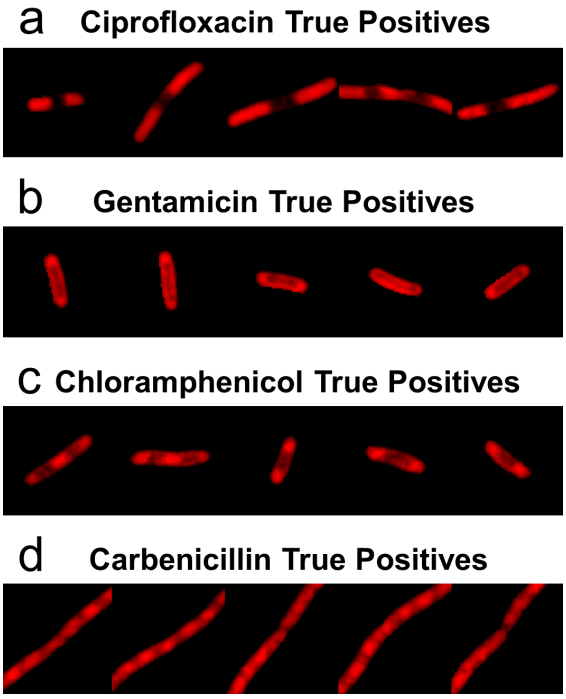


Figure 3.11: Representative classifications for the ciprofloxacin, gentamicin, chloramphenicol, and carbenicillin phenotype models. Example ribosome phenotypes are shown for correctly classified *E. coli* MG1655 cells treated with ciprofloxacin (a), gentamicin (b), chloramphenicol (c), and carbenicillin (d) as described in the *Methods* (2.5). Each 64×64 single-cell image is 7.5×7.5 μm^2. The ciprofloxacin cells show elongated cells with a central, compact nucleoid region. One image shows a cell where the poles have been cut off by the standardised 64×64 image passed to the CNN. The gentamicin cells show elongation and a nucleoid region that has compacted along the midline. The chloramphenicol cells show central, compact nucleoid regions. The carbenicillin cells show multiple nucleoid regions and filamentation. All of the carbenicillin-treated cells are longer than the 64×64 image passed to the CNN.

We then examined the confusion matrices for the four holdout datasets, summed together. The ciprofloxacin phenotype model was one of the most accurate, with a very high average balanced accuracy of 98.9%, sensitivity of 97.5%, and specificity of 99.5% (Figure 3.12, Cip). This may be because the ciprofloxacin response causes two phenotypic changes – elongated cells and condensed, central nucleoid – both of which can be used by the model in the classification task (Figure 3.13). The gentamicin phenotype model was also highly accurate, achieving an average balanced accuracy of 99.0%, sensitivity of 98.5%, and specificity of 99.5% (Figure 3.12, Gent). The chloramphenicol phenotype model had a balanced accuracy of 95.8%, sensitivity of 93.1%, and specificity of 98.5% (Figure 3.12, Cam). Inspection

of the chloramphenicol-treated cells that were misclassified as untreated suggests that this model's increased number of False Negative classifications was driven by cells that did not adopt the expected chloramphenicol-treated phenotype within the antibiotic treatment period, having multiple, diffuse nucleoid regions (Figure 3.14a), whereas False Positive classifications tended to have a central nucleoid region (Figure 3.14b). Finally, the carbenicillin phenotype model had a balanced accuracy of 99.4%, sensitivity of 99.2%, and specificity of 99.6% (Figure 3.12, Carb). The balanced accuracies of >90% reported here are for single cells and therefore the cumulative accuracy of the classifier on a collection of cells is essentially 100%. Accuracy on a sample of cells is discussed further in Chapter 3 (4.4).

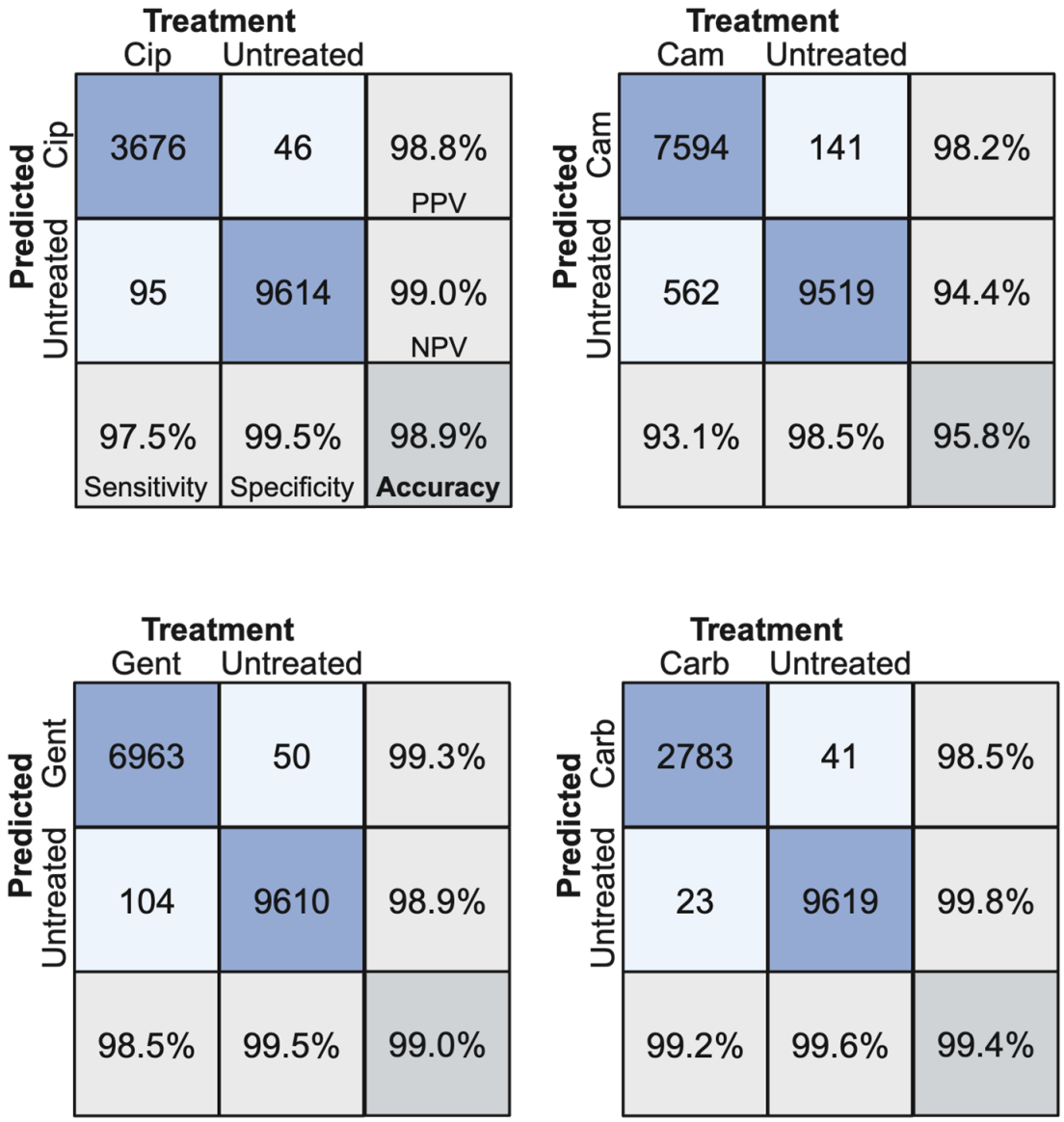


Figure 3.12: Confusion matrices for the ribosome phenotype classifier by antibiotic. A confusion matrix is shown for each of the antibiotic phenotypes tested in *E. coli* MG1655: ciprofloxacin (Cip), gentamicin (Gent), chloramphenicol (Cam), and carbenicillin (Carb). The total number of cells is a sum of the results from four experiments, each with a model trained on three biological replicates and tested on a fourth holdout replicate. The number of images in each class is shown, along with the percentage of cells for each treatment condition. The treatment condition is shown on the columns and the model's predicted classification is shown on the columns. Right column: positive predictive value (PPV) and negative predictive value (NPV) of the model's predictions are shown. Bottom row: accuracy of the model on antibiotic-treated cells (Sensitivity), accuracy of the model on untreated cells (Specificity), and the Balanced Accuracy (Accuracy) are shown. See *Accuracy Metrics* for details (2.9).

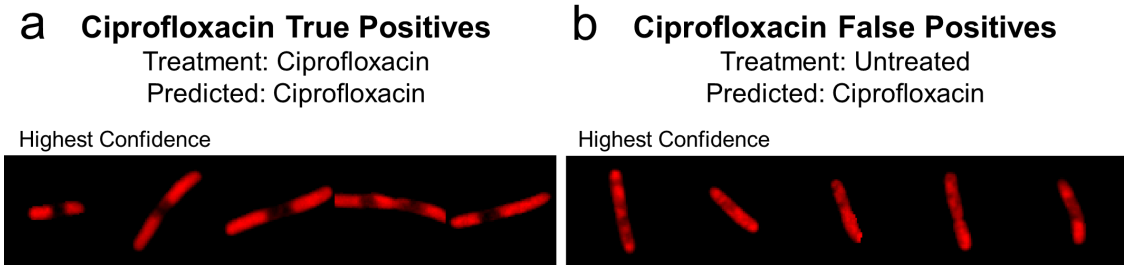


Figure 3.13: Images of cells classified as ciprofloxacin-treated by the model. Representative ribosome phenotypes of *E. coli* MG1655 are shown for (a) True Positives (ciprofloxacin-treated, predicted ciprofloxacin-treated) and (b) False Positives (untreated, predicted ciprofloxacin-treated) that were classified with the highest confidence by the model. Each 64×64 single-cell image is 7.5×7.5 μm^2. The True Positive images show an elongated phenotype and a central nucleoid region with lower ribosome intensity, especially in the images classified with the highest confidence. The False Positive images that were misclassified are elongated, or have a central nucleoid region, or both. This suggests that the ciprofloxacin model is using cell length and the existence of a central nucleoid to classify cells as ciprofloxacin-treated.

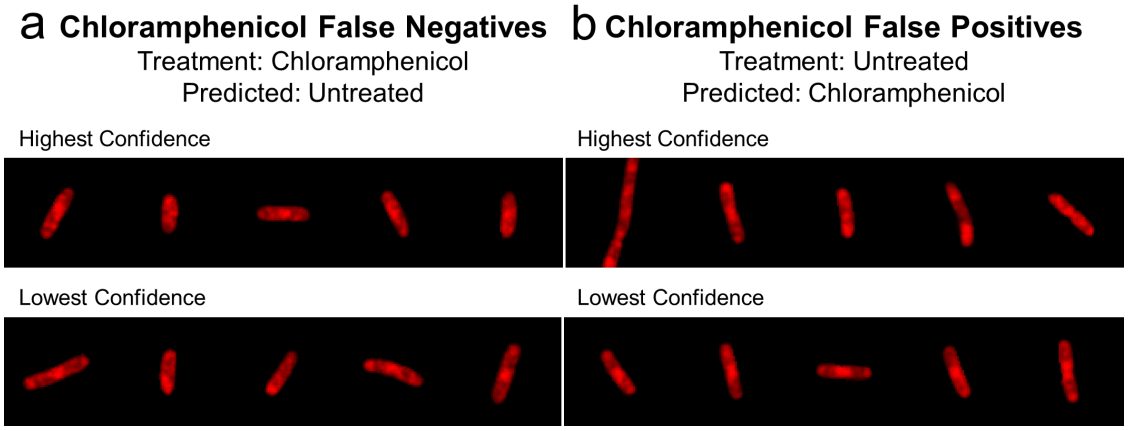


Figure 3.14: Images of *E. coli* MG1655 incorrectly classified by the chloramphenicol model. Example ribosome phenotypes of (a) False Negatives (chloramphenicol-treated, predicted untreated) and (b) False Positives (untreated, predicted chloramphenicol-treated). The examples are stratified by the confidence of the model's prediction. Each 64×64 single-cell image is 7.5×7.5 μm^2. The False Negative cells show multiple, diffuse nucleoid regions in chloramphenicol-treated cells, whereas the False Positives show central nucleoid regions in untreated cells.

3.5 Discussion

We have shown that ribosome-targeted FISH probes (EUB338-Cy3) can be used to visualise intracellular antibiotic response phenotypes in *E. coli* that differ based on the mechanism of action of the antibiotic. In *E. coli* MG1655, we demonstrated that distinct ribosome phenotypes exist for four antibiotics from clinically relevant classes

(ciprofloxacin, a fluoroquinolone; gentamicin, an aminoglycoside; chloramphenicol, an amphenicol; and carbenicillin, a β-lactam). The ribosome phenotypes detected with EUB338-Cy3 labelling correspond with the structural and morphological changes found previously by other imaging methods [93, 146, 195]. We trained CNNs with a cross-validation method to show that antibiotic-treated ribosome phenotypes can be detected reliably and accurately across biological replicates with >95% accuracy.

3.5.1 Strains and antibiotics

The DNA and ribosome phenotypes shown here are characteristic of each antibiotic's mechanism of action. The single, compact central nucleoid found in ciprofloxacin-treated cells is thought to be due to DNA super-compaction as a result of DNA damage and the SOS response, with the RecN DNA-binding protein forming bridges and attachments to double-stranded breaks, although the mechanisms are not yet fully understood [196]. The extremely long lengths of carbenicillin-treated cells are also expected based on its mechanism of action. Carbenicillin, like other β-lactams, has been shown to permit lengthwise, filamentous growth, but prevents cell division, before leading to cell lysis [197]. It is interesting that the median length of carbenicillin-treated cells in our dataset is approximately 4-times that of untreated cells, which is longer than what might be expected after 30 minutes of growth in rich media. While the mechanism of this extreme filamentation and elongation is not yet fully understood, our results are consistent with the average cell lengths seen in other studies of the effects of β-lactam antibiotics on *E. coli* cultured in rich media at 37°C [198].

This initial study was limited to one lab strain of *E. coli* treated with four antibiotics during rapid, exponential growth in rich liquid media. Although these initial results were promising, it is likely that cells exposed to antibiotics in more nutrient-poor or stressed states would have less well-defined nucleoid and ribosome phenotypes, due to the effects of growth rate on nucleoid organisation. In future work, it will be important to collect data on the phenotypes that result from antibiotic

treatment in different growth media, especially because cells filtered directly from clinical samples are more likely to be in a stationary or slow-growth phase [183].

In addition to the effects of cell growth rate, some antibiotics may have mechanisms of action that are less likely to produce a ribosome phenotype within a short treatment time. Further testing of other *E. coli* strains or other species might also yield biological diversity that makes the classification task more challenging. However, it is promising that bacterial cytological profiling has already been demonstrated to correlate with antibiotic susceptibility in other studies of *E. coli* [93] and in other species, including cocci such as *Staphylococcus aureus* [91, 111].

Over the course of this project, we tested two other antibiotics, kasugamycin and amoxicillin-clavulanic acid, in *E. coli* MG1655 for ribosome phenotype profiling. Kasugamycin is another member of the aminoglycoside class that has been used in high concentrations (5,000 mg/L) in previous work as a translation inhibitor [146]. The mechanism of action of kasugamycin inhibits translation initiation, which should result in an increased fraction of dissociated 30S and 50S ribosomal subunits that mix with the nucleoid and result in nucleoid decompaction [199]. In our work with *E. coli* MG1655, we were challenged by that strain's high resistance to kasugamycin (MIC = 5,440 mg/L). Our kasugamycin experiments were conducted at a lower concentration (400 mg/L), which resulted in poor conversion to the treated phenotype and False Negative rates of around 50%. We would predict that with a higher concentration of kasugamycin, above the MIC of the strain, the phenotype would develop more reliably and could be accurately classified by the CNN. Ultimately, we chose to use gentamicin as a representative of the aminoglycoside class because it is commonly used to treat severe *E. coli* infections, whereas kasugamycin is more commonly used in research.

We also conducted initial experiments with amoxicillin-clavulanic acid, a combined penicillin and β-lactamase inhibitor that is frequently used for *E. coli* infections and showed a subtle phenotype with treatment at 20× EUCAST for 60 minutes in previous work [93]. We tested amoxicillin-clavulanic acid with 1× EUCAST (8 mg/L) for up to 120 minutes and found no visually apparent phenotypic changes. After

treatment with 20× EUCAST for 30 minutes in our conditions, only 1-2 cells per mL remained, with no morphological changes. Based on the mechanisms of action of amoxicillin and clavulanic acid, we expect that the rest of the cells had compromised cell membranes and burst during the treatment period. It is possible that other concentrations of amoxicillin-clavulanic acid between 1× and 20× EUCAST would result in a reliable phenotype during the 30 minute treatment period. Or, if cells are being imaged before and after treatment, bursting could be used as a read-out of antibiotic susceptibility. We chose to move to carbenicillin to characterise the ribosome phenotype of a typical penicillin without a beta-lactamase inhibitor.

With gentamicin and carbenicillin, we were able to find antibiotic treatment concentrations above 1× that reliably induced a clear phenotype within 30 minutes of antibiotic treatment. Our ability to explore other antibiotics and antibiotic concentrations was limited by the throughput of our agarose pad imaging method. It is possible that other higher-throughput imaging strategies, such as microfluidic devices or 96-well plates, could be used to characterise more antibiotics and antibiotic concentrations in future work.

3.5.2 Neural network training and augmentations

In applying models to our datasets, we explored the effects of different image processing and augmentation strategies. Two of our antibiotics resulted in filamentous and elongated cells — ciprofloxacin-treated cells were around twice as long as our untreated cells and carbenicillin-treated cells were around 4-times as long. Many of these cells would not fit in the 64×64 image required by the CNN without cropping or resizing. For these antibiotics, we found that both cropping and resizing resulted in similar accuracies. However, there are cases where this would need to be carefully considered, for example if the cells are resized to the point where intracellular morphology is completely lost, or if important phenotype data is lost by cropping the cell poles. This decision should be tailored to the application and dataset.

We also explored the effects of blurring and shearing augmentations on the model's ability to learn ribosome phenotypes (Figure 3.15). For ciprofloxacin

and carbenicillin, there was little difference in the accuracy between models with and without these augmentations. However, learning and accuracy improved for gentamicin and chloramphenicol when blurring and shearing were removed from the augmentations to the training data. The representative training images show that our application of shear and blur distorted or erased the ribosome phenotype and made it difficult to infer the nucleoid morphology from the ribosome signal. Although shear and blur are commonly used in augmenting training data for CNNs and have been used in other image classification tasks for clinical microbiology [200], for our task, performance on the test dataset improved without these augmentations. It is possible that using less shear and blur would maintain the ribosome phenotype and yield some of the benefits for model training. When considering training a model for a future diagnostic test, it will need to handle blurry or out-of-focus cells. The pipeline's ability to handle blurry cells could potentially be improved by including slightly out-of-focus cells in the dataset, for example by collecting images with some Δz from the focal plane; by reintroducing a smaller amount of blur to the training data augmentations; or by filtering blurry cells from the pipeline, for example by creating a threshold for the image Laplacian [201].

Some of these problems could be avoided if an alternative machine learning method was used, for example a network trained to classify phenotypes based on the intensity profile along the long axis of the cell, cell length or cell width (Figure 3.5). Because of the predictable phenotypes seen in *E. coli* MG1655, a model designed to use a simplified data structure might be able to generate predictions with similar accuracy to our CNN. In a more complicated classification task, the CNN would be expected to learn additional spatial features that would give it a performance advantage over a simpler model. A CNN could also be trained to classify cropped images of cell populations, rather than single cells.

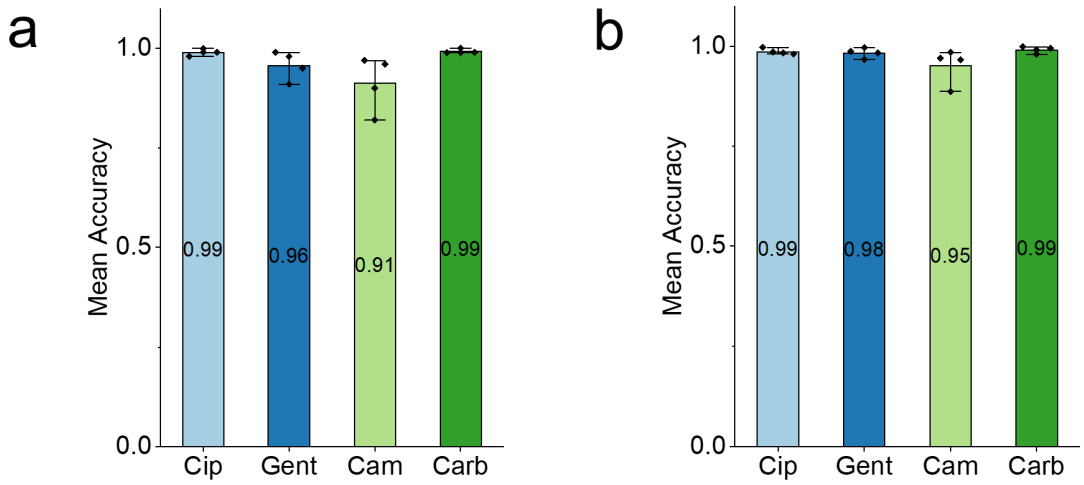


Figure 3.15: Model accuracy with and without shear and blur.
(a) Repeated from Figure 3.10, the balanced accuracy of the ribosome phenotype classifier is shown for each antibiotic using the augmentations demonstrated in Figure 3.8b, which include shear and blur. Each point represents a biological replicate. The mean balanced accuracy is shown on each column and the error bars indicate the 95% confidence interval of the mean on the four biological replicates. The mean balanced accuracy for gentamicin is 0.96 and the mean balanced accuracy for chloramphenicol is 0.91.
(b) As in (a), for models trained without shear or blur. The mean balanced accuracy for gentamicin has increased to 0.98 and the mean balanced accuracy for chloramphenicol has increased to 0.95.

3.6 Conclusions

In this chapter, we showed that ribosome labelling with FISH probes can be used to report changes in cell spatial organisation and morphology that result from antibiotic treatment. We used these ribosome phenotypes to train CNNs with high single-cell accuracy (>95%) in classifying antibiotic-treated and untreated *E. coli* MG1655 in cross-validation tests. We showed that ribosome phenotypes can be used to detect the antibiotic response in four clinically relevant antibiotic classes with different mechanisms of action.

4

Detecting antimicrobial resistance in clinical isolates by ribosome phenotype

Contents

4.1 Introduction 61
4.2 Model trained on lab strain has variable accuracy on clinical isolates 63
4.3 Developing a susceptible-resistant clinical isolate classifier 65
4.4 Necessary sample size for testing 69
4.5 Model performance on unseen strains and antibiotic concentrations 71
4.6 Discussion 74
4.7 Conclusions 76

In the previous chapter, we showed that ribosome-targeting FISH probes can be used to visualise changes in bacterial cellular morphology associated with antibiotic treatment. In this chapter, we apply the ribosome phenotyping method in *E. coli* strains isolated from bloodstream infections, using a selection of ciprofloxacin-susceptible and resistant strains. We also explore some of the challenges associated with moving from a lab strain to clinical strains, and test the ability of models to generalise to other strains and treatment conditions. This chapter is adapted from [186].

4.1 Introduction

Having demonstrated that ribosome phenotypes can reliably report antibiotic response and can be classified by CNN models with high accuracy, we moved to test this methodology on bacteria isolated from clinical samples. Until this point, our models were trained and tested on a laboratory-adapted strain, *Escherichia coli* MG1655, and all of our samples were considered to be genetically identical.

Compared to our laboratory strain, clinical isolates of *E. coli* will exhibit greater diversity in morphology and antimicrobial response. Pathogenic strains of *E. coli* are categorised as one of more than 160 serotypes, defined by their polysaccharide O-antigens, flagellar H-antigens, and capsular K-antigens [202]. In most communities, a subset of these serotypes are responsible for the majority of infections [170]. As whole-genome sequencing of bacteria becomes more common in clinical laboratories, prevalent strains are increasingly also classified by multi-locus sequence types [170]. These sequencing studies continue to find extraordinary genetic diversity in *E. coli* isolates. *E. coli* has an open and large pangenome of more than 13,000 genes, which means that the species is actively evolving by gene acquisition and diversification [203]. Genes can also be shared by horizontal gene transfer between *E. coli* strains. In particular, virulence factors and AMR-associated genes are known to be transferred between *E. coli* strains by mobile genetic elements such as transposons, plasmids, bacteriophages, and pathogenicity islands [163, 204]. This means that many pathogenic *E. coli* isolates have a collection of resistance and virulence factors, including mutations in target genes and efflux pumps. Most of these resistance factors introduce a fitness cost in the absence of antibiotics, which means that antibiotic-susceptible and resistant clinical isolates might show phenotypic differences even before antibiotic treatment [46]. We would also expect there to be more variance in the single-cell ribosome phenotypes of susceptible and resistant strains because of their different tolerance mechanisms. For example, a strain with high expression of efflux pumps or target genes with lower binding efficiency for the antibiotic will be effectively experiencing a lower antibiotic treatment concentration.

Sampling all of this biodiversity will take time. However, we decided to apply the ribosome phenotyping method on a selection of *E. coli* clinical isolates, aiming to learn principles that can be applied as other species and strains are phenotyped. As part of a collaboration with the John Radcliffe Hospital microbiology laboratory in Oxford, we were able to access *E. coli* isolates collected and whole-genome sequenced for an epidemiological study of *E. coli* bloodstream infections [170].

In previous work from this collaborative project, single-cell phenotyping with DNA and membrane stains could be used to provide equivalent information to a 24-hour growth curve for *E. coli* treated with ciprofloxacin at 20× EUCAST for 30 minutes — the proportion of cells classified as susceptible by the neural network corresponded to the MIC as measured by the growth curve [93]. In that work, a model trained on *E. coli* MG1655 was directly applied to clinical isolates with the expectation that susceptible cells would look like ciprofloxacin-treated MG1655, while resistant cells would look like untreated MG1655.

Ciprofloxacin is a commonly used fluoroquinolone antibiotic that is a first-line therapy for hospitalised patients with urinary tract infections and a third-line therapy for patients with community-acquired simple cystitis who cannot tolerate other antibiotics [205]. To demonstrate the potential to use ribosome phenotypes instead of DNA phenotypes for the classification of susceptible and resistant clinical strains, we chose to narrow our focus to ciprofloxacin. This enabled us to characterise ribosome phenotypes for a clinically important antibiotic and allowed us compare our ribosome phenotypes to the DNA phenotypes found in previous work [182].

Here, we show how the ribosome phenotyping method can be applied to classify *E. coli* clinical isolates as susceptible or resistant to ciprofloxacin. First, we apply the ribosome phenotype model trained on *E. coli* MG1655 treated with ciprofloxacin to our clinical isolates. Then, we move to training and testing our models on the clinical isolates themselves. Finally, we perform simulations to estimate how many cells would need to be tested with the ribosome phenotyping method to classify a clinical sample with >99% confidence.

4.2 Model trained on lab strain has variable accuracy on clinical isolates

In previous work [93], a CNN trained on the DNA and cell membrane phenotypes of untreated and ciprofloxacin-treated *E. coli* MG1655 was used successfully to classify ciprofloxacin-susceptible and resistant *E. coli* clinical isolates. In their dataset, the DNA and cell membrane of resistant clinical isolates resembled those of untreated MG1655 and susceptible clinical isolates resembled ciprofloxacin-treated MG1655. We hypothesised that the same domain transfer would be possible with the ribosome-trained CNN.

For our clinical isolates dataset, we tested three susceptible strains (S1, S2, S3) and three resistant strains (R1, R2, R3) with different mutations in ciprofloxacin resistance-associated genes (Table 2.2). Each of the susceptible strains has a mutation in one of these genes, whereas the resistant strains all have three or more resistance-associated mutations. For example, strain R2 has two mutations in *GyrA*, which encodes the A subunit of DNA gyrase, and three mutations in the genes that encode DNA topoisomerase IV (two mutations in *parC*, for the A subunit, and one in *parE*, for the B subunit). Each of these mutations has been associated with increased ciprofloxacin resistance in previous work [206, 207, 208, 209]. Instead of the 1× EUCAST concentration used to treat MG1655, each isolate was treated with ciprofloxacin at the higher concentration that was shown to robustly induce phenotypic changes within 30 minutes in the previous single-cell phenotyping work (10 mg/L, 20× EUCAST breakpoint) [93].

Following ciprofloxacin treatment, each of the susceptible *E. coli* strains (S1, S2, S3) demonstrated a ribosome phenotype similar to ciprofloxacin-treated *E. coli* MG1655, with a compact, central nucleoid region that we can detect indirectly because it results in low ribosome density in the central region (Figure 4.1a: S1). However, while some cells from the resistant strains resembled untreated *E. coli* MG1655 as expected (Figure 4.1a: R2) and were classified correctly, most cells from the resistant strains showed a different ciprofloxacin response, wherein the cells were elongated but retained diffuse nucleoid regions that spanned the cell

length (Figure 4.1a: R1). When the ciprofloxacin ribosome phenotype classifier from *Chapter 3* was applied to this dataset, it frequently classified these elongated cells as ciprofloxacin-susceptible.

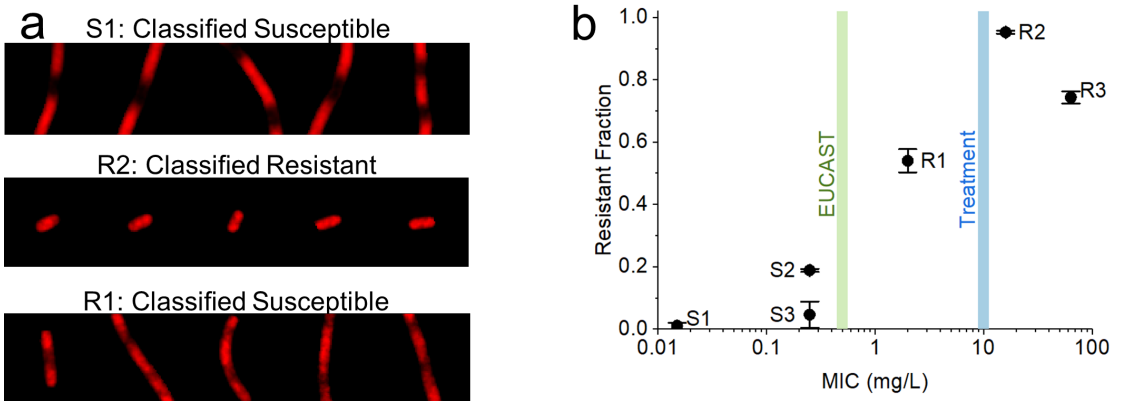


Figure 4.1: Classifier trained on *E. coli* MG1655 applied to clinical isolates has variable accuracy.
(a) Representative ribosome phenotypes from S1 cells correctly classified as susceptible, showing an elongated phenotype and compact nucleoid; R2 cells correctly classified as resistant, showing a phenotype more typical of untreated E. coli MG1655; and R1 cells incorrectly classified as susceptible, showing an elongated phenotype and diffuse nucleoid. Each 64×64 single-cell image is 7.5×7.5 μm^2.
(b) The fraction of cells in the sample called Resistant by the MG1655 ciprofloxacin classifier is plotted against the MIC of the strain (mg/L) on a logarithmic scale. Error bars represent the 95% confidence interval, equal to ± 2 standard deviations of the mean. The test dataset is composed of images from the six *E. coli* clinical isolates treated with 10 mg/L ciprofloxacin for 30 minutes (Table 2.2). The EUCAST breakpoint (0.5 mg/L) and the treatment condition (10 mg/L) are shown with shaded vertical lines. All strains with an MIC below the EUCAST breakpoint have a resistant fraction less than 0.2, whereas there is no clear relationship between the MIC and the resistant fraction for the resistant strains.

This led to variable accuracy when the MG1655 classifier was applied to our dataset of ciprofloxacin-treated clinical isolates (Figure 4.1b). The MG1655 classifier recognised ciprofloxacin-susceptible phenotypes with high accuracy. For S1 and S3, it classified cells as ciprofloxacin-susceptible with 99.0 ± 1.0% and 95.4 ± 4.2% accuracy, respectively. The accuracy was lower for S2 cells (81.2 ± 0.5%), possibly because these cells are less likely to be elongated than ciprofloxacin-treated MG1655 (Mann-Whitney non-parametric hypothesis test $p<0.05$; Figure 4.2: S2).

For resistant isolates, the MG1655 classifier was less reliable. Compared to untreated MG1655, the ciprofloxacin-treated resistant isolates had similarly diffuse

nucleoid regions but were elongated (Figure 4.2: R1, R2, R3). The MG1655 classifier classified 95.2 ± 5.8% of R2 cells as ciprofloxacin-resistant, but only 54.0 ± 3.7% of R1 cells and 74.3 ± 1.9% of R3 cells (Figure 4.1b). Representative cell images show that the R1 cells that were misclassified as susceptible had an elongated cell shape and a diffuse nucleoid (Figure 4.1a: R1), whereas R2 cells that were correctly classified as resistant had a shape and nucleoid phenotype more similar to the untreated MG1655 (Figure 4.1a: R2). The variability in accuracies for R1, R2, and R3 clearly shows that the ribosome phenotypes resulting from ciprofloxacin treatment at this concentration are too diverse to be reliably recognised by the MG1655 classifier, especially for resistant strains that develop an elongated shape with a diffuse nucleoid.

4.3 Developing a susceptible-resistant clinical isolate classifier

Given the variable accuracy of the MG1655 classifier on the clinical isolates in our dataset, we hypothesised that a CNN trained to learn the more variable ribosome phenotypes of susceptible and resistant isolates would have more robust performance. In addition, training the model on images of clinical isolates more closely approximates the real-world application, in which a diagnostic platform could learn from the images it sees.

For this model, the training dataset was composed of 34,205 images of clinical *E. coli* strains isolated from bloodstream infections and treated with 20× EUCAST ciprofloxacin for 30 minutes. As was done for the MG1655 model, the fluorescence images of the ribosome phenotypes were segmented, zero-filled, and augmented for training. We trained two six-strain models as a cross validation to check consistency on different biological replicates. For each of the six-strain susceptible-resistant models, two biological replicates were used for the training and validation datasets and one was used for a holdout test to assess the model's accuracy on unseen data. In total, the testing dataset comprised 28,448 cells from three susceptible and three resistant isolates (S1, S2, S3, R1, R2, R3 in Table 2.2).

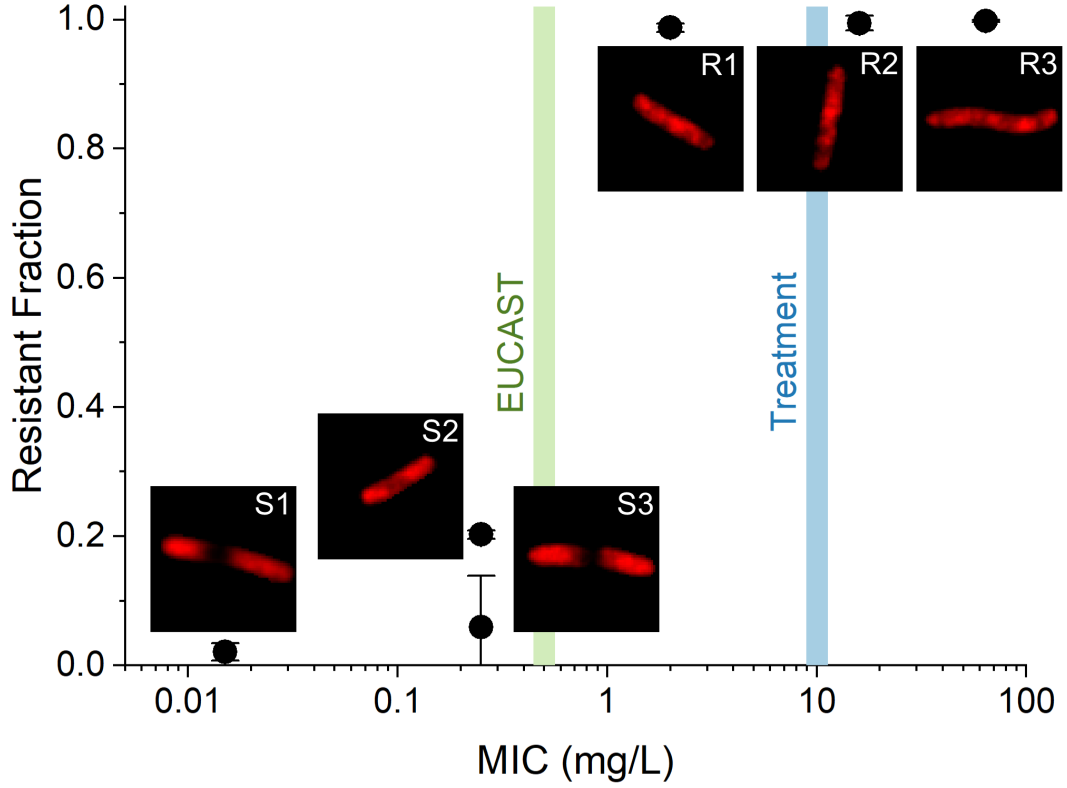


Figure 4.2: Susceptible (MIC < EUCAST breakpoint) and resistant (MIC > EUCAST breakpoint) *E. coli* clinical isolates can be differentiated by the fraction of cells called resistant by the model. The fraction of cells in the sample called resistant by the susceptible-resistant classifier (Resistant Fraction) is plotted against the MIC of the strain (mg/L) on a logarithmic scale, with error bars indicating the 95% confidence interval of the mean on two biological replicates. The test dataset is composed of holdout images, previously unseen by the classifier, from each clinical isolate. The EUCAST breakpoint (0.5 mg/L, green) and the treatment condition (10 mg/L, blue) are shown with shaded vertical lines. All strains with an MIC below the EUCAST breakpoint have a resistant fraction less than 0.2, whereas the fraction classified resistant is nearly 1.00 for the strains with an MIC above the EUCAST breakpoint. Representative, correctly classified images of ribosome phenotypes from each of the clinical isolates are shown for each point, labelled by strain (S1, S2, S3, R1, R2, R3). Each 64×64 single-cell image is 7.5×7.5 μm^2.

The susceptible-resistant CNN learned to identify phenotypes associated with ciprofloxacin-treated susceptible and resistant strains with a single-cell balanced accuracy of 94.9 ± 0.3%. It displayed an accuracy of 99.3 ± 0.2% in classifying resistant cells and an accuracy of 90.5 ± 0.5% in classifying susceptible cells (Figure 4.3). These extremely high single-cell accuracies lead to very high confidence in classifying a population of *E. coli* as antibiotic-susceptible, antibiotic-resistant, or

a mixture of the two, as would happen in a clinical testing scenario. Stratifying the classifications by strain, we found that our classification accuracy remained greater than or equal to 80% for all six clinical isolates (Figure 4.4).

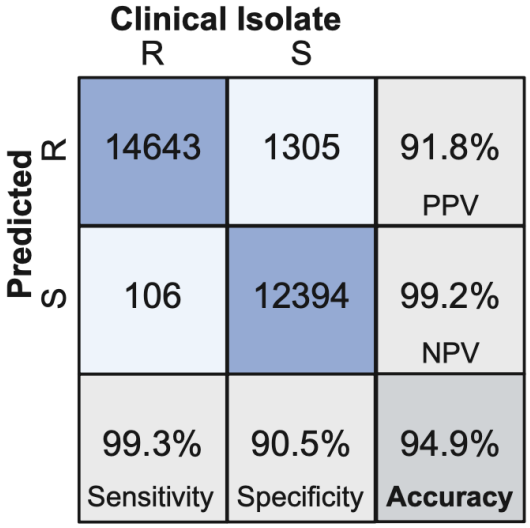


Figure 4.3: Confusion matrix for the susceptible-resistant clinical isolates classifier. The confusion matrix for the ciprofloxacin-susceptible and ciprofloxacin-resistant classifier trained on 6 clinical isolates and tested on a holdout, unseen dataset of biological replicates from those 6 strains. The testing dataset was composed of 28,448 single-cell ribosome fluorescence images. See 2.9 for details on Accuracy, Sensitivity, Specificity, PPV, and NPV.

In the context of a clinical test, the relationship between the fraction of cells classified as resistant and the strain's MIC as determined by broth micro-dilution is important for test validation and trustworthiness. When examining the relationship between a given strain's MIC and the fraction of cells classified as resistant, $<$ 9% of cells from strains with an MIC less than the EUCAST breakpoint were called resistant, whereas nearly 100% of cells from strains with an MIC above the EUCAST breakpoint were called resistant (Figure 4.3).

Compared to the application of MG1655 model on the clinical isolates dataset, the model trained on susceptible and resistant strains had similar or higher accuracy for all susceptible strains and was more accurate on all resistant strains ($+44.8 \pm 4.5\%$ for R1, $+4.3 \pm 1.6\%$ for R2, $+25.5 \pm 2.1\%$ for R3) (Figure 4.5). This result

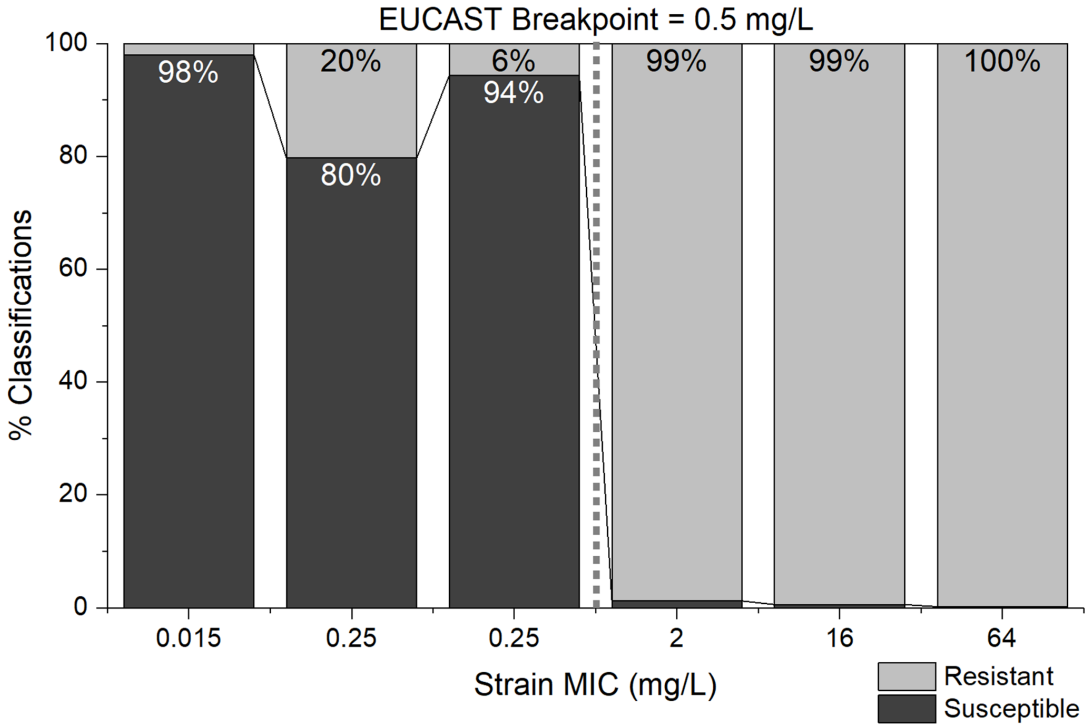


Figure 4.4: The susceptible-resistant model can differentiate each of the susceptible (MIC < 0.5 mg/L) and resistant (MIC > 0.5 mg/L) clinical isolates. The percentage of cells in the test dataset classified as susceptible (dark grey) or resistant (light grey) is shown for each of the clinical isolates, ordered by the MIC of the strain. For strains with an MIC below the EUCAST breakpoint (0.5 mg/L), the % of susceptible classifications is greater than or equal to 80% (98%, 80%, 94%). For strains with an MIC above the EUCAST breakpoint, the % of resistant classifications approaches 100% (99%, 99%, 100%).

was expected, given the greater domain specificity of the susceptible-resistant model, however it reinforces the importance of using clinical isolates when available.

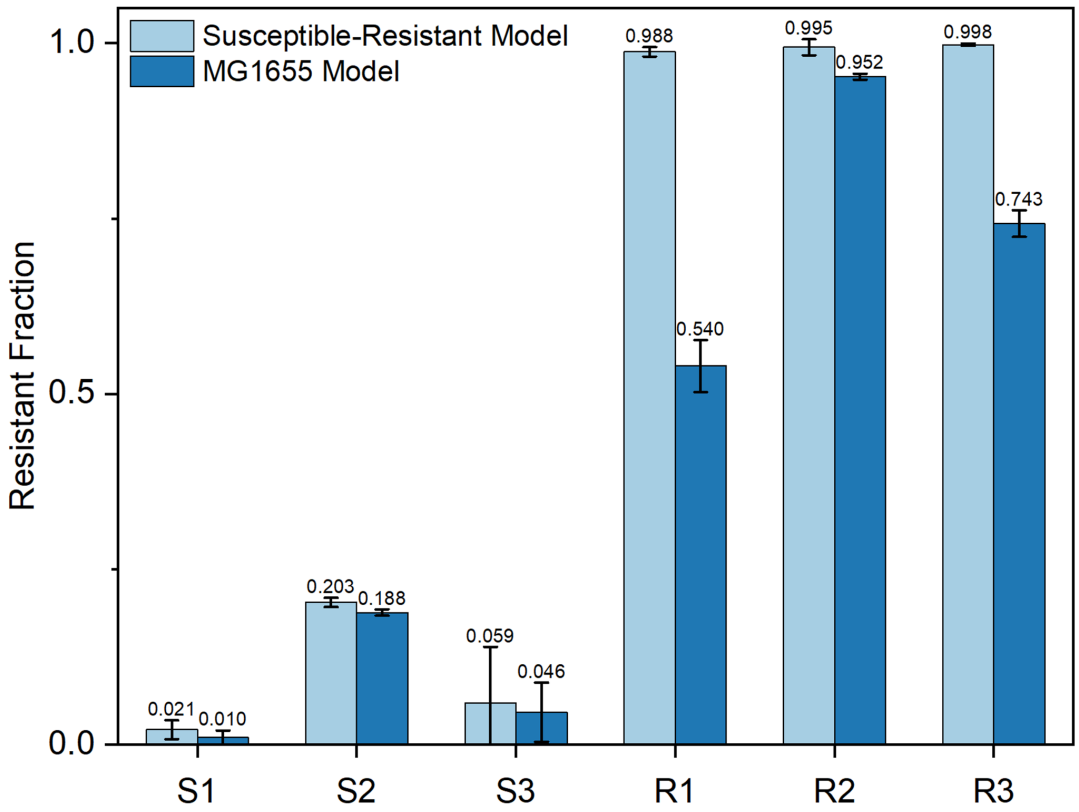


Figure 4.5: The susceptible-resistant classifier trained on clinical isolates is more accurate than the MG1655-trained classifier. The fraction of cells classified as resistant (Resistant Fraction) is shown for each of the clinical strains (S1, S2, S3, R1, R2, R3), based on classifications by the susceptible-resistant model (light blue) or the MG1655 model (blue). The error bars denote the 95% confidence interval of the biological replicates. The test dataset is composed of images from the six *E. coli* clinical isolates treated with 10 mg/L ciprofloxacin (20× EUCAST) for 30 minutes. Compared to the MG1655 model, the susceptible-resistant model had similar or higher accuracy on all sensitive strains and was more accurate on all resistant strains.

4.4 Necessary sample size for testing

Within a diagnostic workflow, a certain number of cells will need to be isolated from the clinical sample in order to give the test sufficient power to confidently classify the sample as antibiotic-resistant, susceptible, or multiple strains of mixed resistance. To calculate this necessary sample size, we simulated cell samples with 100% resistant cells, 100% susceptible cells, or a 50-50 mixture, using the sensitivity and specificity of our assay. (See *Sample size simulation* 2.11 for discussion of the methods.)

The results of these simulations suggest that given its high sensitivity (99.3%) and

specificity (90.5%), our susceptible-resistant classifier has the power to differentiate a 100% resistant sample from a 100% susceptible sample with 99% confidence after sampling as few as 2 cells (Figure 4.6). With a reasonable sample size of 10 to 100 bacteria isolated from a clinical specimen using a microfluidic device [210], the confidence level of our prediction would increase. In the case of a mixed infection or contaminated sample, we could differentiate a mixed sample from a resistant sample (7 cells) or a mixed sample from a susceptible sample (12 cells) with the same level of confidence (Figure 4.6).

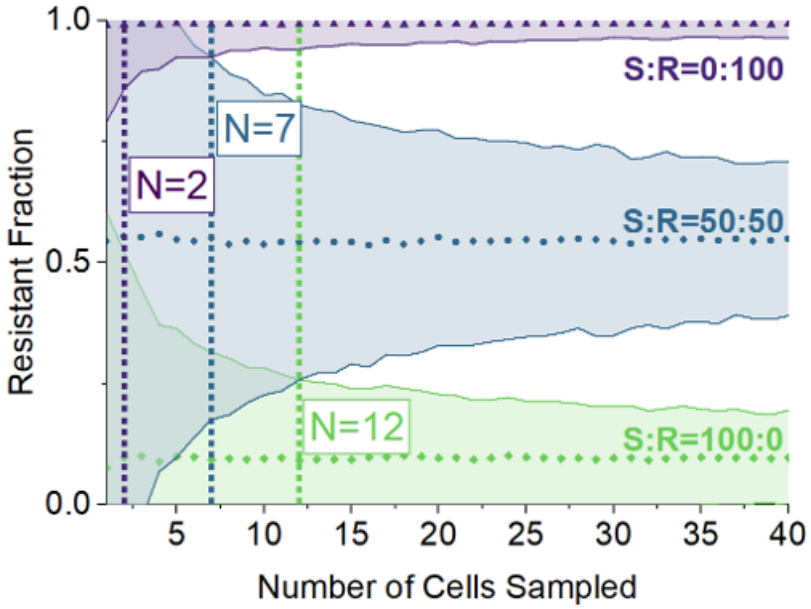


Figure 4.6: Necessary sample size simulation for the susceptible-resistant classifier. The number of cells necessary to classify a sample as coming from a population of susceptible or resistant bacteria. Simulated samples of different susceptible:resistant ratios (S:R) were transformed through the sensitivity and specificity of the susceptible-resistant classifier to determine the minimum number of cells necessary to differentiate them. Here, we plot the mean Resistant Fraction and 99% confidence interval after 1,000 trials with samples ranging from 1 to 40 cells sampled (N) for susceptible:resistant ratios of 0:100 (purple triangles), 50:50 (blue circles), and 100:0 (green diamonds). As the number of cells sampled increases, the confidence interval of the Resistant Fraction narrows. Susceptible samples can be differentiated from resistant samples with a sample of 2 cells (purple dotted line). A mixed sample can be differentiated from a resistant sample with 7 cells (blue dotted line) or from a susceptible sample with 12 cells (green dotted line). The confidence interval for resistant cells is narrower than that of susceptible cells because the classifier is more sensitive than it is specific. See *Sample size simulation* (2.11) for more details.

4.5 Model performance on unseen strains and antibiotic concentrations

We found variable accuracy when applying the model trained on MG1655 to the clinical isolates (Figures 4.1 and 4.5). Therefore, we sought to investigate the ability of a ribosome phenotype CNN to generalise to previously unseen strains and antibiotic concentrations.

We trained a CNN on just one susceptible strain (S2) and one resistant strain (R4) (Table 2.2). These strains showed characteristic ciprofloxacin-susceptible and ciprofloxacin-resistant phenotypes after 30 minutes of treatment with 1× EUCAST ciprofloxacin (0.5 mg/L) (Figure 4.7a). At this treatment concentration, the resistant strain (R4) showed two, diffuse nucleoid regions, as was often seen in untreated *E. coli* MG1655 (Figure 3.2). When treated at 1× EUCAST, this strain was less filamentous than the resistant strains treated at 20× EUCAST. The susceptible strain (S2) showed a compact, central nucleoid, but was not as elongated as some of the ciprofloxacin-treated MG1655 or clinical isolate cells (Figures 3.2 and 4.2).

The model was trained on a dataset of 2,888 cell images with an 80:20 training-validation split, using the same pipeline as described in *Chapter 2: Neural network training* (2.8). The 1× EUCAST model was tested on the same holdout dataset as the susceptible-resistant model, composed of unseen cells from the six clinical isolates treated at 20× EUCAST (10 mg/L) for 30 minutes. The 1× EUCAST model was tested on each of the 3 biological replicates in the 20× EUCAST dataset and performed with an average accuracy of 73.8 ± 5.3% on cells from susceptible strains and an average accuracy of 89.6 ± 3.8% on cells from resistant strains (Figure 4.7b). Although lower in accuracy than the six-strain 20× EUCAST model, the two-strain 1× EUCAST model demonstrates an ability to reliably differentiate susceptible and resistant cells with relatively high accuracy, even given images of different strains treated at a different concentration of ciprofloxacin. In this case, the model was also learning the ribosome phenotype on cells that were shorter than the cells treated at 20× EUCAST (Figures 4.7a, 4.2). This demonstrates

the robustness of the ribosome phenotype classification method, so long as the model has seen sufficiently similar training data.

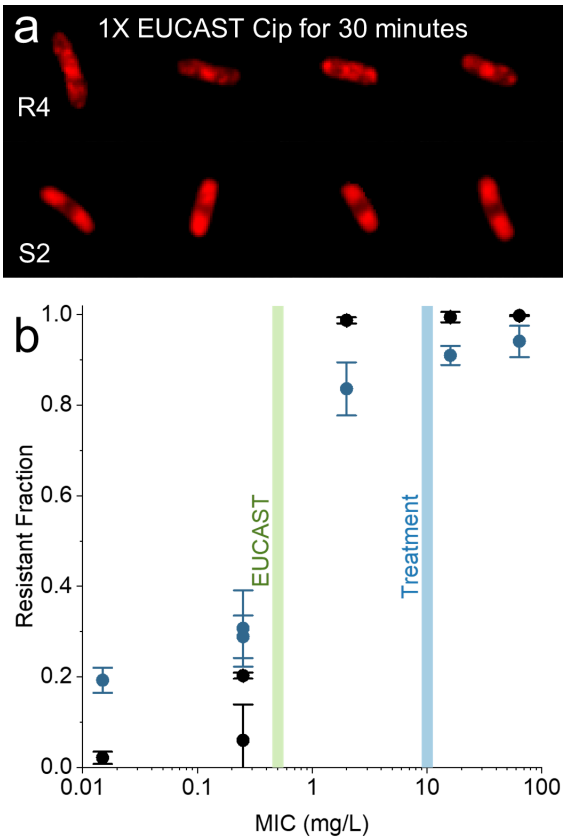


Figure 4.7: Model performs with >80% accuracy when trained on cells treated with 1× EUCAST and tested on cells treated with 20× EUCAST.
(a) Representative, correctly classified images of the ribosome phenotypes of strains R4 and S2 treated at 1× EUCAST (0.5 mg/L) for 30 minutes. Each 64×64 single-cell image is 7.5×7.5 μm^2.
(b) The 20× EUCAST classifier trained on 6 clinical isolates of *E. coli* (black circles) is compared to the 1× EUCAST classifier trained on 2 clinical isolates of *E. coli* (R4, S2) (blue circles). The error bars denote the 95% confidence interval of the mean accuracy. The 20× EUCAST model was tested on 28,448 holdout test images from 2 biological replicates of the 6 clinical *E. coli* isolates treated at 20× EUCAST ciprofloxacin for 30 minutes. The 1× EUCAST model was tested on 50,681 unseen images from the same 20× EUCAST dataset. For every isolate, the 20× EUCAST model is more likely to call resistant cells resistant and less likely to call susceptible cells resistant. However, the 1× EUCAST model maintains an accuracy of 73.8 ± 5.3% on cells from susceptible strains and an accuracy of 89.6 ± 3.8% on cells from resistant strains, despite being trained on images of cells treated at a different concentration and classifying previously unseen strains.

Conversely, we tried applying the model trained on our 20× EUCAST dataset to the two isolates treated at 1× EUCAST. The model performed well in classifying

R4. At 1× EUCAST, this strain had a similar phenotype to the resistant strains treated at 20× EUCAST, and most of the misclassified cells were very short with a central nucleoid region (Figure 4.8a: R4). On the resistant strain, the 20× EUCAST model had an accuracy of 93.0 ± 12.3% (Figure 4.8b). However, the model was unable to reliably classify S2. With 1× EUCAST treatment, this strain had a much less elongated phenotype than the other susceptible strains (Figure 4.8a: S2). Many of the cells also exhibited a less defined nucleoid region. The model was only 52.3 ± 11.0% accurate in classifying S2.

The S2 strain was also the most challenging to classify for the 20× EUCAST classifier when tested on 20× EUCAST cells (80% accuracy, Figures 4.2, 4.4). Looking at the images of misclassified cells, it appears that many of these misclassifications are the result of intrinsic heterogeneity in the ribosome phenotype of S2 (Figure 4.8a: S2).

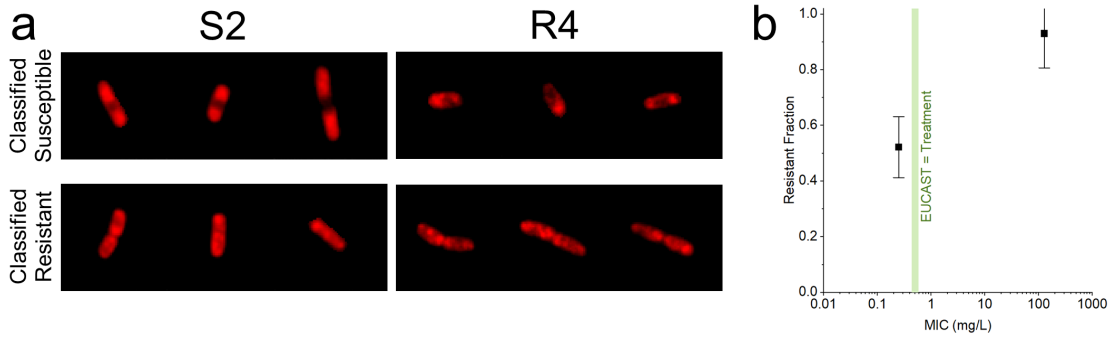


Figure 4.8: Classifier trained on clinical isolates treated at 20× EUCAST tested on isolates treated at 1× EUCAST.
(a) Representative ribosome phenotypes from S2 cells correctly classified as susceptible, showing an elongated phenotype and compact nucleoid; S2 cells incorrectly classified as resistant, showing shorter cells and a less defined nucleoid region; R4 cells incorrectly classified as susceptible, showing a compact central nucleoid; and R4 cells correctly classified as resistant, with diffuse nucleoid regions. Each 64×64 single-cell image is 7.5×7.5 μm^2.
(b) The fraction of cells in the 1× EUCAST sample called Resistant by the 20× EUCAST ciprofloxacin classifier is plotted against the MIC of the strain (mg/L) on a logarithmic scale. Error bars represent the 95% confidence interval, equal to ± 2 standard deviations of the mean. The test dataset is composed of images of S2 and R4 treated with 0.5 mg/L ciprofloxacin for 30 minutes. The EUCAST breakpoint and treatment concentration (0.5 mg/L) is shown with a vertical line. The classifier had high accuracy on R4, but was unable to reliably classify S2.

4.6 Discussion

Here, we aimed to demonstrate how the ribosome phenotyping method could be applied in *Escherichia coli* clinical isolates. For this study, we aimed to use a selection of clinical isolates to learn principles that could be applied when studying the single-cell antibiotic response of other strains or species.

We found that in our treatment conditions, phenotype development was more complicated than expected. The previous deep-learning-based AST found a close relationship between the proportion of cells classified as antibiotic-susceptible and the MIC of the strain [93]. We also found a strong relationship between the MIC of the clinical isolate and its morphology after ciprofloxacin treatment; ciprofloxacin-susceptible cells developed compact nucleoid regions whereas resistant cells became elongated but maintained diffuse nucleoid regions. However, while the previous method used untreated *E. coli* MG1655 as a proxy for the resistant phenotype, we found that our MG1655-trained classifier had unreliable accuracy when classifying our resistant strains with a diffuse nucleoid but elongated cell shape. Our resistant isolates, cultured and treated at 37°C, exhibited a more filamentous morphology than those in the previous study, which could explain this difference in results [93].

These results, and the results we found when applying the 1× EUCAST model to the 20× EUCAST data and vice versa, highlight the importance of training models on data from treatment conditions and strains similar to their ultimate use case. Therefore, it is best to characterise antibiotic response phenotypes in clinical isolates when possible, although lab strains can be used as a starting point. When applying models to clinical isolates treated at different concentrations than the training data, there is some transferability, but it depends on the similarity of the training data to the tested phenotypes. Here, we directly applied trained models to new domains to test their performance and generalisability. It would also be possible to use transfer learning, where a model trained on one domain is retrained in a new domain, often with a smaller dataset than the original training dataset. In this case, a model trained on a lab strains could be augmented with data from clinical isolates.

Both the *E. coli* MG1655 model and the clinical isolates susceptible-resistant model performed differently when classifying strains S2 and S3, even though both strains have the same MIC as measured by broth micro-dilution (0.25 mg/L ciprofloxacin). Around 94% of S3 cells were classified as susceptible, but only 80% of S2 cells. Most of the misclassified S2 cells were smaller and had a less defined nucleoid region than was seen in the S1 and S3 strains, while the correctly classified cells had a more typical ciprofloxacin-susceptible phenotype. S2 is also the only isolate we studied with a mutation in *marR* (Table 2.2). However, although *marR* encodes for a regulatory locus associated with multi-drug resistance, the S3N mutation found in S2 results in wild-type expression levels [211, 212]. The mechanism behind the variable phenotypes seen in S2 is probably caused by another AMR-associated gene not investigated by the Lipworth study [170].

When we moved to using a model trained on the 6 clinical isolates, we achieved an average single-cell classification accuracy of 95.0% on holdout test images. For this model, the difference in sensitivity and specificity was influenced by the low classification accuracy for clinical isolate S2. The other susceptible isolates had classification accuracies above 94% (97.9% for S1 and 94.1% for S3). Nearly 100% of resistant cells were correctly classified by the model. A high accuracy on resistant cells could be especially useful in clinical practice. In several clinical trials, rapid ASTs led to significantly faster times to effective antibiotic therapy in patients with resistant organisms, even when they did not observe a difference in the overall study populations [75, 213, 214]. In the context of a diagnostic test, our current level of accuracy means that with only 2 cells, we could differentiate a susceptible from a resistant sample with 99% confidence and could identify mixed infections with a sample of 12 bacterial cells. On a realistic scale of between 10 to 100 bacteria captured from a dilute sample such as blood or cerebrospinal fluid [160, 90], this level of accuracy could enable confident diagnosis even in less ideal imaging conditions. Our results with the ribosome phenotyping method provide additional validation for the use of CNNs to detect single-cell changes associated with antibiotic susceptibility.

To further improve the accuracy on a wide variety of isolates, more data collection is needed. The lower accuracy of the model on clinical isolate S2 demonstrates how an unexpected morphology causes confusion for the current model. As the model has access to more varied susceptible and resistant ribosome phenotypes, its ability to generalise should improve [134]. Data augmentation strategies could also be further optimised, especially to improve the model's ability to classify blurred cells, but, as isolate S2 demonstrates, the ribosome phenotype of an unknown strain cannot necessarily be predicted *a priori*. Providing the model with more data, particularly from a wider selection of clinical isolates and treatment conditions, would be the best way to improve real-world accuracy.

4.7 Conclusions

In this chapter, we applied ribosome phenotyping to 3 ciprofloxacin-susceptible and 3 ciprofloxacin-resistant clinical isolates of *E. coli*. A model trained on the ribosome phenotypes of these isolates was able to classify susceptible cells with 91% single cell accuracy and resistant cells with 99% accuracy, for an overall balanced accuracy of 95%. A model with this level of performance would be able to determine whether bacteria came from an antibiotic-susceptible or resistant strain by sampling as few as 2 cells. We also demonstrated the performance of ribosome phenotype models when transferred to different domains, applying a model trained on a lab strain to clinical isolates and applying models to classify cells treated at a different ciprofloxacin concentration than the training data.

5

Citizen science for image feature analysis of antibiotic phenotype misclassifications

Contents

5.1 Introduction 78
5.2 Methods 81
5.2.1 Image dataset 81
5.2.2 Accuracy analysis 84
5.2.3 Image feature analysis 84
5.3 Development of the Infection Inspection project 86
5.3.1 User participation 87
5.3.2 Accuracy analysis 89
5.4 Results of feature analysis of sensitive and resistant phenotypes 90
5.4.1 Differences in DNA morphology drive confusion 90
5.4.2 Incorrectly classified images have more ambiguous phenotypes 94
5.4.3 Feature importance in ciprofloxacin-sensitive and resistant cells 97
5.4.4 Phenotypic differences in resistant *E. coli* treated with high ciprofloxacin concentrations 97
5.5 Discussion 99
5.5.1 Feature analysis 99
5.5.2 Volunteer participation 100
5.5.3 Limitations 101
5.5.4 Future directions 102
5.6 Conclusions 103

In the previous chapter, we trained models to classify antibiotic-susceptible and resistant *E. coli* from clinical isolates by changes in their cellular phenotype. However, a certain number of cells from each sample are misclassified. In this chapter we use data from our citizen science project Infection Inspection to explore the image features associated with misclassification of ciprofloxacin-treated *E. coli* isolated from bloodstream infections. The work presented in this chapter is adapted from "Infection Inspection: using the power of citizen science for image-based prediction of antibiotic resistance in *Escherichia coli* treated with ciprofloxacin" [215]. The Infection Inspection project was a collaborative effort. The dataset was produced by Alexander Zagajewski and curated by Piers Turner [93, 182]. Conor Feehily and I developed the Zooniverse project. More details can be found in the *Statement of Originality*. In this chapter, we call the clinical isolates ciprofloxacin-sensitive instead of ciprofloxacin-susceptible to match the terminology used in the Infection Inspection project.

5.1 Introduction

In Chapter 2, we profiled the canonical phenotypes displayed by antibiotic-sensitive *E. coli* in response to antibiotic treatment for several classes of antibiotics. In Chapter 3, we saw that these phenotypes can vary between strains and between cells from the same culture. The causes of this phenotypic heterogeneity are an area of active research, and are known to include differential expression of efflux pumps, cell cycle stage, point mutations in target genes, plasmid-borne genes, persister cells, and viable but non-culturable cells [86, 63, 64, 65, 209]. Here, we focus on the practical ramifications of phenotypic heterogeneity for the CNN-based classification of AMR. What features of the cell morphology are common to correctly-classified cells? What features make misclassification more likely? Are human classifiers able to leverage contextual information to identify more nuanced features than the machine? To seek answers to these questions, we developed a

citizen science project called Infection Inspection to recruit volunteers to classify images of ciprofloxacin-treated clinical isolates.

Citizen science is a research method that integrates public outreach with data collection, typically with members of the public collecting or analysing data related to the natural world as part of a collaborative project with professional scientists [216, 217]. Citizen science collaborations between volunteers, often called citizen scientists, and research teams can play an important role in educating the public about scientific concepts and have been instrumental in recognising complex patterns within biological data, starting with research in ecology and extending to various biological fields, including protein folding, DNA sequence alignment, electron microscopy, and microbiology [218, 219, 220]. These projects enable individuals of diverse backgrounds and expertise levels to contribute to scientific data collection and analysis, empowering them to actively advance and acquire knowledge in various disciplines. Public involvement broadens the spectrum of available data, perspectives, and ideas, leading to more comprehensive and innovative research outcomes. Successful examples, such as the Great Backyard Bird Count and the Merlin Bird ID app [221, 222], demonstrate citizen science's potential to enable large-scale data collection and analysis, raise awareness, inspire future scientists, and promote environmental and civic responsibility.

The public can be effectively engaged in citizen science projects using various strategies, including hosting events, utilising social media platforms, partnering with educational institutions and community organisations, and offering training and educational resources, but most of these approaches engage only tens to hundreds of individuals. Online platforms that simplify access for large-scale, global public engagement in targeted or diverse citizen science projects include Zooniverse [223, 224], SciStarter [225] and Foldit [226], amongst others [227]. For participants, Zooniverse offers a unique and engaging way to learn about science, participate in real research, and connect with like-minded individuals [228]. Zooniverse is currently the largest and most popular citizen science platform, with more than 1 million volunteers [223].

A previous project hosted on Zooniverse, Bash the Bug [220], successfully engaged citizen scientists to look at images of bacterial growth and identify the sample's resistance to antimicrobial drugs. This demonstrated how citizen science can be used for antimicrobial resistance research and the development of novel diagnostic tools. We were interested in what insights could be gained from a citizen science collaboration with volunteers providing human interpretations of images of ciprofloxacin-treated sensitive and resistant *E. coli* with DNA and membrane stains. We hypothesised that a large-scale volunteer analysis of these images could provide insights into why about 20% of cells were misclassified by our deep learning model [93, 182].

We therefore developed a project on Zooniverse called Infection Inspection to leverage the power of citizen scientists towards optimising our phenotypic classification method, and to engage the public in an antibiotic resistance-focused project. We first trained volunteers to recognise cellular phenotypes associated with ciprofloxacin-sensitive and ciprofloxacin-resistant *E. coli*, and then used their classifications to learn which features facilitate accurate classification, and which lead to ambiguity and misclassifications (Figure 5.1).

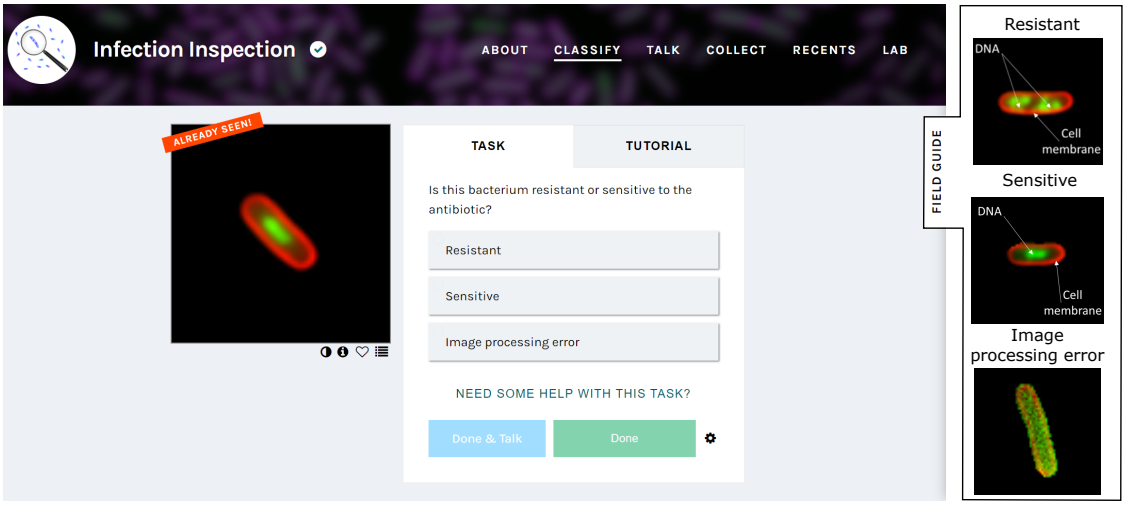


Figure 5.1: Infection Inspection workflow. The Classify task page for users of Infection Inspection presented users with an image of a bacterial cell and asked them to select from one of three options to classify the image. Accessing the field guide provided examples of each cell type, illustrated here with selected images of a Resistant cell, a Sensitive cell, and an Image processing error. Adapted from figure made by Conor Feehily.

Ciprofloxacin is a fluoroquinolone antibiotic that inhibits enzymes involved in bacterial DNA replication and repair [229]. In sensitive bacteria, this results in the compaction of the nucleoid and blockage of cell division, leading to filamentous *E. coli* with centralised DNA regions [230]. In more resistant strains, this phenotype develops more slowly, if at all, so the DNA appears more diffuse even after 30 minutes of antibiotic exposure.

Our aim was to use the volunteer's classifications and misclassifications to see if humans were able to identify more nuanced features than our machine, to make our machine learning-based classifier more robust to atypical phenotypes, while simultaneously educating the citizen scientists about antibiotic resistance. We found that our machine learning model picks up a greater breadth of features, but by investigating the common features of images, we identified the key features necessary for *E. coli* to be classified as ciprofloxacin-sensitive or resistant.

5.2 Methods

5.2.1 Image dataset

The Infection Inspection project dataset was made up of 49,074 individual images of ciprofloxacin-treated *E. coli* cells generated by Alexander Zagajewski for previous work [93, 182]. Before imaging, all bacteria were chemically fixed and stained using 4',6-diamidino-2-phenylindole (DAPI) as the nucleic acid stain and Nile Red as the membrane stain. The dataset for the first upload was composed of 11,074 256x256 Red-Green-Blue (RGB) images of *E. coli* cells from 5 clinical strains (EC1, EC2, EC3, EC5, and EC6, reported previously [93] and detailed in Table 5.1), with strains defined as ciprofloxacin-resistant (minimum inhibitory concentration [MIC] $>$ 0.5 mg/L) or ciprofloxacin-sensitive (MIC $\leq$ 0.25 mg/L) using European Union Council on Antimicrobial Susceptibility Testing (EUCAST) breakpoints [169]. The minimum inhibitory concentration is a standard metric of antimicrobial resistance and is the lowest concentration of antibiotic that inhibits growth of bacteria in an overnight culture [231]. The MICs of the strains were calculated empirically

by E-test strip (Liofilchem), or, where the MIC exceeded the maximum range of the strip, by a 1:1.5 broth dilution of the antibiotic [93].

In the initial dataset, all strains were treated at 10 mg/L ciprofloxacin for 30 minutes; this concentration, well in excess of the EUCAST breakpoint, was chosen to induce an antibiotic response within the short, 30-minute treatment period. EC4 was excluded because its MIC is equal to the treatment concentration, and we define the expected phenotype based on whether the treatment concentration was greater than or less than the MIC (see 5.3.2). A second dataset of 38,000 images included two of the *E. coli* strains (EC1 and EC3) treated at 9 concentrations ranging above and below their MICs (from 16 mg/L to 0.001 mg/L ciprofloxacin) for 30 minutes.

All bacteria were imaged in an automated workflow as agarose-mounted samples in phosphate-buffered saline (PBS) on a Nanoimager-S fluorescence microscope (Oxford Nanoimaging) using the multiple acquisition capability of the microscope with autofocusing on each field of view. The image segmentation for background removal was done with an optimised model of Mask-RCNN adapted from a standard implementation [93, 232, 233].

Strain	**EC1**	**EC2**	**EC3**	**EC5**	**EC6**
Year	2016	2015	2018	2017	2014
Cip S/R/I	S	S	I	R	R
Cip MIC (mg/L)	0.008	0.03	0.5	8	72
Number of Images	15,900	3,883	21,529	2,882	4,877
Correct / Total Classifications	187,729 / 280,469	41,013 / 53,251	238,057 / 377,612	41,732 / 45,599	42,650 / 68,575
Zooniverse Accuracy	66.9±0.02%	77.0±0.4%	63.0±0.2%	91.5±0.3%	62.2±0.4%
Model Accuracy	84.9±0.6%	81.6±1.2%	71.0±0.6%	86.2±1.3%	76.3±1.2%

Table 5.1: *Escherichia coli* clinical isolates used for the Infection Inspection project. Strains are listed, along with their year of collection; ciprofloxacin-sensitive, resistant, or intermediate (Cip S/R/I); ciprofloxacin MIC (mg/L); number of images; correct classifications and total classifications; Zooniverse user accuracy (95% confidence interval); and model accuracy (95% confidence interval). All strains are Escherichia coli isolated from bloodstream infections in the United Kingdom. EC1 and EC3 have many more images than the other strains because we included images from samples treated at multiple antibiotic titrations.

5.2.2 Accuracy analysis

For the accuracy determination, any classifications of images that were part of the training/feedback dataset or classifications of "Image processing error" were removed. Summary statistics were performed in R (version 4.2.3) using the R package vegan [234] (version 2.6-4) and plotted with ggplot2 [235] (version 3.4.4). Accuracy was graded as whether the user's classification matched the image's predicted classification as defined by the MIC in *Accuracy analysis* (5.3.2). We report accuracy with the 95% confidence interval of the mean, by the Wald method with continuity correction.

5.2.3 Image feature analysis

CellProfiler [236, 237] (version 4.2.6) was used to extract image features from the dataset. The RGB .tif images were split into grayscale single-colour images using the ColorToGray module. Then, the IdentifyPrimaryObjects module was used with default settings to identify the Membrane object from the red channel. For the Nucleoid object, two-class Otsu thresholding was used with default settings because it segmented diffuse nucleoid regions more accurately. Intensity measurements for each object were measured using the MeasureObjectIntensity module. Size and shape measurements were extracted using the MeasureObjectSizeShape module. All measurement data were exported using CellProfiler to an SQLite database [238] and selected measurements were converted to .csv files with DB Browser for SQLite [239] (version 3.12.2).

Further image feature analysis was completed in Python and R scripts, available at `https://github.com/KapanidisLab/infection_inspection`. Images were excluded from analysis if more than half of their classifications were "Image Processing Error." An accuracy threshold of 0.5 was chosen to compare images that were most frequently classified correctly or incorrectly. For example, if a cell's predicted classification was Sensitive, based on its MIC and the treatment concentration, and it was classified as Sensitive by more than 50% of users, it would be labelled as Correct Sensitive. A cell from the same strain and treatment condition that was

classified as Resistant by more than 50% of users would be labelled as Incorrect Sensitive. This yielded four sets of images whose features could be compared: Correct Sensitive, Correct Resistant, Incorrect Sensitive, and Incorrect Resistant. Images were called Most Correct if they were classified correctly with a ratio greater than 0.94, corresponding to roughly 19 correct classifications of 20.

For feature comparisons between groups of cells displaying ciprofloxacin-resistant or sensitive phenotypes, we performed two-sided t-tests with Bonferroni corrections for multiple comparisons using the ggpubr [240] (version 0.6.0) and Rstatix [241] (version 0.7.2) packages. We compared the distributions of the values associated with cellular phenotypic features to a normally distributed Random Noise feature generated by numpy.random.normal [242]. A principal component analysis with 2 principal components was performed using the 7 measured image features and the Scikit-learn PCA function [243]. Before analysis, all feature measurements were normalised with Standard Scaler from Scikit-learn [243]. For the dataset with multiple ciprofloxacin concentrations, the principal component analysis with 2 principal components was performed in R with the prcomp [244] function from the stats library (version 4.1.3) and plotted with ggplot2 [235] (version 3.4.3).

Independently, we extracted the feature importance values from a Random Forest classifier with 100 trees and a minimum of 3 samples per leaf that had been trained using Scikit-learn [243] on images that were randomly allocated to a 75-25 train-test split and then scaled with Standard Scaler. The Random Forest classifier was evaluated by cross-validation by the Mean Absolute Error and was then applied to the test dataset to make predictions.

For each image, we calculated SHAP (SHapley Additive exPlanation) values for each feature using Kernel SHAP, a model-agnostic implementation for Python [245]. This method, which derives from game theory approaches, measures an importance value for each feature for each image classification, and has been shown to correspond well to intuitive human feature impact estimates.

5.3 Development of the Infection Inspection project

Infection Inspection was designed as a citizen science project on the Zooniverse platform [223] using the Project Builder [246], a free-to-use web browser application enabling research teams to build and contribute projects to the site (https://www.zooniverse.org/projects/conor-feehily/infection-inspection). During the building process, we developed an initial workflow, tutorial, and project field guide. Datasets were added to the project as .png images using the Subject Set upload tool within the Zooniverse Project Builder.

We wanted volunteers to classify our images of *E. coli* as "Sensitive" or "Resistant" (Figure 5.1). We encouraged users to make their best classification, even if they were unsure. We hypothesised that by forcing this binary choice and aggregating data from a large number of users, cells with a clear phenotype would be given the same label by most users, and cells with less clear phenotypes would have more mixed classifications. Infection Inspection was submitted for internal review in August 2022 and went to beta reviewers in September 2022. Based on their feedback, a label of "Image Processing Error" was added to account for a small number of images where a problem in the imaging pipeline made the image impossible to classify, such as a membrane with no DNA, DNA and membrane channel mismatch, or a segmentation error. At this stage, we also improved our project terminology and added explanations to the field guide and instructions for how to classify ambiguous or unusual cells.

During the beta test, we noticed that user accuracy did not improve with the number of classifications done. To help users learn from their own misclassifications, we implemented user feedback for a set of 30 tutorial images. These images had a ground truth classification of "Sensitive," "Resistant," or "Image Processing Error" and users would receive feedback on their accuracy immediately after submitting a classification for one of these images. Tutorial images were shown to users with decreasing probability: 0.5 for the first 5 to 10 images, declining to 0.25 by 20 images, and 0.05 after 50 images. The retirement limit was set to 20, meaning that each image was considered complete once it was classified by 20 unique volunteers.

The Zooniverse platform allows anyone to volunteer for projects and does not require any specialised training or qualifications beyond the project's tutorial and Field Guide. The tutorial and field guides were written in line with guidance on communicating with the public on antibiotic resistance from the Wellcome Trust [33]. We solicited and implemented feedback from non-experts, public engagement experts, and a secondary school biology teacher on the language used in the project before submitting for beta testing. While the terms antibiotic-susceptible and antibiotic susceptibility test are more standard in the clinical microbiology field, the labels sensitive and resistant were recommended by the public engagement expert for better reading accessibility.

When Infection Inspection launched on the 7th of February 2023, it was promoted through the Zooniverse email list, social media, and in-person outreach events. Before the launch of the final dataset, it was covered in the Citizen Science column in *New Scientist* magazine [247]. The initial dataset comprised 30 training and 5,000 dataset images and was completed within just 18 hours. Subsequently, a second dataset of 6,074 images was uploaded and was completed in 72 hours. The final dataset of 38,000 images, including images treated at different antibiotic concentrations, was uploaded later and was completed in 35 days (840 hours).

5.3.1 User participation

The Infection Inspection project was live from the 7th of February 2023 17:40 UTC to the 10th of May 2023 21:40 UTC, when all images in the dataset were completed. Only image classifications from that period were included in the analysis. All classifications from beta testing or those done on completed subjects were excluded. Before user accuracy and contribution analysis, all usernames were anonymised to User_1, User_2, etc. The Zooniverse system provides usernames and IP addresses for data analysis, but there is no way to differentiate multiple individuals using the same account or to aggregate an anonymous user contributing on several devices.

Over the course of the project, a total of 5,273 unique users performed at least one image classification and a total of 1,045,199 classifications were made. After

removing classifications of the training dataset, a total of 4,927 users and a total of 1,003,588 classifications remained. There was a large variation in the number of classifications per user, with the median at 38 classifications and with 56 users contributing more than 2,000 classifications (Figure 5.2). The user with the greatest number of image classifications contributed 46,289 labels. It is possible that this user, and many of our top contributors, represent several individuals using the same account. Among the 20 most engaged users, return to the project was common, with these users returning to the project on several occasions throughout (Figure 5.3).

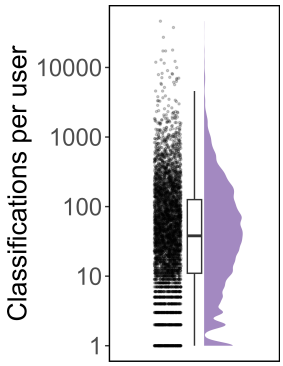


Figure 5.2: Distribution of the number of classifications per user (N=5,273 users, N=1,045,199 classifications). Each dot represents an individual user who performed a classification on at least 1 non-training dataset image. The box represents the middle 50% (IQR) of the users and the mid-line indicates the median number of classifications (median=38). The half violin plot shows the density for each number of classifications. Figure made by Conor Feehily.

We actively participated in the project discussion board throughout the project. Often, volunteers would start discussion threads about images that were difficult to classify. At the start of the project, we quickly responded to all of these discussions to help create an educational resource and to clarify how volunteers should think about unusual or confusing phenotypes. Later, experienced volunteers often stepped into this role. Most discussions had to do with images of unusual cells or questions about cell division and the *E. coli* life cycle. There were more than 450 discussion threads over the course of the project.

We used the Gini coefficient to characterise the extent of inequality in the distribution of classifications by volunteers. The Gini coefficient derives from a metric

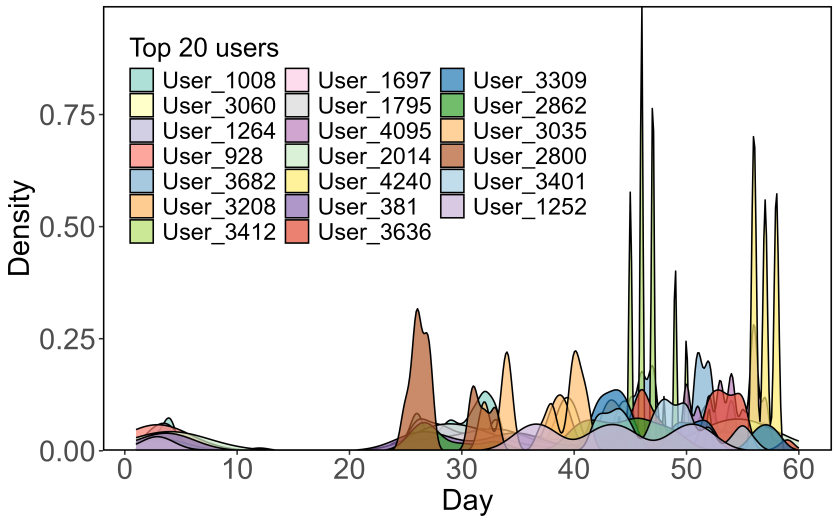


Figure 5.3: User engagement with Infection Inspection. Density mapping of contributions for each of the top 20 users (by number of classifications) over the course of the project, showing the different patterns of activity that led to significant numbers of classifications. Day 1 on the x-axis represents the first day that the user engaged with the project. Figure made by Conor Feehily.

for income inequality [248] and has been applied to measure inequality in volunteer contributions [228]. We calculated the Gini coefficient with the following formula:

$$1 - (2 \times \frac{\sum proportion\ of\ total\ classifications}{total\ users})$$

The Gini coefficient for our user participation was 0.81, which means that the most prolific volunteers contributed a large proportion of our project's classifications, or our project attracted many casual users, or both. The Gini coefficient for Infection Inspection is close to the mean Gini coefficients of the most popular ecology (0.80), astronomy (0.82), and transcription (0.81) projects on Zooniverse and higher than the average biomedical project score (0.67), based on a previous analysis [228].

5.3.2 Accuracy analysis

Image identifiers were matched back to the original strain and metadata including known MIC, treatment concentration, and clinical antibiotic susceptibility phenotype. Then, a predicted classification was assigned to each data point. The predicted classification was defined as the expected response of the strain to the antibiotic. If the MIC was less than the treatment concentration, the expected

phenotype was sensitive. If the MIC was greater than the treatment concentration, the expected phenotype was resistant. For instance, if the MIC for a particular strain was 0.03 mg/L and the treatment concentration was 10 mg/L, that strain was categorised as sensitive. Conversely, if the MIC was 72 mg/L and the treatment concentration was 10 mg/L, the strain was labelled as resistant.

5.4 Results of feature analysis of sensitive and resistant phenotypes

5.4.1 Differences in DNA morphology drive confusion

We used the proportion of misclassifications, or accuracy, as a measure for how confusing an image was to users. Accuracy was determined by comparing a user's classification to the predicted classification as defined by the MIC and treatment concentration. Because each image was classified by 20 users, we could look at the aggregated accuracy and inter-rater agreement to see how closely a cell matches the expected phenotype. Looking at the overall accuracy also minimised the impact of accidental or intentional misclassifications by individual users.

Overall, Infection Inspection volunteers classified ciprofloxacin phenotypes with 66.4±0.2% accuracy on sensitive cells and 67.3±0.2% on resistant cells. This is comparable to the accuracy of the published deep learning model [93] (Figure 5.4). The model was slightly less accurate in classifying resistant cells (62.5±0.9%), but much more accurate in classifying sensitive cells (88.2±0.4%).

Some images were more frequently misclassified than others. In the first uploaded dataset of ciprofloxacin-sensitive and ciprofloxacin-resistant clinical isolates treated at 10 mg/L, the classification accuracy histograms are skewed, with many images almost always classified correctly, and others almost never (Figure 5.5). This suggests that while many cells display the expected cellular phenotype when exposed to ciprofloxacin, there were sub-populations with atypical features.

To look at the common image features that were associated with misclassifications, we separated images by their aggregated classification accuracy. Images were assigned to the "Incorrect" class if they were classified with less than 50%

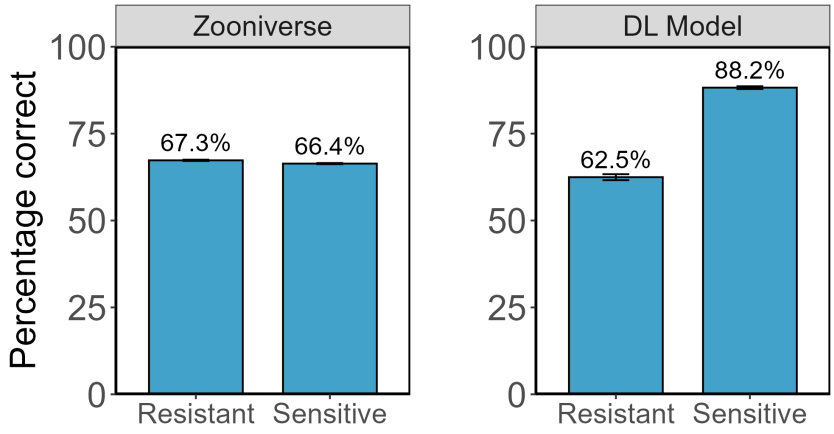


Figure 5.4: Image classification accuracy. Zooniverse user classification accuracy (Zooniverse) compared to the deep learning model (DL Model), stratified by the expected phenotype (Resistant or Sensitive). The percentage correct is defined as the classifications that matched the expected phenotype for a given sample. Error bars indicate the 95% confidence interval. Adapted from a figure made by Conor Feehily.

accuracy (1,755 images), and otherwise to the "Correct" class (7,074 images). Cells were excluded from the image feature analysis if they were labelled as an "Image Processing Error" by more than half of the users who classified the image. This removed 230 images.

To describe the commonalities and differences between these groups, human-interpretable image features were measured with CellProfiler [236, 237]. Seven features were chosen for their potential biological relevance to the ciprofloxacin response in *E. coli* (illustrated in Figure 5.6). To characterise the compaction, heterogeneity, and quantity of DNA, we measured the number of DNA regions per cell, the mean and standard deviation of the integrated intensity of the DNA regions, the mean standard deviation of the DNA intensity, and the area fraction occupied by the nucleoid regions. For these automated measurements, the term DNA region is not synonymous with the biological nucleoid, rather it refers to a spot of higher DNA density that comprises part or all of the nucleoid. These measurements also capture the variations in nucleoid morphology that can be seen within the same cell and may reflect differences in cell cycle stage. The cell shape was described by the form factor of the membrane and the major axis length was

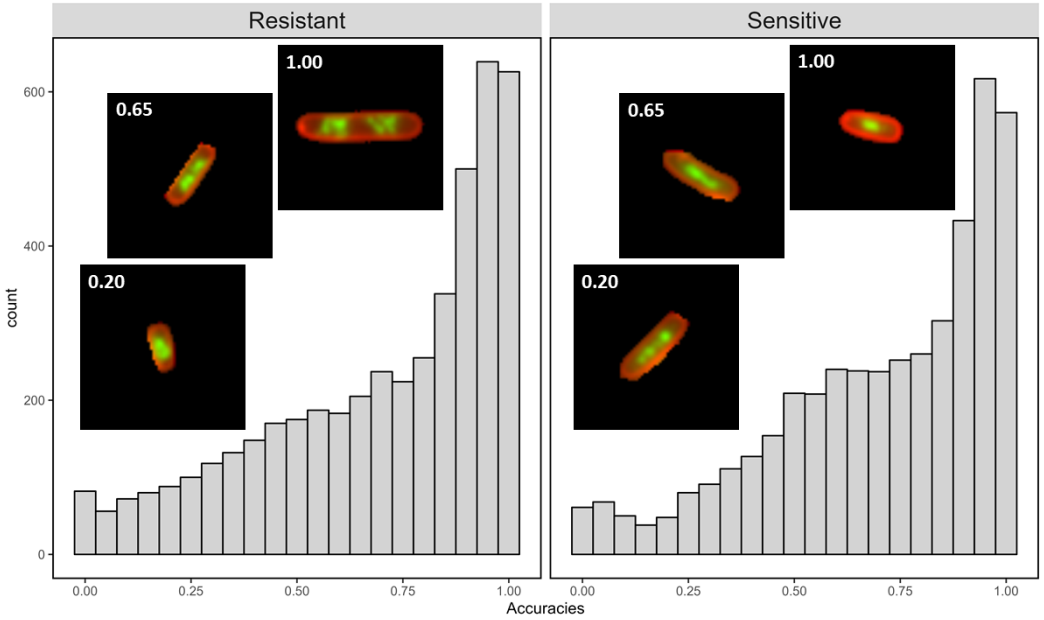


Figure 5.5: User accuracy varied by image for both resistant and sensitive cells. Histograms of the user accuracy on images of *E. coli* treated at 10 mg/L for resistant and sensitive cells. Representative images of resistant and sensitive cells with low classification accuracy (0.20), intermediate classification accuracy (0.65) and high accuracy ($\geq$0.95) are shown. Both resistant and sensitive cells show a skew, with many cells being classified correctly nearly always and some almost never. However, both populations also have many ambiguous images that were classified correctly by around half of the users.

used to measure the cell size. The Form Factor of a perfectly round object is equal to 1, so most bacilli will have form factors <1.

The image features of sensitive and resistant cells that were most often classified correctly were all significantly different (Corrected *t*-test p<0.001), indicating that differences in morphology were associated with each class. The image features of Correct Sensitive and Incorrect Resistant (truly sensitive) and of Correct Resistant and Incorrect Sensitive (truly resistant) cells were also significantly different from each other (p<0.001). Interestingly, however, when comparing the features of cells that were most frequently classified incorrectly, there was no significant difference in the mean integrated intensity of the DNA regions (p=1), the mean standard deviation of intensity of the DNA regions (p=0.34), and the nucleoid area fraction (p=1) between sensitive and resistant bacteria (Figure 5.6). This suggests that users were unable to classify cells accurately when these DNA-associated features were too ambiguous or too similar to what is expected for the other class.

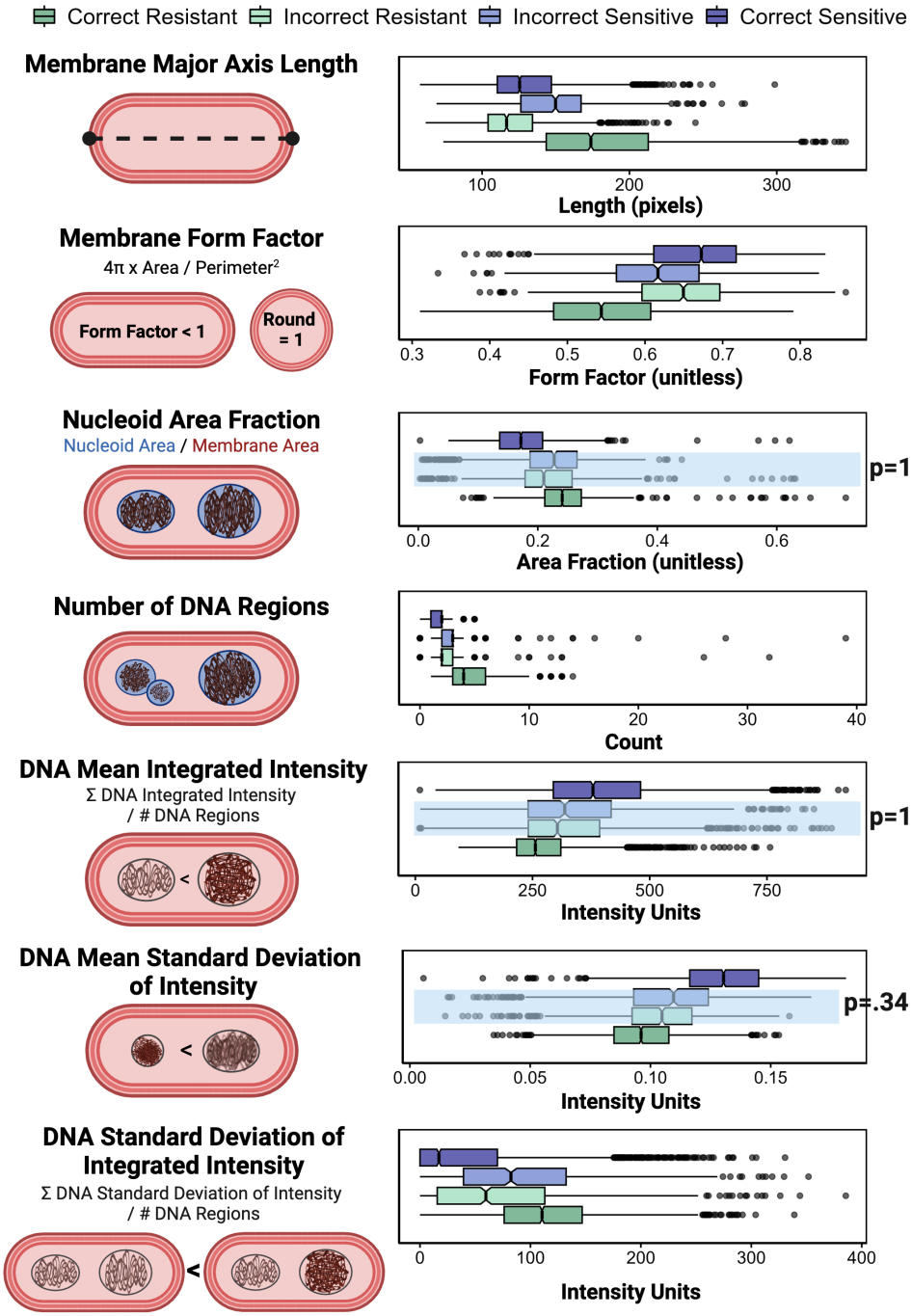


Figure 5.6: Feature measurements associated with correct and incorrect classifications. Seven feature measurements were chosen for their potential to reflect responses to ciprofloxacin treatment. The features, with illustrative diagrams, are shown on the left; box plots of the feature distributions for all cells are shown on the right, coloured by whether the cell was Sensitive or Resistant and whether they were most often classified Correctly or Incorrectly. For each feature, pairwise *t*-tests were performed for the Correct Resistant, Incorrect Resistant, Incorrect Sensitive, and Correct Resistant distributions. The p-value with Bonferroni correction for multiple comparisons is <0.0001 except where highlighted and labelled. Figure created with BioRender.com.

5.4.2 Incorrectly classified images have more ambiguous phenotypes

To further test the hypothesis that cells with atypical or ambiguous phenotypes are more likely to be classified incorrectly, a principal component analysis (PCA) was performed on the image features. The principal component analysis allowed us to visualise the phenotypic variance in our image dataset by projecting the feature measurements of each cell into a 2-dimensional space such that images with more similar features would cluster together. In addition, the loading vectors of each feature showed the magnitude of that feature's contribution to the ciprofloxacin-sensitive and ciprofloxacin-resistant phenotypes.

The variation in the first principal component was primarily driven by the number of DNA regions, the standard deviation of the integrated intensity of the DNA regions, and the cell major axis length; the second principal component was driven by the nucleoid area fraction and the mean integrated intensity of the nucleoid (Figure 5.7a). In descriptive terms, the cell length, nucleoid copy number, and phenotypic variance between nucleoids in the same cell were the largest contributions to the first principal component. The second principal component was more driven by the quantity of DNA (mean integrated intensity), and the cell volume it occupies (area fraction).

Correctly classified images of sensitive or resistant *E. coli* formed separate phenotypic clusters, with some overlap (Figure 5.7a). Incorrectly classified images clustered in the centre of the phenotype space, with greater variation in the second principal component (Figure 5.7b). Therefore, images that were frequently classified incorrectly had intermediate phenotypes, with DNA morphology that did not clearly demonstrate signs of ciprofloxacin resistance or sensitivity, or suggested features of the opposite class. The greater variation in the second principal component suggests that these cells also had different quantities of DNA than the typical cell.

To see the distribution of the prototypical sensitive or resistant *E. coli*, we defined a "Most Correct" class containing images that were classified with greater

than 94% accuracy. For these images, there was distinct clustering with less overlap than was shown for the other correct images (Figure 5.7c).

We were also interested in the phenotype of images that were classified differently by the deep learning model and the Zooniverse users. Those cells clustered almost exclusively in the centre of the plot, in between the distribution of Most Correct images (Figure 5.7d); they also had more variation in the second principal component. Because the model had higher accuracy on images of ciprofloxacin-sensitive cells, this indicated that the model was using more nuanced features from the second principal component to classify images that the users were more likely to misclassify.

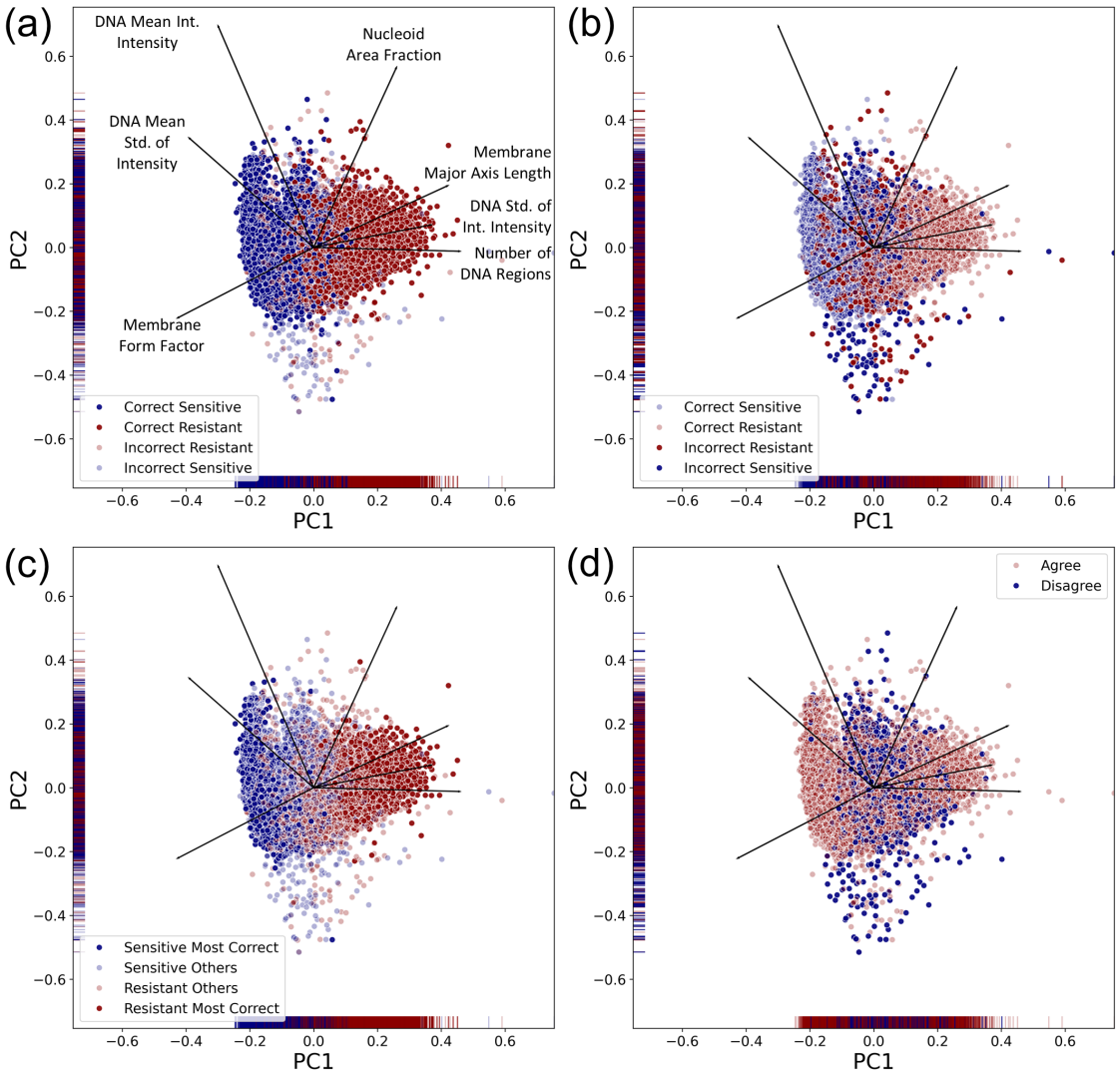


Figure 5.7: Phenotype analysis of correctly and incorrectly classified cells by PCA. (a) In a 2-dimensional Principal Component Analysis (PC1, PC2), images of Resistant and Sensitive cells that were classified correctly more than 50% of the time (Correct Sensitive, blue dots and Correct Resistant, red dots) cluster with other members of their class, with some overlap. The feature loadings are shown as black lines. The most influential features for the first Principal Component (PC1) are the Number of DNA Regions and the DNA Standard Deviation (Std.) of Integrated (Int.) Intensity, a measure of the variation in nucleoid region brightness within the cell. (b) Images that were classified incorrectly more than 50% of the time cluster near the centre of the principal component plot, or near images of the opposite class, with greater variance in the second principal component than correctly classified images. (c) Images that were classified with more than 94% accuracy were called "Most Correct." The Sensitive Most Correct cells and Resistant Most Correct cells form distinct clusters, indicating that there are certain populations of cells that exhibit characteristic features and are therefore likely to be classified accurately. (d) Images where the model and the consensus of users disagreed are highlighted in blue. The users were more likely to disagree with the model on images with intermediate features, in the centre, or on images with more atypical features, with greater variance in PC2.

5.4.3 Feature importance in ciprofloxacin-sensitive and resistant cells

The PCA clusters provided clues to which features primarily differentiate the ciprofloxacin-sensitive and ciprofloxacin-resistant phenotypes. To investigate which features might be the most influential in ciprofloxacin phenotype classification, we computed SHAP contributions (SHapley Additive exPlanations) [245] for each measured image feature. Although this is a machine learning approach, SHAP values have shown good agreement with human interpretations [245].

For the SHAP analysis, a Random Forest classifier was trained to classify images from our dataset as sensitive or resistant using our image feature measurements. An additional feature of normally distributed random numbers was added to determine which features held significance greater than random noise. The Random Forest classifier achieved a Mean Absolute Error of 0.15±0.01 on a holdout dataset. The trained Random Forest model was then used to compute SHAP feature contribution scores for each image in the test holdout dataset. The average importance of a feature can be measured by the mean absolute value of the SHAP contribution for all images in the dataset.

Using this approach, the most important features in both sensitive and resistant images were the DNA mean standard deviation of intensity (median SHAP=0.109, *t*-test with Bonferroni correction for multiple comparisons $p<0.0001$), the number of DNA regions (median SHAP=0.085, $p<0.0001$), and the nucleoid area fraction (median SHAP=0.081, $p<0.0001$) (Figure 5.8). All of the measured features contributed more to the classification task than the normally distributed random noise ($p<0.0001$).

5.4.4 Phenotypic differences in resistant *E. coli* treated with high ciprofloxacin concentrations

In addition to the stark phenotypic differences between ciprofloxacin-treated sensitive and resistant bacteria treated at the same antibiotic concentration, we used the dataset for an *E. coli* strain with intermediate resistance (EC3, MIC=0.5 mg/L) to

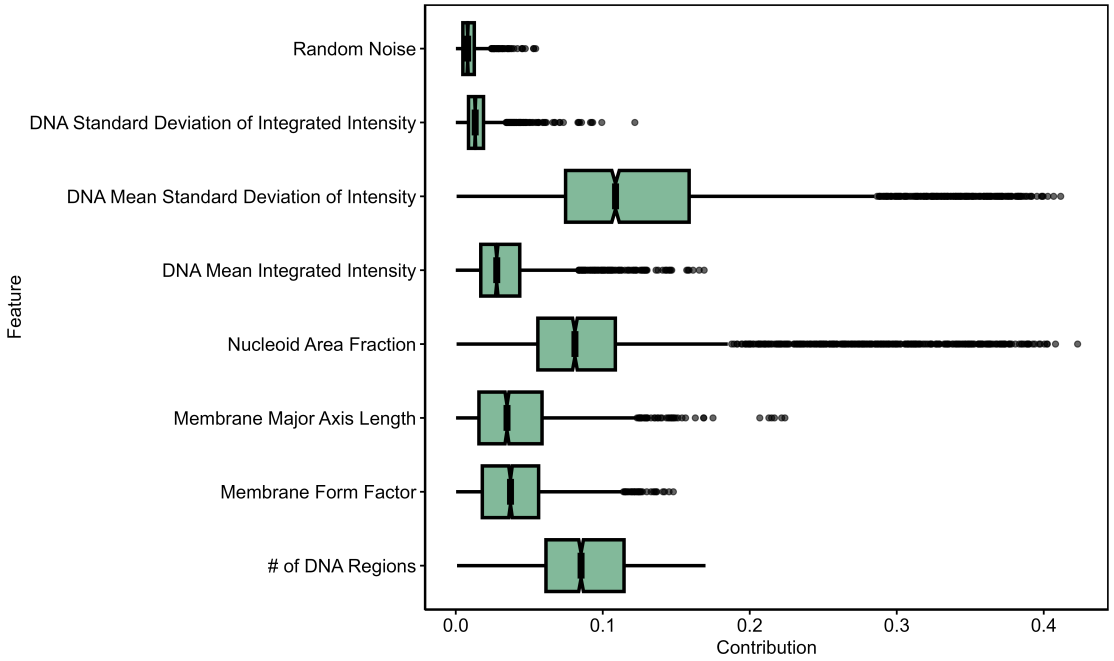


Figure 5.8: Feature significance measured by SHAP values for all images. The absolute value of the SHAP contribution for each feature is shown on a box plot. In order of their median SHAP contribution, the features are DNA Mean Standard Deviation of Intensity (0.109), Number of DNA Regions (0.085), Nucleoid Area Fraction (0.081), Membrane Form Factor (0.037), Membrane Major Axis Length (0.035), DNA Mean Integrated Intensity (0.028), DNA Standard Deviation of Integrated Intensity (0.013), and Random Noise (0.008). All the features are significantly different from Random Noise ($p<0.0001$), and all features are significantly different from each other ($p<0.0001$), except Membrane Major Axis Length and Membrane Form Factor (no significance).

investigate phenotypic differences when a strain is treated with different ciprofloxacin concentrations above its MIC for 30 minutes. A principal component analysis showed that there was a phenotypic difference between cells treated at or below 2- and 4-times the MIC (1 and 2 mg/L, respectively) and cells treated at 8-, 16-, and 32-times the MIC (4, 8, and 16 mg/L, respectively) (Figure 5.9). At these high treatment concentrations, there was a shift in the centroid of the cluster towards different features.

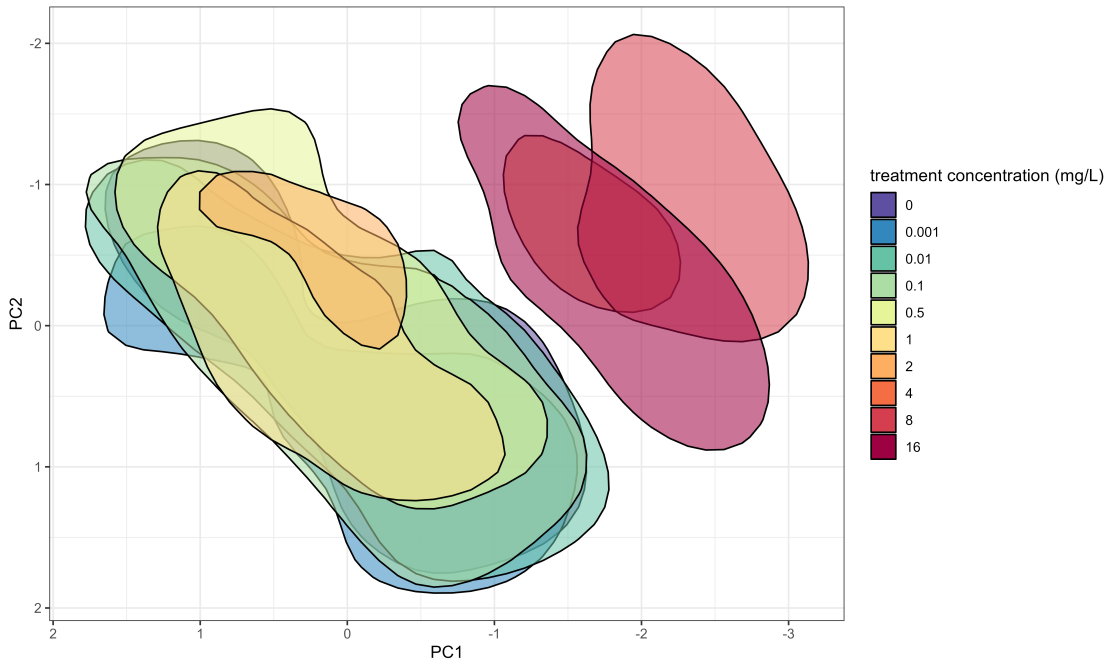


Figure 5.9: Resistant *E. coli* treated with high concentrations of ciprofloxacin cluster together. This ciprofloxacin-resistant clinical isolate (EC3; MIC = 0.5 mg/L) was treated at varying concentrations of ciprofloxacin for 30 minutes. This principal component analysis shows that features of cells treated at extremely high concentrations of ciprofloxacin (8-, 16-, and 32-times MIC; 4, 8, and 16 mg/L) cluster with a different centroid than those treated at lower concentrations, even when those concentrations are above the MIC. Geometric shapes are plotted with different colours to show the regions with point density above 0.07.

5.5 Discussion

5.5.1 Feature analysis

The Infection Inspection showed that misclassifications of 5 tested ciprofloxacin-sensitive and ciprofloxacin-resistant *E. coli* strains are associated with diversity in the appearance of the bacterial DNA after treatment with ciprofloxacin. On this dataset of 49,074 images of cells from clinical isolates, the machine learning-based classifier achieved a classification accuracy as high as 80%, but it was unclear what image features led to misclassifications. The model is much more accurate when classifying ciprofloxacin-sensitive cells, but Infection Inspection users had similar accuracies on sensitive and resistant cells (Figure 5.4). By image feature analysis, we found that images that were most likely to be classified incorrectly

by users did not show the phenotypic features of correctly classified ciprofloxacin-sensitive or resistant cells, indicating that these bacteria develop ambiguous or intermediate phenotypes. Using Principal Component Analysis, we found that accuracy differences between the machine classifier and the users were greatest for images with more atypical features, suggesting that the model out-performs human classifiers in recognising these signs.

We combined feature analysis with SHAP feature contribution scores to determine that DNA mean standard deviation of intensity, the nucleoid area fraction, and the number of DNA regions were the most important features when deciding how to classify an image. This means that the degree of DNA compaction and heterogeneity, and the space it occupies within the cell, are the key features that can be used to determine whether an *E. coli* bacterium is responding to ciprofloxacin treatment. Because we chose rational, pre-determined features, it is likely that there are other features that could be added to better characterise the phenotypes seen. For example, the sensitive cells have radially compacted DNA regions, which could be characterised by measuring the DNA to membrane distance.

Our use of automated image feature quantification enabled us to compare the features of all of the cells in our dataset; however, a limitation of this method is that the detected feature measurements sometimes diverge from human or biological interpretation. For example, CellProfiler measured more than 10 regions of DNA density in some images, which is biologically implausible (Figure 5.6). While these outliers may not match human annotations, the law of large numbers makes our average measurements meaningful.

5.5.2 Volunteer participation

The successful participation of the public with Infection Inspection and the speed at which users classified the images highlight the interest in and the value of the public in tackling the problem of antimicrobial resistance and interpreting large biomedical datasets. Despite this, biomedical projects are a minority on the Zooniverse platform. A 2019 study showed only 3 biomedical projects of 63

projects surveyed on the platform (5%) [228]. As of November 13, 2023, this fraction remained low, with only 5 of 100 (5%) active projects in a biomedical discipline. Users also interact differently with biomedical projects on the Zooniverse. The Gini coefficient is a measure of inequality that has been used to assess the degree to which many casual volunteers and some super-users contribute to the shared work of Zooniverse projects. On average, biomedical projects were found to have a notably lower average Gini coefficient than astronomy projects, which could be the result of fewer return volunteers, or because biomedical projects more successfully attract many casual contributors [228]. Infection Inspection had a Gini coefficient of 0.81, higher than the average biomedical project studied. We speculate that our single-step, fast workflow encouraged some users to contribute more classifications than the average biomedical research project.

Comments from volunteers who participated in the Infection Inspection project demonstrate how citizen science methods can be mutually beneficial for research teams and for participants. Several volunteers mentioned that they were motivated to contribute to a project about antibiotic resistance because people close to them had been affected by bacterial infections. One volunteer stated that participating in citizen science projects "gives a sense of pride and usefulness." Another citizen scientist said that they love this kind of volunteering and that it was "interesting to learn...how bacteria can be acting under the influence of antibiotics." These comments and others on the project discussion board, as well as the enthusiasm with which the volunteers made more than 1 million classifications, show the interest in biomedical citizen science projects and how such projects can be both educational and meaningful for participants.

5.5.3 Limitations

While the results of the Infection Inspection project are informative, the project had limitations. It relied on voluntary contributions from citizen scientists, which introduced variability in data quality and quantity. We had no information on the users participating in the study and no quantitative feedback on the impact

of our tutorials on informing the public about AMR. This study focused on a single antibiotic, ciprofloxacin, and used cells obtained from a small number of bacterial strains, limiting the ability of these findings to be generalised to other antibiotics and pathogens.

We believe our efficient, single-step classification workflow increased user engagement with our project, but other workflow designs could capture more data on the volunteer's decision-making process. On the project discussion board, some volunteers started discussions about images that appeared to be cells in the process of cell division or images that looked unusual. While this project was not designed to classify images in such detail, it is encouraging to realise that users could be asked to consider stages of cell growth in a future task. Researchers working with citizen scientists can take advantage of this scientific intuition in understanding their datasets.

5.5.4 Future directions

Our image feature analysis revealed that sub-populations of sensitive and resistant cells develop DNA features that lead to misclassifications, likely because they resemble cells from the opposite class or because they exhibit atypical features. With more training data, the model may be able to learn to classify these atypical features. Otherwise, given a certain number of cells with intermediate or ambiguous features, the contribution of each single-cell classification to the sample classification could be weighted by the confidence of the CNN in its prediction. For a binary classification task, this can be extracted as the value of the activation function for the final output layer. With this approach, the classifications with higher confidence would contribute more to the final classification of the sample. Additionally, a threshold could be set to flag samples for which there are many cells with low classification confidence.

In addition to the feature measurements, Principal Component Analysis, and SHAP analysis shown here, misclassifications could also be investigated by looking at the regions of cells that have the greatest impact on classification decisions. In image classification, this is called saliency mapping. One method plots the gradient

from the output to the feature map of the final convolutional layer in order to identify important regions of the image [179]. Using this method on our image data, perhaps along with a method to plot the average cell [180], could help to show which regions of misclassified cells are the most misleading to the classifier.

Overall, the findings of the Infection Inspection project demonstrate the potential for a citizen science project design to be used to characterise other bacterial strains, antibiotics, and experimental treatment conditions, such as the dynamic responses of bacterial cells to antibiotics in a time course. The antibiotic response phenotypes of the most common bacterial pathogens and antibiotics will need to be characterised before a phenotypic AST based on intracellular morphology can be brought to the clinic.

Looking ahead, citizen scientists can continue to play a pivotal role in biomedical research and in addressing global health challenges related to antibiotic resistance. Future engagements could involve exploring other bacterial species and antibiotics, improving training materials and guidelines, raising public awareness, and integrating an assessment of the impact of these tools on user education about the scientific topics being studied.

5.6 Conclusions

In this chapter, we investigated the image features and phenotypes associated with ciprofloxacin sensitivity and resistance in *E. coli* clinical isolates with the help of an online citizen science project, Infection Inspection. We found that misclassifications are associated with atypical or ambiguous DNA features in ciprofloxacin-treated cells, and that the machine learning-based classifier outperforms human volunteers in classifying nuanced phenotypes.

6 Conclusion

This thesis was motivated by the potential for rapid antibiotic susceptibility tests to slow the development of antimicrobial resistance, to conserve existing antibiotics, and to improve clinical outcomes for patients with antibiotic-resistant infections. When rapid ASTs are available, patients receive optimal antibiotic treatment faster, spend less time in the hospital, and have lower healthcare costs and improved clinical outcomes [72, 74]. This motivation has led to the development of a variety of novel methods for antimicrobial susceptibility testing, including molecular and phenotypic methods, some of which are approved for clinical use or are in clinical trials [249, 250, 103, 87, 106, 104]. Still, in most clinical settings the time-to-result for gold-standard ASTs remains around 12 to 48 hours, leading to a lag in treatment optimisation and the prescription of empiric, broad-spectrum antibiotic therapies. Ultra-fast tests, which skip the primary culture step and use cells captured directly from the patient sample, like the Longitude prize-winning PA-100 urine AST [87] and the nanoparticle, micro-growth blood AST [90], are necessary to return results in less than 2 hours.

Compared to other rapid ASTs on the market, the single-cell, image-based phenotyping method is promising because it could be conducted directly on captured bacterial cells and could simultaneously identify bacterial species. Recent publications have demonstrated the capture of >90% of cells from dilute clinical samples with micro-fluidics [160] or with sβ2GPI nanoparticles [90]. Combined with

these methods, single-cell phenotyping could subsequently be used to determine the antibiotic susceptibility of the captured cells. This could cut down the time-to-result to less than 2 hours, which would fit into inpatient and outpatient clinical decision-making workflows. With such a rapid test, patients could be started on targeted therapy directly, decreasing the risk of antibiotic toxicity and decreasing inappropriate antibiotic use.

In this thesis, we contributed to the development of a rapid, single-cell phenotypic test by exploring the use of ribosome phenotypes for antibiotic susceptibility testing. The use of ribosome phenotypes for single-cell phenotyping represents a step forward in this area of research because it enables a single label to be used for cell segmentation, AST classification, and potentially species ID. We also explored failure modes of these assays, in particular the causes of misclassified cells. In Chapter 2, we detailed the methods we used to capture high-quality, single-cell ribosome and DNA phenotypes, and how we trained neural networks to classify these images. In Chapter 3, we showed that ribosome-targeting FISH probes can be used to visualise changes in intracellular morphology associated with antibiotic susceptibility for 4 antibiotics from clinically relevant classes with different mechanisms of action, namely ciprofloxacin, gentamicin, chloramphenicol, and carbenicillin. In Chapter 4, we developed a convolutional neural network that could classify ciprofloxacin-susceptible and resistant isolates with 95% overall single-cell accuracy based on their ribosome phenotypes. In Chapter 5, we showed that most misclassifications are associated with atypical or ambiguous DNA, rather than mistakes made by the classifier.

In the process of this work, we have also learned some generalisable principles related to using neural networks for the classification of single-cell phenotypes. We found that our model training improved when we removed shear and blur from our data augmentation strategy. These models outperformed models trained with shear and blur on holdout testing data. We also found that for our dataset, cropping or resizing long cells resulted in similar model performance. Our models also demonstrated some ability to transfer to previously unseen strains or antibiotic

treatment concentrations, although this performance was variable and depended on the similarity of the testing data to the training data.

Because we used a CNN instead of using measured cell features, it would be interesting to explore the ability of the model to classify fields of view from cropped images showing populations of cells. This could potentially improve the accuracy problems caused by individual cells with unusual phenotypes, although it would be more susceptible to bias from the image background.

Through the Infection Inspection citizen science project, we engaged more than 5,000 volunteers in the development of our single-cell phenotyping methods. From their classifications and in discussions with them on the message board, we learned what features make some cells difficult to classify. Our results suggest that intra-population variance in antibiotic response is the biggest driver of misclassifications, rather than deep learning classifier mistakes. Therefore, future work should aim to capture this biological variability for model training, to maximise real-world accuracy. Developing and participating in the Infection Inspection project was a meaningful process for the research team and for many of the participants. The engagement we received should encourage future biomedical research teams to launch citizen science projects, if they have the time and resources to invest in a sustained project. It took around 6 months from our initial project submission to the project launch, and it is important to stay active in communicating with volunteers over the course of the project and afterwards. We hope volunteers learned from their participation in the project, and feel more optimistic and empowered about the challenge of antimicrobial resistance.

This investigation into single-cell phenotyping was designed to demonstrate and explore a novel method, and it was limited in its choice of antibiotics, isolates, and species. Before this method can have clinical utility, it will need to be extended to other bacterial species and antibiotics, but the basic principles should remain. Species-specific FISH probes targeting the ribosomal RNA have been used to identify a variety of Gram positive and negative species for clinical applications [161, 153, 251], and we expect that these probes could be used for ribosomal

phenotyping. It is likely that the extension of this method from *E. coli* to other Gram-negative bacilli will be more straightforward, while the smaller size of cocci and lower permeability of Gram-positive bacteria to FISH probes may be more challenging [252, 253]. Similarly, we have shown that three antibiotics with intracellular targets (ciprofloxacin, gentamicin, and chloramphenicol) and one membrane-targeting antibiotic (carbenicillin) cause characteristic ribosome phenotypes. We expect that this method can be extended to other antibiotics, as long as they reliably induce a visible change in the ribosome phenotype within the time scale of the test. As was the case for gentamicin and carbenicillin, the benchmark treatment concentration may need to be adjusted for the mechanism of action of the antibiotic. The length of the antibiotic treatment may also need to be adjusted, especially for slower-growing species.

In future work, it will be important to characterise the ribosome phenotypes for cells treated with antibiotics in different growth conditions. Growth rate and growth medium are known to have a profound effect on *E. coli* nucleoid structure [184], with slower growing *E. coli* having a more diffuse DNA region and less ribosome-nucleoid segregation. It would be beneficial to collect data from cells grown in nutrient-poor media, or to simulate the conditions found in clinical specimens such as blood or urine. These data will show whether cells captured from a clinical sample need to be pre-cultured in rich media before they will show clear ribosome phenotypes, or whether it will be possible to begin antibiotic treatment immediately after cell capture.

Here, we also focused on a 30-minute treatment time for all antibiotics. It would be interesting to investigate whether faster treatment times can be used for at least some antibiotics. In some studies, significant changes in DNA morphology can be seen within minutes of chloramphenicol treatment [146, 190]. These experiments could be conducted in live cells to see how early a classifiable difference in the ribosome phenotype can be detected. Future work could also explore whether antibiotic-susceptible and resistant isolates show phenotypic differences before antibiotic treatment, which could potentially be used to classify cells without an antibiotic treatment phase.

Phenotypic assays have the benefit of directly measuring the ability of the antibiotic to inhibit bacterial growth or cause cell death. Compared to growth or viability-based methods, single-cell phenotyping by fluorescence microscopy has the advantage of offering insight into the mechanism of action of the antibiotic and potentially into the mechanism of resistance. This could be especially useful in the development of new antimicrobial compounds or when used in conjunction with genotype information in research. However, the cells must be fixed and permeabilised in order to conduct a FISH assay, which limits the time-dependent data that can be collected. A single time point for the interrogation must be pre-selected, which, if incorrectly timed, could miss important data for susceptibility classification. Phenotypic methods such as single-cell ribosome phenotyping are more complex than many other methods, such as label-free growth assays. A significant engineering challenge of the ribosome phenotyping method is that it requires separate interrogations for each antibiotic. This could be parallelised, especially on a microfluidics platform, but it introduces complexity to the testing procedure. Using an image-based phenotyping method also requires a microscopy platform and a method for sample preparation. Adapting the ribosome phenotyping method presented here to clinical use will require significant engineering development in order to develop a platform that can capture the cells from the specimen, expose them to antibiotic treatment, and conduct the FISH assay and imaging. We envision that this technique could be most easily implemented in a well-resourced clinical laboratory or research setting, although it could be adapted to a lower-cost imaging platform than was demonstrated here, such as a 60× optical microscope.

Despite these challenges, the biggest advantage of using a CNN alongside rich imaging data for the detection of AMR is that the benefits feed forward. This thesis has already resulted in >50,000 high-resolution, curated single-cell images that are being made available online for further research. Many mechanisms contribute to antibiotic resistance beyond mutations in target genes, such as the over-expression of efflux pumps and plasmid-borne *Qnr* genes [86, 209, 254]. Large datasets like these, showing morphological differences in antibiotic response, could be powerful

in developing our understanding of bacterial antibiotic response, especially when combined with whole-genome sequencing. This diversity in antibiotic response is one of the greatest motivators for a CNN-based phenotypic AST, because when it is able to continuously learn from additional, real-world data, its performance would only be expected to improve [134]. We could expect the model to perform well, even on previously unseen strains and phenotypes, so long as they are sufficiently similar to something it had seen before. With these tools, the extraordinary diversity and constant evolution of pathogenic bacteria is not an insurmountable challenge, but an opportunity to learn.

References

[1] Ramanan Laxminarayan et al. “Antibiotic Resistance—the Need for Global Solutions”. In: *The Lancet infectious diseases* 13.12 (2013), pp. 1057–1098.

[2] David S. Jones, Scott H. Podolsky, and Jeremy A. Greene. “The Burden of Disease and the Changing Task of Medicine”. In: *New England Journal of Medicine* 366.25 (June 2012), pp. 2333–2338.

[3] Michael Padget. “Antimicrobial Resistance: A Frightening and Complex Public Health Challenge”. In: *Stemming the Superbug Tide*. OECD, 2019.

[4] World Health Organization. *Global Excess Deaths Associated with COVID-19 (Modelled Estimates)*. May 2023.

[5] Abigail Bartlett et al. “A Comprehensive List of Bacterial Pathogens Infecting Humans”. In: *Microbiology* 168.12 (2022), p. 001269.

[6] Stilianos Louca et al. “A Census-Based Estimate of Earth’s Bacterial and Archaeal Diversity”. In: *PLOS Biology* 17.2 (Feb. 2019), e3000106.

[7] Kevin S Ikuta et al. “Global Mortality Associated with 33 Bacterial Pathogens in 2019: A Systematic Analysis for the Global Burden of Disease Study 2019”. In: *The Lancet* (Nov. 2022).

[8] Daniel J. Diekema et al. “The Microbiology of Bloodstream Infection: 20-Year Trends from the SENTRY Antimicrobial Surveillance Program”. In: *Antimicrobial Agents and Chemotherapy* 63.7 (June 2019), 10.1128/aac.00355–19.

[9] Stuart B. Levy and Bonnie Marshall. “Antibacterial Resistance Worldwide: Causes, Challenges and Responses”. In: *Nature Medicine* 10.12 (Dec. 2004), S122–S129.

[10] Jules Brunel. “Antibiosis from Pasteur to Fleming”. In: *Journal of the History of Medicine and Allied Sciences* 6.3 (1951), pp. 287–301.

[11] J. F. Linares et al. “Antibiotics as Intermicrobial Signaling Agents Instead of Weapons”. In: *Proceedings of the National Academy of Sciences of the United States of America* 103.51 (Dec. 2006), pp. 19484–19489.

[12] Khaled Elsayad. “What Ancient Egyptian Medicine Can Teach Us”. In: *JCO Global Oncology* 9 (June 2023), e2300146.

[13] Matthew I Hutchings, Andrew W Truman, and Barrie Wilkinson. “Antibiotics: Past, Present and Future”. In: *Current Opinion in Microbiology*. Antimicrobials 51 (Oct. 2019), pp. 72–80.

[14] Mark S. Butler et al. “A Review of Antibacterial Candidates with New Modes of Action”. In: *ACS Infectious Diseases* (July 2024).

[15] Garima Kapoor, Saurabh Saigal, and Ashok Elongavan. "Action and Resistance Mechanisms of Antibiotics: A Guide for Clinicians". In: *Journal of Anaesthesiology, Clinical Pharmacology* 33.3 (2017), pp. 300–305.

[16] Michael A. Kohanski, Daniel J. Dwyer, and James J. Collins. "How Antibiotics Kill Bacteria: From Targets to Networks". In: *Nature Reviews Microbiology* 8.6 (June 2010), pp. 423–435.

[17] Robert W. Finberg et al. "The Importance of Bactericidal Drugs: Future Directions in Infectious Disease". In: *Clinical Infectious Diseases* 39.9 (Nov. 2004), pp. 1314–1320.

[18] Eili Y. Klein et al. "Global Increase and Geographic Convergence in Antibiotic Consumption between 2000 and 2015". In: *Proceedings of the National Academy of Sciences of the United States of America* 115.15 (Apr. 2018), E3463.

[19] Public Health England. *Total Consumption of Antibiotics in England in 2023, by Antibiotic Group (in DDD per 1,000 Population).*

[20] Christopher J. L. Murray et al. "Global Burden of Bacterial Antimicrobial Resistance in 2019: A Systematic Analysis". In: *The Lancet* 399.10325 (Feb. 2022), pp. 629–655.

[21] The Review on Antimicrobial Resistance. *Tackling Drug-Resistant Infections Globally: Final Report and Recommendations.* Tech. rep. 2016.

[22] World Bank. *Drug-Resistant Infections: A Threat to Our Economic Future.* Tech. rep. Washington, DC, 2017.

[23] Herman Goossens et al. "Outpatient Antibiotic Use in Europe and Association with Resistance: A Cross-National Database Study". In: *The Lancet* 365.9459 (Feb. 2005), pp. 579–587.

[24] Andrew Cole. "GPs Feel Pressurised to Prescribe Unnecessary Antibiotics, Survey Finds". In: *BMJ (Clinical research ed.)* 349 (Aug. 2014), g5238.

[25] António Teixeira Rodrigues et al. "Understanding Physician Antibiotic Prescribing Behaviour: A Systematic Review of Qualitative Studies". In: *International Journal of Antimicrobial Agents* 41.3 (Mar. 2013), pp. 203–212.

[26] Government of the United Kingdom, Global and Public Health Group, Emergency Preparedness and Health Protection Policy Directorate. *Contained and Controlled: The UK's 20-Year Vision for Antimicrobial Resistance.* Tech. rep. 2019.

[27] Daniel J. Morgan et al. "Non-Prescription Antimicrobial Use Worldwide: A Systematic Review". In: *The Lancet. Infectious Diseases* 11.9 (Sept. 2011), pp. 692–701.

[28] Peter Katona and Judit Katona-Apte. "The Interaction between Nutrition and Infection". In: *Clinical Infectious Diseases: An Official Publication of the Infectious Diseases Society of America* 46.10 (May 2008), pp. 1582–1588.

[29] World Health Organization. *Global Action Plan on Antimicrobial Resistance.* Tech. rep. Jan. 2016, p. 45.

[30] Joseph A. Lewnard et al. "Burden of Bacterial Antimicrobial Resistance in Low-Income and Middle-Income Countries Avertible by Existing Interventions: An Evidence Review and Modelling Analysis". In: *The Lancet* 403.10442 (June 2024), pp. 2439–2454.

[31] Global Leaders Group on Antimicrobial Resistance. *Annex to the GLG Report: Towards Specific Commitments and Action in the Response to Antimicrobial Resistance*. Position Paper. Apr. 2024.

[32] Jimmy Nkaiwuatei. "Turning the Tide against Antimicrobial Resistance in Africa". In: *Nature Microbiology* 8.11 (Nov. 2023), pp. 1937–1938.

[33] Wellcome Trust. *Reframing Resistance*. Tech. rep. Oct. 2019.

[34] Rainer Tan et al. "A Digital Health Algorithm to Guide Antibiotic Prescription in Pediatric Outpatient Care: A Cluster Randomized Controlled Trial". In: *Nature Medicine* 30.1 (Jan. 2024), pp. 76–84.

[35] Remco P. H. Peters et al. "Novel Lateral Flow Assay for Point-of-Care Detection of Neisseria Gonorrhoeae Infection in Syndromic Management Settings: A Cross-Sectional Performance Evaluation". In: *Lancet (London, England)* 403.10427 (Feb. 2024), pp. 657–664.

[36] Fernandez-Pittol M et al. "Multicenter Study to Assess the Use of BL-DetecTool for the Detection of CTX-M-type ESBLs and Carbapenemases Directly from Clinical Specimens". In: *Journal of clinical microbiology* 62.3 (Mar. 2024).

[37] Abraham Tolley et al. "Cost-Effectiveness of Point-of-Care Diagnostics for AMR: A Systematic Review". In: *Journal of Antimicrobial Chemotherapy* 79.6 (June 2024), pp. 1248–1269.

[38] Andreea-Loredana Golli et al. "Prevalence of Multidrug-Resistant Pathogens Causing Bloodstream Infections in an Intensive Care Unit". In: *Infection and Drug Resistance* 15 (Oct. 2022), pp. 5981–5992.

[39] Mark S. Butler et al. "Antibiotics in the Clinical Pipeline as of December 2022". In: *The Journal of Antibiotics* 76.8 (Aug. 2023), pp. 431–473.

[40] Michael Schlander et al. "How Much Does It Cost to Research and Develop a New Drug? A Systematic Review and Assessment". In: *Pharmacoeconomics* 39.11 (2021), pp. 1243–1269.

[41] Trevor Mundel. *A New Model of Drug Discovery Could Change the Game on Superbugs*. June 2024.

[42] J. Cama et al. "To Push or To Pull? In a Post-COVID World, Supporting and Incentivizing Antimicrobial Drug Development Must Become a Governmental Priority". In: *ACS Infectious Diseases* 7.8 (Aug. 2021), pp. 2029–2042.

[43] Global AMR R&D Hub. *Dynamic Dashboard Antibacterials in Clinical Development*. https://dzifhelmholtzdashboard.azurewebsites.net/reports/pipelines/pipelines.

[44] Waqas Nasir Chaudhry et al. "Synergy and Order Effects of Antibiotics and Phages in Killing Pseudomonas Aeruginosa Biofilms". In: *PloS One* 12.1 (2017), e0168615.

[45] Iwona J. Fijalkowska, Roel M. Schaaper, and Piotr Jonczyk. "DNA Replication Fidelity in Escherichia Coli: A Multi-DNA Polymerase Affair". In: *FEMS microbiology reviews* 36.6 (Nov. 2012), pp. 1105–1121.

[46] Dan I. Andersson and Diarmaid Hughes. "Antibiotic Resistance and Its Cost: Is It Possible to Reverse Resistance?" In: *Nature Reviews Microbiology* 8.4 (Apr. 2010), pp. 260–271.

[47] Salvador Castañeda-Barba, Eva M. Top, and Thibault Stalder. "Plasmids, a Molecular Cornerstone of Antimicrobial Resistance in the One Health Era". In: *Nature Reviews Microbiology* 22.1 (Jan. 2024), pp. 18–32.

[48] Beata Kowalska-Krochmal and Ruth Dudek-Wicher. "The Minimum Inhibitory Concentration of Antibiotics: Methods, Interpretation, Clinical Relevance". In: *Pathogens* 10.2 (Feb. 2021), p. 165.

[49] Nagendra N. Mishra et al. "Daptomycin Resistance in Enterococci Is Associated with Distinct Alterations of Cell Membrane Phospholipid Content". In: *PloS One* 7.8 (2012), e43958.

[50] Draper P. "The Outer Parts of the Mycobacterial Envelope as Permeability Barriers". In: *Frontiers in bioscience : a journal and virtual library* 3 (Dec. 1998).

[51] Matthew Levitus, Ayesan Rewane, and Thomas B. Perera. "Vancomycin-Resistant Enterococci". In: *StatPearls*. Treasure Island (FL): StatPearls Publishing, 2024.

[52] Lucía Fernández and Robert E. W. Hancock. "Adaptive and Mutational Resistance: Role of Porins and Efflux Pumps in Drug Resistance". In: *Clinical Microbiology Reviews* 25.4 (Oct. 2012), pp. 661–681.

[53] Emily E. Whittle et al. "Efflux Impacts Intracellular Accumulation Only in Actively Growing Bacterial Cells". In: *mBio* 12.5 (Oct. 2021), e0260821.

[54] Arup Panda et al. "Genome-Wide Analysis of Horizontally Acquired Genes in the Genus Mycobacterium". In: *Scientific Reports* 8.1 (Oct. 2018), p. 14817.

[55] Ruchika Bhujbalrao and Ruchi Anand. "Deciphering Determinants in Ribosomal Methyltransferases That Confer Antimicrobial Resistance". In: *Journal of the American Chemical Society* 141.4 (Jan. 2019), pp. 1425–1429.

[56] Karen Bush. "Alarming β-Lactamase-Mediated Resistance in Multidrug-Resistant Enterobacteriaceae". In: *Current Opinion in Microbiology* 13.5 (Oct. 2010), pp. 558–564.

[57] Jesper Larsen et al. "Emergence of Methicillin Resistance Predates the Clinical Use of Antibiotics". In: *Nature* 602.7895 (Feb. 2022), pp. 135–141.

[58] Pavel A. Nazarov. "MDR Pumps as Crossroads of Resistance: Antibiotics and Bacteriophages". In: *Antibiotics* 11.6 (June 2022), p. 734.

[59] Xian-Zhi Li, Patrick Plésiat, and Hiroshi Nikaido. "The Challenge of Efflux-Mediated Antibiotic Resistance in Gram-Negative Bacteria". In: *Clinical Microbiology Reviews* 28.2 (Mar. 2015), pp. 337–418.

[60] Elizabeth M. Darby et al. "Molecular Mechanisms of Antibiotic Resistance Revisited". In: *Nature Reviews Microbiology* (Nov. 2022).

[61] P. S. Stewart and J. W. Costerton. "Antibiotic Resistance of Bacteria in Biofilms". In: *Lancet (London, England)* 358.9276 (July 2001), pp. 135–138.

[62] Lucas R. Hoffman et al. "Aminoglycoside Antibiotics Induce Bacterial Biofilm Formation". In: *Nature* 436.7054 (Aug. 2005), pp. 1171–1175.

[63] Nathalie Q. Balaban et al. "Bacterial Persistence as a Phenotypic Switch". In: *Science* 305.5690 (Sept. 2004), pp. 1622–1625.

[64] María Antonia Sánchez-Romero and Josep Casadesús. "Contribution of Phenotypic Heterogeneity to Adaptive Antibiotic Resistance". In: *Proceedings of the National Academy of Sciences* 111.1 (Jan. 2014), pp. 355–360.

[65] Rosemary A. Bamford et al. "Investigating the Physiology of Viable but Non-Culturable Bacteria by Microfluidics and Time-Lapse Microscopy". In: *BMC Biology* 15.1 (Dec. 2017), p. 121.

[66] Gaby E. Pfyffer and Franziska Wittwer. "Incubation Time of Mycobacterial Cultures: How Long Is Long Enough To Issue a Final Negative Report to the Clinician?" In: *Journal of Clinical Microbiology* 50.12 (Dec. 2012), pp. 4188–4189.

[67] Ina Gajic et al. "Antimicrobial Susceptibility Testing: A Comprehensive Review of Currently Used Methods". In: *Antibiotics* 11.4 (Apr. 2022), p. 427.

[68] Alistair McGregor et al. "The MicroScan WalkAway Diagnostic Microbiology System — an Evaluation". In: *Pathology* 27.2 (Jan. 1995), pp. 172–176.

[69] J.-L. Donay et al. "Evaluation of the Automated Phoenix System for Potential Routine Use in the Clinical Microbiology Laboratory". In: *Journal of Clinical Microbiology* 42.4 (Apr. 2004), pp. 1542–1546.

[70] Emma Jonasson, Erika Matuschek, and Gunnar Kahlmeter. "The EUCAST Rapid Disc Diffusion Method for Antimicrobial Susceptibility Testing Directly from Positive Blood Culture Bottles". In: *The Journal of Antimicrobial Chemotherapy* 75.4 (Apr. 2020), pp. 968–978.

[71] National Institute for Health and Care Excellence. *Suspected Sepsis: Recognition, Diagnosis and Early Management*. NICE Guideline NG51. Mar. 2024, p. 53.

[72] J. Barenfanger, C. Drake, and G. Kacich. "Clinical and Financial Benefits of Rapid Bacterial Identification and Antimicrobial Susceptibility Testing". In: *Journal of Clinical Microbiology* 37.5 (May 1999), pp. 1415–1418.

[73] Shawn H. MacVane and Hari P. Dwivedi. "Evaluating the Impact of Rapid Antimicrobial Susceptibility Testing for Bloodstream Infections: A Review of Actionability, Antibiotic Use and Patient Outcome Metrics". In: *The Journal of Antimicrobial Chemotherapy* 79.Supplement_1 (Sept. 2024), pp. i13–i25.

[74] Do Hyeon Park et al. "A Direct Rapid Phenotypic Antimicrobial Susceptibility Test Enables Early Selection of Optimal Antibiotics to Treat Bacteremia in COVID-19 Patients". In: *Infection & Chemotherapy* 53.4 (Dec. 2021), pp. 776–785.

[75] Ritu Banerjee et al. "Randomized Trial Evaluating Clinical Impact of RAPid IDentification and Susceptibility Testing for Gram-negative Bacteremia: RAPIDS-GN". In: *Clinical Infectious Diseases* 73.1 (July 2021), e39–e46.

[76] Chanu Rhee et al. "Prevalence of Antibiotic-Resistant Pathogens in Culture-Proven Sepsis and Outcomes Associated With Inadequate and Broad-Spectrum Empiric Antibiotic Use". In: *JAMA network open* 3.4 (Apr. 2020), e202899.

[77] Alex Van Belkum and W. Michael Dunne. "Next-Generation Antimicrobial Susceptibility Testing". In: *Journal of Clinical Microbiology* 51.7 (July 2013), pp. 2018–2024.

[78] U. Eigner et al. “Analysis of the Comparative Workflow and Performance Characteristics of the VITEK 2 and Phoenix Systems”. In: *Journal of Clinical Microbiology* 43.8 (Aug. 2005), pp. 3829–3834.

[79] Alex van Belkum et al. “Developmental Roadmap for Antimicrobial Susceptibility Testing Systems”. In: *Nature Reviews Microbiology* 17.1 (Jan. 2019), pp. 51–62.

[80] Lindsay G. Stevenson, Steven K. Drake, and Patrick R. Murray. “Rapid Identification of Bacteria in Positive Blood Culture Broths by Matrix-Assisted Laser Desorption Ionization-Time of Flight Mass Spectrometry”. In: *Journal of Clinical Microbiology* 48.2 (Feb. 2010), pp. 444–447.

[81] Simon Lévesque et al. “A Side by Side Comparison of Bruker Biotyper and VITEK MS: Utility of MALDI-TOF MS Technology for Microorganism Identification in a Public Health Reference Laboratory”. In: *PLOS ONE* 10.12 (Dec. 2015), e0144878.

[82] Katrin Sparbier et al. “MALDI Biotyper-Based Rapid Resistance Detection by Stable-Isotope Labeling”. In: *Journal of Clinical Microbiology* 51.11 (Dec. 2020), pp. 3741–3748.

[83] Hamza Rasheed et al. “Discrepancies between Phenotypic and Genotypic Identification Methods of Antibiotic Resistant Genes Harboring *Staphylococcus Aureus*”. In: *Microbial Pathogenesis* 184 (Nov. 2023), p. 106342.

[84] Suhail Ahmad et al. “Discordance across Phenotypic and Molecular Methods for Drug Susceptibility Testing of Drug-Resistant Mycobacterium Tuberculosis Isolates in a Low TB Incidence Country”. In: *PLOS ONE* 11.4 (Apr. 2016), e0153563.

[85] Michael Feldgarden et al. “Validating the AMRFinder Tool and Resistance Gene Database by Using Antimicrobial Resistance Genotype-Phenotype Correlations in a Collection of Isolates”. In: *Antimicrobial Agents and Chemotherapy* 63.11 (Oct. 2019), 10.1128/aac.00483–19.

[86] Kenneth P. Smith and James E. Kirby. “Rapid Susceptibility Testing Methods”. In: *Clinics in laboratory medicine* 39.3 (Sept. 2019), pp. 333–344.

[87] Özden Baltekin et al. “Antibiotic Susceptibility Testing in Less than 30 Min Using Direct Single-Cell Imaging”. In: *Proceedings of the National Academy of Sciences* 114.34 (Aug. 2017), pp. 9170–9175.

[88] Zining Hou et al. “Time Lapse Investigation of Antibiotic Susceptibility Using a Microfluidic Linear Gradient 3D Culture Device”. In: *Lab on a Chip* 14.17 (July 2014), pp. 3409–3418.

[89] Marlene Fredborg et al. “Real-Time Optical Antimicrobial Susceptibility Testing”. In: *Journal of Clinical Microbiology* 51.7 (Dec. 2020), pp. 2047–2053.

[90] Tae Hyun Kim et al. “Blood Culture-Free Ultra-Rapid Antimicrobial Susceptibility Testing”. In: *Nature* 632.8026 (Aug. 2024), pp. 893–902.

[91] D. T. Quach et al. “Bacterial Cytological Profiling (BCP) as a Rapid and Accurate Antimicrobial Susceptibility Testing Method for Staphylococcus Aureus”. In: *EBioMedicine* 4 (Feb. 2016), pp. 95–103.

[92] Jungil Choi et al. “Direct, Rapid Antimicrobial Susceptibility Test from Positive Blood Cultures Based on Microscopic Imaging Analysis”. In: *Scientific Reports* 7.1 (Apr. 2017), p. 1148.

[93] Alexander Zagajewski et al. “Deep Learning and Single-Cell Phenotyping for Rapid Antimicrobial Susceptibility Detection in Escherichia Coli”. In: *Communications Biology* 6.1 (Nov. 2023), pp. 1–12.

[94] Maxim Kalashnikov et al. “Rapid Phenotypic Stress-Based Microfluidic Antibiotic Susceptibility Testing of Gram-negative Clinical Isolates”. In: *Scientific Reports* 7.1 (Aug. 2017), p. 8031.

[95] Urszula Łapińska et al. “Fast Bacterial Growth Reduces Antibiotic Accumulation and Efficacy”. In: *eLife* 11 (June 2022). Ed. by Vaughn S Cooper and Gisela Storz, e74062.

[96] Kieran Mulroney et al. “Same-Day Confirmation of Infection and Antimicrobial Susceptibility Profiling Using Flow Cytometry”. In: *eBioMedicine* 82 (Aug. 2022).

[97] Benjamin Lundquist Thomsen et al. “Accurate and Fast Identification of Minimally Prepared Bacteria Phenotypes Using Raman Spectroscopy Assisted by Machine Learning”. In: *Scientific Reports* 12.1 (Sept. 2022), p. 16436.

[98] Chia-Ying Liu et al. “Rapid Bacterial Antibiotic Susceptibility Test Based on Simple Surface-Enhanced Raman Spectroscopic Biomarkers”. In: *Scientific Reports* 6 (Mar. 2016), p. 23375.

[99] Anna K. Boardman et al. “Rapid Detection of Bacteria from Blood with Surface-Enhanced Raman Spectroscopy”. In: *Analytical Chemistry* 88.16 (Aug. 2016), pp. 8026–8035.

[100] Daniel C. Spencer et al. “A Fast Impedance-Based Antimicrobial Susceptibility Test”. In: *Nature Communications* 11.1 (Oct. 2020), p. 5328.

[101] Sharath Narayana Iyengar et al. “Identifying Antibiotic-Resistant Strains via Cell Sorting and Elastic-Light-Scatter Phenotyping”. In: *Applied Microbiology and Biotechnology* 108.1 (July 2024), p. 406.

[102] Alexander Sturm et al. “Accurate and Rapid Antibiotic Susceptibility Testing Using a Machine Learning-Assisted Nanomotion Technology Platform”. In: *Nature Communications* 15.1 (Mar. 2024), p. 2037.

[103] Patrycja Zalas-Więcek, Tomasz Bogiel, and Eugenia Gospodarek-Komkowska. “The Accelerate Pheno™ System-A New Tool in Microbiological Diagnostics of Bloodstream Infections: A Pilot Study from Poland”. In: *Pathogens (Basel, Switzerland)* 11.12 (Nov. 2022), p. 1415.

[104] Kristin R. Baker et al. “Multicenter Evaluation of the Selux Next-Generation Phenotyping Antimicrobial Susceptibility Testing System”. In: *Journal of Clinical Microbiology* 62.1 (Dec. 2023), e00546–23.

[105] James W Snyder, Nadia Chaudry, and Wesley Hoffman. “345. Performance Evaluation of the LifeScale AST Rapid Automated Antimicrobial Susceptibility System”. In: *Open Forum Infectious Diseases* 9.Supplement_2 (Dec. 2022), ofac492.423.

[106] Jan Esse et al. “Rapid Phenotypic Antimicrobial Susceptibility Testing of Gram-negative Rods Directly from Positive Blood Cultures Using the Novel Q-linea ASTar System”. In: *Journal of Clinical Microbiology* 61.11 (Oct. 2023), e00549–23.

[107] *BARDA-Supported Antimicrobial Susceptibility Test (AST) Receives FDA Clearance*. https://medicalcountermeasures.gov/newsroom/2024/vitek/. Aug. 2024.

[108] Huang Th, Tzeng Yl, and Dickson Rm. “FAST: Rapid Determinations of Antibiotic Susceptibility Phenotypes Using Label-Free Cytometry”. In: *Cytometry. Part A : the journal of the International Society for Analytical Cytology* 93.6 (June 2018).

[109] Thomas M. Hooton et al. “Voided Midstream Urine Culture and Acute Cystitis in Premenopausal Women”. In: *The New England Journal of Medicine* 369.20 (Nov. 2013), pp. 1883–1891.

[110] L J La Scolea and D Dryja. “Quantitation of Bacteria in Cerebrospinal Fluid and Blood of Children with Meningitis and Its Diagnostic Significance”. In: *Journal of Clinical Microbiology* 19.2 (Feb. 1984), pp. 187–190.

[111] Sushmita Sridhar et al. “High-Content Imaging to Phenotype Antimicrobial Effects on Individual Bacteria at Scale”. In: *mSystems* (May 2021).

[112] Htut Htut Htoo et al. “Bacterial Cytological Profiling as a Tool To Study Mechanisms of Action of Antibiotics That Are Active against Acinetobacter Baumannii”. In: *Antimicrobial Agents and Chemotherapy* 63.4 (Apr. 2019), e02310–18.

[113] Poochit Nonejuie et al. “Bacterial Cytological Profiling Rapidly Identifies the Cellular Pathways Targeted by Antibacterial Molecules”. In: *Proceedings of the National Academy of Sciences of the United States of America* 110.40 (Oct. 2013), pp. 16169–16174.

[114] Lidia Araújo-Bazán et al. “Cytological Profile of Antibacterial FtsZ Inhibitors and Synthetic Peptide MciZ”. In: *Frontiers in Microbiology* 7 (2016), p. 1558.

[115] Sannah Zoffmann et al. “Machine Learning-Powered Antibiotics Phenotypic Drug Discovery”. In: *Scientific Reports* 9.1 (Mar. 2019), p. 5013.

[116] Maxime Mistretta et al. “Dynamic Microfluidic Single-Cell Screening Identifies Pheno-Tuning Compounds to Potentiate Tuberculosis Therapy”. In: *Nature Communications* 15.1 (May 2024), p. 4175.

[117] Tuan-Anh Tran et al. “Combining Machine Learning with High-Content Imaging to Infer Ciprofloxacin Susceptibility in Isolates of Salmonella Typhimurium”. In: *Nature Communications* 15.1 (June 2024), p. 5074.

[118] Danilo Bzdok, Naomi Altman, and Martin Krzywinski. “Statistics versus Machine Learning”. In: *Nature Methods* 15.4 (Apr. 2018), pp. 233–234.

[119] M. I. Jordan and T. M. Mitchell. “Machine Learning: Trends, Perspectives, and Prospects”. In: *Science (New York, N.Y.)* 349.6245 (July 2015), pp. 255–260.

[120] Danilo Bzdok, Martin Krzywinski, and Naomi Altman. “Machine Learning: A Primer”. In: *Nature methods* 14.12 (Nov. 2017), pp. 1119–1120.

[121] Alexandre Drouin et al. “Interpretable Genotype-to-Phenotype Classifiers with Performance Guarantees”. In: *Scientific Reports* 9 (Mar. 2019), p. 4071.

[122] Marcus Nguyen et al. “Using Machine Learning To Predict Antimicrobial MICs and Associated Genomic Features for Nontyphoidal Salmonella”. In: *Journal of Clinical Microbiology* 57.2 (Feb. 2019), e01260–18.

[123] Yang Yang et al. “DeepAMR for Predicting Co-Occurrent Resistance of Mycobacterium Tuberculosis”. In: *Bioinformatics (Oxford, England)* 35.18 (Sept. 2019), pp. 3240–3249.

[124] Samaneh Kouchaki et al. “Application of Machine Learning Techniques to Tuberculosis Drug Resistance Analysis”. In: *Bioinformatics* 35.13 (July 2019), pp. 2276–2282.

[125] Jan Zrimec et al. “Deep Learning Suggests That Gene Expression Is Encoded in All Parts of a Co-Evolving Interacting Gene Regulatory Structure”. In: *Nature Communications* 11.1 (Dec. 2020), p. 6141.

[126] Yann LeCun, Yoshua Bengio, and Geoffrey Hinton. “Deep Learning”. In: *Nature* 521.7553 (May 2015), pp. 436–444.

[127] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E Hinton. “ImageNet Classification with Deep Convolutional Neural Networks”. In: *Advances in Neural Information Processing Systems*. Vol. 25. Curran Associates, Inc., 2012.

[128] Karen Simonyan, Andrea Vedaldi, and Andrew Zisserman. *Deep Inside Convolutional Networks: Visualising Image Classification Models and Saliency Maps*. Apr. 2014.

[129] Mingzhu Liang et al. “Low-Dose CT Screening for Lung Cancer: Computer-aided Detection of Missed Lung Cancers”. In: *Radiology* 281.1 (Oct. 2016), pp. 279–288.

[130] Thijs Kooi et al. “Large Scale Deep Learning for Computer Aided Detection of Mammographic Lesions”. In: *Medical Image Analysis* 35 (Jan. 2017), pp. 303–312.

[131] Ke Yan et al. *DeepLesion: Automated Deep Mining, Categorization and Detection of Significant Radiology Image Findings Using Large-Scale Clinical Lesion Annotations*. Oct. 2017.

[132] Pierre Buyssens, Abderrahim Elmoataz, and Olivier Lezoray. “Multiscale Convolutional Neural Networks for Vision-based Classification of Cells”. In: *Asian Conference on Computer Vision (ACCV)*. Vol. 7725. LNCS. South Korea, 2012, pp. 342–352.

[133] William J Godinez et al. “A Multi-Scale Convolutional Neural Network for Phenotyping High-Content Cellular Images”. In: *Bioinformatics* 33.13 (July 2017), pp. 2010–2019.

[134] Chao Luo et al. “How Does the Data Set Affect CNN-based Image Classification Performance?” In: *2018 5th International Conference on Systems and Informatics (ICSAI)*. Nov. 2018, pp. 361–366.

[135] Tyler M. Earnest et al. “Ribosome Biogenesis in Replicating Cells: Integration of Experiment and Theory”. In: *Biopolymers* 105.10 (2016), pp. 735–751.

[136] Marat M. Yusupov et al. “Crystal Structure of the Ribosome at 5.5 Å Resolution”. In: *Science* 292.5518 (May 2001), pp. 883–896.

[137] J. H. Cate et al. “X-Ray Crystal Structures of 70S Ribosome Functional Complexes”. In: *Science (New York, N.Y.)* 285.5436 (Sept. 1999), pp. 2095–2104.

[138] P Traub and M Nomura. “Structure and Function of E. Coli Ribosomes. V. Reconstitution of Functionally Active 30S Ribosomal Particles from RNA and Proteins.” In: *Proceedings of the National Academy of Sciences of the United States of America* 59.3 (Mar. 1968), pp. 777–784.

[139] Zoe L Watson et al. “Structure of the Bacterial Ribosome at 2 Å Resolution”. In: *eLife* 9 (Sept. 2020). Ed. by Sjors HW Scheres et al., e60482.

[140] Xiongfeng Dai et al. “Reduction of Translating Ribosomes Enables Escherichia Coli to Maintain Elongation Rates during Slow Growth”. In: *Nature Microbiology* 2.2 (Dec. 2016), pp. 1–9.

[141] Hans Bremer and Patrick P. Dennis. “Modulation of Chemical Composition and Other Parameters of the Cell at Different Exponential Growth Rates”. In: *EcoSal Plus* 3.1 (Oct. 2008).

[142] N. O. Kjeldgaard, O. Maaloe, and M. Schaechter. “The Transition between Different Physiological States during Balanced Growth of Salmonella Typhimurium.” In: *Journal of general microbiology* 19.3 (Dec. 1958), pp. 607–616.

[143] M. Schaechter, O. MaalØe, and N. O.YR 1958 Kjeldgaard. “Dependency on Medium and Temperature of Cell Size and Chemical Composition during Balanced Growth of Salmonella Typhimurium”. In: *Microbiology* 19.3 (1958), pp. 592–606.

[144] Somenath Bakshi et al. “Superresolution Imaging of Ribosomes and RNA Polymerase in Live Escherichia Coli Cells”. In: *Molecular Microbiology* 85.1 (2012), pp. 21–38.

[145] Arash Sanamrad et al. “Single-Particle Tracking Reveals That Free Ribosomal Subunits Are Not Excluded from the Escherichia Coli Nucleoid”. In: *Proceedings of the National Academy of Sciences* 111.31 (Aug. 2014), pp. 11413–11418.

[146] Somenath Bakshi et al. “Time-Dependent Effects of Transcription- and Translation-Halting Drugs on the Spatial Distributions of the Escherichia Coli Chromosome and Ribosomes”. In: *Molecular Microbiology* 94.4 (2014), pp. 871–887.

[147] Somenath Bakshi et al. “Nonperturbative Imaging of Nucleoid Morphology in Live Bacterial Cells during an Antimicrobial Peptide Attack”. In: *Applied and Environmental Microbiology* 80.16 (Aug. 2014), pp. 4977–4986.

[148] Jagannath Mondal et al. “Entropy-Based Mechanism of Ribosome-Nucleoid Segregation in E. Coli Cells”. In: *Biophysical Journal* 100.11 (June 2011), pp. 2605–2613.

[149] Alexander Rich. “A Hybrid Helix Containing Both Deoxyribose and Ribose Polynucleotides and Its Relation to the Transfer of Information between the Nucleic Acids”. In: *Proceedings of the National Academy of Sciences* 46.8 (Aug. 1960), pp. 1044–1053.

[150] G Milman, R Langridge, and M J Chamberlin. “The Structure of a DNA-RNA Hybrid.” In: *Proceedings of the National Academy of Sciences* 57.6 (June 1967), pp. 1804–1810.

[151] Alexander P. Young, Daniel J. Jackson, and Russell C. Wyeth. "A Technical Review and Guide to RNA Fluorescence in Situ Hybridization". In: *PeerJ* 8 (Mar. 2020), e8806.

[152] R. I. Amann. "In Situ Identification of Micro-Organisms by Whole Cell Hybridization with rRNA-targeted Nucleic Acid Probes". In: *Molecular Microbial Ecology Manual*. Ed. by A. D. C. Akkerman, J. D. Van Elsas, and F. J. De Bruijn. Dordrecht: Kluwer Academic Publishers, 1995, pp. 1–15.

[153] Volkhard A. J. Kempf, Karlheinz Trebesius, and Ingo B. Autenrieth. "Fluorescent In Situ Hybridization Allows Rapid Identification of Microorganisms in Blood Cultures". In: *Journal of Clinical Microbiology* 38.2 (Feb. 2000), pp. 830–838.

[154] Sven Poppert et al. "Rapid Diagnosis of Bacterial Meningitis by Real-Time PCR and Fluorescence In Situ Hybridization". In: *Journal of Clinical Microbiology* 43.7 (July 2005), pp. 3390–3397.

[155] J.G.J. Bauman et al. "A New Method for Fluorescence Microscopical Localization of Specific DNA Sequences by in Situ Hybridization of Fluorochrome-Labelled RNA". In: *Experimental Cell Research* 128.2 (Aug. 1980), pp. 485–490.

[156] Alex M. Valm et al. "Systems-Level Analysis of Microbial Community Organization through Combinatorial Labeling and Spectral Imaging". In: *Proceedings of the National Academy of Sciences* 108.10 (Mar. 2011), pp. 4152–4157.

[157] A. M. Femino et al. "Visualization of Single RNA Transcripts in Situ". In: *Science (New York, N.Y.)* 280.5363 (Apr. 1998), pp. 585–590.

[158] Kok Hao Chen et al. "Spatially Resolved, Highly Multiplexed RNA Profiling in Single Cells". In: *Science* 348.6233 (Apr. 2015).

[159] Hao Shi et al. "Highly Multiplexed Spatial Mapping of Microbial Communities". In: *Nature* 588.7839 (Dec. 2020), pp. 676–681.

[160] Stelios Chatzimichail et al. "Rapid Identification of Bacterial Isolates Using Microfluidic Adaptive Channels and Multiplexed Fluorescence Microscopy". In: *Lab on a Chip* (Sept. 2024).

[161] Vinodh Kandavalli et al. "Rapid Antibiotic Susceptibility Testing and Species Identification for Mixed Samples". In: *Nature Communications* 13.1 (Oct. 2022), p. 6215.

[162] Zachary D. Blount. "The Unexhausted Potential of E. Coli". In: *eLife* 4 (Mar. 2015), e05826.

[163] James B. Kaper, James P. Nataro, and Harry L. T. Mobley. "Pathogenic Escherichia Coli". In: *Nature Reviews Microbiology* 2.2 (Feb. 2004), pp. 123–140.

[164] Heloisa H. Nascimento et al. "Phenotypic and Genotypic Characteristics Associated with Biofilm Formation in Clinical Isolates of Atypical Enteropathogenic Escherichia Coli (aEPEC) Strains". In: *BMC microbiology* 14 (July 2014), p. 184.

[165] Valeria Bortolaia et al. "ResFinder 4.0 for Predictions of Phenotypes from Genotypes". In: *Journal of Antimicrobial Chemotherapy* 75.12 (2020), pp. 3491–3500.

[166] Katherine A. Lau et al. "Proficiency Testing for Bacterial Whole Genome Sequencing in Assuring the Quality of Microbiology Diagnostics in Clinical and Public Health Laboratories". In: *Pathology* 53.7 (Dec. 2021), pp. 902–911.

[167] Sam Lipworth et al. "Antimicrobial Resistance Genes and Clonal Success in Escherichia Coli Isolates Causing Bloodstream Infection". In: *The Lancet Microbe* 2.10 (Oct. 2021), e492.

[168] Gary W. Procop and Elmer W. Koneman. *Koneman's Color Atlas and Textbook of Diagnostic Microbiology*. Seventh edition. Philadelphia: Wolters Kluwer Health, 2017.

[169] European Committee on Antimicrobial Susceptibility Testing. *EUCAST: Clinical Breakpoints and Dosing of Antibiotics*. Tech. rep. Version 14.0. Jan. 2024.

[170] Samuel Lipworth et al. "Ten-Year Longitudinal Molecular Epidemiology Study of Escherichia Coli and Klebsiella Species Bloodstream Infections in Oxfordshire, UK". In: *Genome Medicine* 13.1 (Sept. 2021), p. 144.

[171] Alex van Belkum et al. "Innovative and Rapid Antimicrobial Susceptibility Testing Systems". In: *Nature Reviews Microbiology* 18.5 (May 2020), pp. 299–311.

[172] Piers Turner. *Napari-Bacseg: Bacterial Segmentation and Analysis Platform than Can Inport/Export Files in Multiple Formats. Integrating Many Tools Such as Cellpose, ColiCoords, Oufti/MicrobeTracker.*

[173] Carsen Stringer et al. "Cellpose: A Generalist Algorithm for Cellular Segmentation". In: *Nature Methods* 18.1 (Jan. 2021), pp. 100–106.

[174] A. Buslaev et al. "Albumentations: Fast and Flexible Image Augmentations". In: *ArXiv e-prints* (2018).

[175] Adam Paszke et al. "PyTorch: An Imperative Style, High-Performance Deep Learning Library". In: *Advances in Neural Information Processing Systems 32*. Ed. by H. Wallach et al. Curran Associates, Inc., 2019, pp. 8024–8035.

[176] Takuya Akiba et al. *Optuna: A Next-generation Hyperparameter Optimization Framework*. July 2019.

[177] Mingxing Tan and Quoc V. Le. *EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks*. Sept. 2020.

[178] Diederik P. Kingma and Jimmy Ba. *Adam: A Method for Stochastic Optimization*. Jan. 2017.

[179] Ramprasaath R. Selvaraju et al. "Grad-CAM: Visual Explanations from Deep Networks via Gradient-Based Localization". In: *2017 IEEE International Conference on Computer Vision (ICCV)*. Venice: IEEE, Oct. 2017, pp. 618–626.

[180] Jochem H. Smit et al. "ColiCoords: A Python Package for the Analysis of Bacterial Fluorescence Microscopy Data". In: *PLOS ONE* 14.6 (June 2019), e0217524.

[181] Mehmet A. Orman and Mark P. Brynildsen. "Inhibition of Stationary Phase Respiration Impairs Persister Formation in E. Coli". In: *Nature Communications* 6.1 (Aug. 2015), p. 7983.

[182] Aleksander Zagajewski et al. *Deep Antimicrobial Susceptibility Phenotyping (DASP) Training and Evaluation Dataset, and Trained Models.* 2023.

[183] Maria Schei Haugan et al. “Growth Rate of Escherichia Coli During Human Urinary Tract Infection: Implications for Antibiotic Effect”. In: *Antibiotics (Basel, Switzerland)* 8.3 (July 2019), p. 92.

[184] Mathew Stracy et al. “Live-Cell Superresolution Microscopy Reveals the Organization of RNA Polymerase in the Bacterial Nucleoid”. In: *Proceedings of the National Academy of Sciences* 112.32 (Aug. 2015), E4390–E4399.

[185] Yao-Kuan Wang et al. “Comparison of Escherichia Coli Surface Attachment Methods for Single-Cell Microscopy”. In: *Scientific Reports* 9 (Dec. 2019), p. 19418.

[186] Alison Farrar et al. *Ribosome Phenotypes Enable Rapid Antibiotic Susceptibility Testing in Escherichia Coli*. June 2024.

[187] Suckjoon Jun and Bela Mulder. “Entropy-Driven Spatial Organization of Highly Confined Polymers: Lessons for the Bacterial Chromosome”. In: *Proceedings of the National Academy of Sciences* 103.33 (Aug. 2006), pp. 12388–12393.

[188] Ding Jun Jin, Cedric Cagliero, and Yan Ning Zhou. “Role of RNA Polymerase and Transcription in the Organization of the Bacterial Nucleoid”. In: *Chemical Reviews* 113.11 (Nov. 2013), pp. 8662–8682.

[189] William T. Gray et al. “Nucleoid Size Scaling and Intracellular Organization of Translation across Bacteria”. In: *Cell* 177.6 (May 2019), 1632–1648.e20.

[190] Christoph Spahn et al. “The Nucleoid of Rapidly Growing Escherichia Coli Localizes Close to the Inner Membrane and Is Organized by Transcription, Translation, and Cell Geometry”. In: *Nature Communications* 16.1 (Apr. 2025), p. 3732.

[191] Qian Chai et al. “Organization of Ribosomes and Nucleoids in Escherichia Coli Cells during Growth and in Quiescence”. In: *The Journal of Biological Chemistry* 289.16 (Apr. 2014), pp. 11342–11352.

[192] Jungil Choi et al. “A Rapid Antimicrobial Susceptibility Test Based on Single-Cell Morphological Analysis”. In: *Science Translational Medicine* 6.267 (Dec. 2014), 267ra174–267ra174.

[193] Trevor Hastie, Robert Tibshirani, and Jerome Friedman. “Cross Validation”. In: *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. 10. [print.], (corr. as of 4. print.) Springer Series in Statistics. New York, NY: Springer, 2008, pp. 241–249.

[194] Sebastien C. Wong et al. *Understanding Data Augmentation for Classification: When to Warp?* Nov. 2016.

[195] Nicholas A. Rossi, Imane El Meouche, and Mary J. Dunlop. “Forecasting Cell Fate during Antibiotic Exposure Using Stochastic Gene Expression”. In: *Communications Biology* 2.1 (July 2019), pp. 1–7.

[196] Krister Vikedal et al. “RecN and RecA Orchestrate an Ordered DNA Supercompaction Response Following Ciprofloxacin-Induced DNA Damage in *Escherichia Coli*”. In: *Nucleic Acids Research* 53.10 (May 2025), gkaf437.

[197] J. Buijs et al. “Concentration-Dependency of β-Lactam-Induced Filament Formation in Gram-negative Bacteria”. In: *Clinical Microbiology and Infection* 14.4 (Apr. 2008), pp. 344–349.

[198] Marlene Fredborg et al. “Automated Image Analysis for Quantification of Filamentous Bacteria”. In: *BMC microbiology* 15 (Nov. 2015), p. 255.

[199] Frank Schluenzen et al. “The Antibiotic Kasugamycin Mimics mRNA Nucleotides to Destabilize tRNA Binding and Inhibit Canonical Translation Initiation”. In: *Nature Structural & Molecular Biology* 13.10 (Oct. 2006), pp. 871–878.

[200] Jennifer B. Rattray et al. “Machine Learning Identification of Pseudomonas Aeruginosa Strains from Colony Image Data”. In: *PLOS Computational Biology* 19.12 (Dec. 2023), e1011699.

[201] J.L. Pech-Pacheco et al. “Diatom Autofocusing in Brightfield Microscopy: A Comparative Study”. In: *Proceedings 15th International Conference on Pattern Recognition. ICPR-2000*. Vol. 3. Barcelona, Spain: IEEE Comput. Soc, 2000, pp. 314–317.

[202] Pina M. Fratamico et al. “Advances in Molecular Serotyping and Subtyping of Escherichia Coli”. In: *Frontiers in Microbiology* 7 (May 2016), p. 644.

[203] David A. Rasko et al. “The Pangenome Structure of Escherichia Coli: Comparative Genomic Analysis of E. Coli Commensal and Pathogenic Isolates”. In: *Journal of Bacteriology* 190.20 (Oct. 2008), pp. 6881–6893.

[204] Jolanta Sarowska et al. “Virulence Factors, Prevalence and Potential Transmission of Extraintestinal Pathogenic Escherichia Coli Isolated from Different Sources: Recent Reports”. In: *Gut Pathogens* 11.1 (Feb. 2019), p. 10.

[205] R. Orenstein and E. S. Wong. “Urinary Tract Infections in Adults”. In: *American Family Physician* 59.5 (Mar. 1999), pp. 1225–1234, 1237.

[206] Simone Bagel et al. “Impact of gyrA and parCMutations on Quinolone Resistance, Doubling Time, and Supercoiling Degree of Escherichia Coli”. In: *Antimicrobial Agents and Chemotherapy* 43.4 (Apr. 1999), pp. 868–875.

[207] Dong Chan Moon et al. “Emergence of a New Mutation and Its Accumulation in the Topoisomerase IV Gene Confers High Levels of Resistance to Fluoroquinolones in *Escherichia Coli* Isolates”. In: *International Journal of Antimicrobial Agents* 35.1 (Jan. 2010), pp. 76–79.

[208] Anna Johnning et al. “Resistance Mutations in gyrA and parC Are Common in Escherichia Communities of Both Fluoroquinolone-Polluted and Uncontaminated Aquatic Environments”. In: *Frontiers in Microbiology* 6 (Dec. 2015), p. 1355.

[209] Douglas L. Huseby et al. “Mutation Supply and Relative Fitness Shape the Genotypes of Ciprofloxacin-Resistant Escherichia Coli”. In: *Molecular Biology and Evolution* 34.5 (May 2017), pp. 1029–1039.

[210] Stelios Chatzimichail et al. *Rapid Identification of Bacterial Isolates Using Microfluidic Adaptive Channels and Multiplexed Fluorescence Microscopy*. July 2023.

[211] M. C. Sulavik, L. F. Gambino, and P. F. Miller. “The MarR Repressor of the Multiple Antibiotic Resistance (Mar) Operon in Escherichia Coli: Prototypic Member of a Family of Bacterial Regulatory Proteins Involved in Sensing Phenolic Compounds.” In: *Molecular Medicine* 1.4 (May 1995), pp. 436–446.

[212] Margret Oethinger et al. “Overexpression of the marA or soxS Regulatory Gene in Clinical Topoisomerase Mutants of Escherichia Coli”. In: *Antimicrobial Agents and Chemotherapy* 42.8 (Aug. 1998), pp. 2089–2094.

[213] Thomas L. Walsh et al. “Impact of an Antimicrobial Stewardship Program-bundled Initiative Utilizing Accelerate Pheno™ System in the Management of Patients with Aerobic Gram-negative Bacilli Bacteremia”. In: *Infection* 49.3 (June 2021), pp. 511–519.

[214] Catherine Anne Hogan et al. “Impact of Rapid Antimicrobial Susceptibility Testing in Gram-Negative Rod Bacteremia: A Quasi-experimental Study”. In: *Journal of Clinical Microbiology* 58.9 (Aug. 2020), e00360–20.

[215] Alison Farrar et al. “Infection Inspection: Using the Power of Citizen Science for Image-Based Prediction of Antibiotic Resistance in Escherichia Coli Treated with Ciprofloxacin”. In: *Scientific Reports* 14.1 (Aug. 2024), p. 19543.

[216] Caren Cooper et al. “Citizen Science as a Tool for Conservation in Residential Ecosystems”. In: *Ecology and Society* 12.2 (Oct. 2007).

[217] Oxford English Dictionary. *Citizen Science, n.* July 2023.

[218] Helen Spiers et al. “Deep Learning for Automatic Segmentation of the Nuclear Envelope in Electron Microscopy Data, Trained with Volunteer Segmentations”. In: *Traffic (Copenhagen, Denmark)* 22.7 (July 2021), pp. 240–253.

[219] Seth Cooper et al. “Predicting Protein Structures with a Multiplayer Online Game”. In: *Nature* 466.7307 (Aug. 2010), pp. 756–760.

[220] Philip W Fowler et al. “A Crowd of BashTheBug Volunteers Reproducibly and Accurately Measure the Minimum Inhibitory Concentrations of 13 Antitubercular Drugs from Photographs of 96-Well Broth Microdilution Plates”. In: *eLife* 11 (May 2022). Ed. by Bavesh D Kana, e75046.

[221] The Cornell Lab. *Merlin Bird ID*. https://merlin.allaboutbirds.org.

[222] Kim Van Vliet and Claybourne Moore. “Citizen Science Initiatives: Engaging the Public and Demystifying Science”. In: *Journal of Microbiology & Biology Education* 17.1 (Mar. 2016), pp. 13–16.

[223] The Zooniverse. *What Is the Zooniverse?* https://www.zooniverse.org/about.

[224] J. Woodcock et al. “Crowdsourcing Citizen Science: Exploring the Tensions between Paid Professionals and Users”. In: *Journal of Peer Production* 10 (2017).

[225] Catherine Hoffman et al. “SciStarter 2.0: A Digital Platform to Foster and Study Sustained Engagement in Citizen Science”. In: *Analyzing the Role of Citizen Science in Modern Research*. IGI Global, 2017, pp. 50–61.

[226] Josh Aaron Miller et al. “Introducing Foldit Education Mode”. In: *Nature Structural & Molecular Biology* 27.9 (Sept. 2020), pp. 769–770.

[227] Hai-Ying Liu et al. “Citizen Science Platforms”. In: *The Science of Citizen Science*. Ed. by Katrin Vohland et al. Cham: Springer International Publishing, 2021, pp. 439–459.

[228] Helen Spiers et al. “Everyone Counts? Design Considerations in Online Citizen Science”. In: *Journal of Science Communication* 18.1 (Jan. 2019), A04.

[229] L. M. Fisher et al. “Ciprofloxacin and the Fluoroquinolones. New Concepts on the Mechanism of Action and Resistance”. In: *The American Journal of Medicine* 87.5A (Nov. 1989), 2S–8S.

[230] Jordi Vila. “Fluoroquinolone Resistance”. In: *Frontiers in Antimicrobial Resistance*. John Wiley & Sons, Ltd, 2005. Chap. 4, pp. 41–52.

[231] European Committee on Antimicrobial Susceptibility Testing. “Methods for the Determination of Susceptibility of Bacteria to Antimicrobial Agents. Terminology”. In: *Clinical Microbiology and Infection: The Official Publication of the European Society of Clinical Microbiology and Infectious Diseases* 4.5 (May 1998), p. 291.

[232] Kaiming He et al. “Mask R-CNN”. In: *2017 IEEE International Conference on Computer Vision (ICCV)*. Oct. 2017, pp. 2980–2988.

[233] Waleed Abdulla. *Mask R-CNN for Object Detection and Instance Segmentation on Keras and TensorFlow*. Matterport, Inc. 2017.

[234] Jari Oksanen et al. *Vegan: Community Ecology Package*. May 2024.

[235] Hadley Wickham. *Ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York, 2016.

[236] Anne E. Carpenter et al. “CellProfiler: Image Analysis Software for Identifying and Quantifying Cell Phenotypes”. In: *Genome Biology* 7.10 (2006), R100.

[237] David R. Stirling et al. “CellProfiler 4: Improvements in Speed, Utility and Usability”. In: *BMC Bioinformatics* 22.1 (Sept. 2021), p. 433.

[238] *SQLite*. Hwaci.

[239] Mauricio Piacentini. *DB Browser for SQLite*.

[240] Alboukadel Kassambara. *Ggpubr: ’ggplot2’ Based Publication Ready Plots*. Feb. 2023.

[241] Alboukadel Kassambara. *Rstatix: Pipe-Friendly Framework for Basic Statistical Tests*. Feb. 2023.

[242] Charles R. Harris et al. “Array Programming with NumPy”. In: *Nature* 585.7825 (Sept. 2020), pp. 357–362.

[243] Fabian Pedregosa et al. “Scikit-Learn: Machine Learning in Python”. In: *Journal of Machine Learning Research* 12.85 (2011), pp. 2825–2830.

[244] The R Core Team. *Prcomp*.

[245] Scott M Lundberg and Su-In Lee. “A Unified Approach to Interpreting Model Predictions”. In: *Advances in Neural Information Processing Systems*. Vol. 30. Curran Associates, Inc., 2017.

[246] The Zooniverse. *Zooniverse Project Builder*.

[247] Layal Liverpool. “Citizen Science: Get the Bug”. In: *New Scientist* 257.3430 (Mar. 2023), p. 51.

[248] Joseph L. Gastwirth. “The Estimation of the Lorenz Curve and Gini Index”. In: *The Review of Economics and Statistics* 54.3 (1972), pp. 306–316.

[249] Amy L. Leber et al. “Multicenter Evaluation of BioFire FilmArray Respiratory Panel 2 for Detection of Viruses and Bacteria in Nasopharyngeal Swab Samples”. In: *Journal of Clinical Microbiology* 56.6 (June 2018). Ed. by Yi-Wei Tang, e01945–17.

[250] *Multi-Center Evaluation of the BioFire FilmArray Blood Culture Identification 2 Panel for the Detection of Microorganisms and Resistance Markers in Positive Blood Cultures*. Poster. 2019.

[251] Dorothee Maria Gescher et al. “Fluorescence in Situ Hybridisation (FISH) Accelerates Identification of Gram-positive Cocci in Positive Blood Cultures”. In: *International Journal of Antimicrobial Agents* 32 (Nov. 2008), S51–S59.

[252] Carsten Roller et al. “In Situprobing of Gram-positive Bacteria with High DNA G + C Content Using 23S rRNA-targeted Oligonucleotides”. In: *Microbiology* 140.10 (1994), pp. 2849–2858.

[253] T. Thurnheer, R. Gmür, and B. Guggenheim. “Multiplex FISH Analysis of a Six-Species Bacterial Biofilm”. In: *Journal of Microbiological Methods* 56.1 (Jan. 2004), pp. 37–47.

[254] Linnéa Garoff, Kavita Yadav, and Diarmaid Hughes. “Increased Expression of Qnr Is Sufficient to Confer Clinical Resistance to Ciprofloxacin in Escherichia Coli”. In: *Journal of Antimicrobial Chemotherapy* 73.2 (Feb. 2018), pp. 348–352.

A

Data Availability and Ethics

A.1 Software and Data

The datasets for Chapters 3 and 4 are available at https://zenodo.org/records/11656505. The code for training ribosome phenotype models is available at https://github.com/KapanidisLab/ribosome_phenotype_classification.

The raw data for Chapter 5 is available from the Oxford University Research Archive https://ora.ox.ac.uk/objects/uuid:12153432-e8b3-4398-a395-abfb980bd84e. The segmented single cell image sand classification metadata are available at https://doi.org/10.5281/zenodo.10301352.

The code used for Infection Inspection project data analysis is available at https://github.com/KapanidisLab/infection_inspection.

A.2 Ethics approval

This work was reviewed by the University of Oxford's Research Governance, Ethics and Assurance team (Oxford, UK). The bacterial cells imaged and shared for classification were obtained as part of a separate study with ethical approval (London – Queens Square Research Ethics Committee, REC ref: 17/LO/1420), analysed anonymously as part of this project and thereby not subject to the Department of Health's UK Policy Framework for Health and Social Care Research (2017), and not requiring further sponsorship, research ethics review, nor Health Research Authority (HRA) approval. Citizen scientists who contributed classifications to the Infection Inspection project were deemed contributors and not research participants.