Strathclyde Institute of Pharmacy and

Biomedical Sciences

University of Strathclyde

Glasgow G4 0RE

Submitted September 2025

Barcodes and bioreactors:

Exploring strain stability and genome dynamics in *Streptomyces clavuligerus*

James Timothy Croxford

Principal Investigator: Professor Paul A Hoskisson

Secondary Supervisor: Professor Iain S Hunter

Doctoral Thesis submitted in fulfilment of the requirement for the degree of Doctor of Philosophy.

Declaration of Authenticity and Author's Rights

This thesis is the result of the author's original research. It has been composed by the author and has not been previously submitted for examination which has led to the award of a degree.

Signed:

Date: 25/9/25

Previously published material

Larcombe, D.E., Braes, R.E., Croxford, J.T., Wilson, J.W., Figurski, D.H., Hoskisson, P.A., 2024. Sequence and origin of the *Streptomyces* intergenetic-conjugation helper plasmid pUZ8002. Access Microbiology 6. https://doi.org/10.1099/acmi.0.000808.v3

Contribution - Formal analysis and draft preparation/review.

Accepted Abstracts

Croxford, J.T., Kendrew, S., Huckle, B., Hunter, I.S. & Hoskisson, P.A., 2025. Barcodes and Bioreactors: Using bar-seq to understand the genome dynamics of *S. clavuligerus* during industrial fermentations. Oral presentation at the 20th Symposium of the International Society for the Biology of Actinomycetes (ISBA), Leiden, Netherlands, June 2025.

Croxford, J.T., Kendrew, S., Huckle, B., Hunter, I.S. & Hoskisson, P.A., 2025. Barcodes & Bioreactors: Tracking Fermentation Stability. Poster presentation at the Microbiology Society Annual Conference, Liverpool, UK, April 2025.

Croxford, J.T., Kendrew, S., Huckle, B., Hunter, I.S. & Hoskisson, P.A., 2025. Understanding the genome dynamics of *Streptomyces clavuligerus* during industrial fermentation. Oral and poster presentations at the IBioIC Annual Conference, Glasgow, UK, March 2025.

Croxford, J.T., Kendrew, S., Huckle, B., Hunter, I.S. & Hoskisson, P.A., 2025. Barcodes and Bioreactors: Using bar-seq to understand the genome dynamics of *S. clavuligerus* during industrial fermentations. Oral presentation at the Glasgow Microbiology Collective, Glasgow, UK, August 2025.

Croxford, J.T., Kendrew, S., Huckle, B., Hunter, I.S. & Hoskisson, P.A., 2024. Understanding the genome dynamics of *Streptomyces clavuligerus* during industrial fermentation. Oral presentation at the Glasgow Microbiology Collective, Glasgow, UK, August 2024.

Croxford, J.T., Kendrew, S., Huckle, B., Hunter, I.S. & Hoskisson, P.A., 2024. Understanding the genome dynamics of *Streptomyces clavuligerus* during industrial fermentation. Oral presentation at the IBioIC Annual Conference, Glasgow, UK, March 2024.

Croxford, J.T., Kendrew, S., Huckle, B., Hunter, I.S. & Hoskisson, P.A., 2024. Barcodes & Bioreactors: Tracking Fermentation Stability. Poster presentation at the Microbiology Society Annual Conference, Edinburgh, UK, April 2024.

Croxford, J.T., Kendrew, S., Huckle, B., Hunter, I.S. & Hoskisson, P.A., 2023. Understanding the genome dynamics of *Streptomyces clavuligerus* during industrial fermentation. Poster presentation at the Microbiology Society Annual Conference, Birmingham, UK, April 2023.

Croxford, J.T., Kendrew, S., Huckle, B., Hunter, I.S. & Hoskisson, P.A., 2023. Understanding the genome dynamics of *Streptomyces clavuligerus* during industrial fermentation. Oral presentation at the IBioIC Annual Conference, Glasgow, UK, March 2023.

Croxford, J.T., Kendrew, S., Huckle, B., Hunter, I.S. & Hoskisson, P.A., 2022. Understanding the genome dynamics of *Streptomyces clavuligerus* during industrial fermentation. Poster presentation at the Microbiology Society Annual Conference, Belfast, UK, April 2022.

Croxford, J.T., Kendrew, S., Huckle, B., Hunter, I.S. & Hoskisson, P.A., 2022. Understanding the genome dynamics of *Streptomyces clavuligerus* during industrial fermentation. Poster presentation at the Meeting of the International Society for the Biology of Actinomycetes, Toronto, Canada, June 2022.

Croxford, J.T., Kendrew, S., Huckle, B., Hunter, I.S. & Hoskisson, P.A., 2022. Understanding the genome dynamics of *Streptomyces clavuligerus* during industrial fermentation. Poster presentation at the MRC National PhD Training Program in AMR Research Annual Conference, Bristol, UK, August 2022.

Croxford, J.T., Kendrew, S., Huckle, B., Hunter, I.S. & Hoskisson, P.A., 2022. Understanding the genome dynamics of *Streptomyces clavuligerus* during industrial fermentation. Oral presentation at the Streptomyces Symposium, GSK, Glasgow, UK, June 2022.

Acknowledgements

Writing these acknowledgments feels entirely insufficient to express how grateful I am to everyone I have met during this PhD and who, in their own way, have played a part in it. But hopefully they go some way towards showing my appreciation.

I would firstly like to thank my supervisor, Professor Paul Hoskisson. Thank you for giving me the opportunity to complete this PhD, for supporting me from the application process all the way through to writing this thesis. Your obvious obsession with science and breadth of knowledge are contagious and have inspired me throughout the past four years. Most importantly, thank you for reminding me that while a PhD is about the science it produces, it is more importantly about the scientist who comes out of it. Thank you for allowing me the time and space to develop, explore new opportunities and begin to find my way in this field.

I would also like to extend my thanks to my second supervisor, Professor Iain Hunter, and to Dr Paul Herron. While Paul ensured that I had a sufficiently gruelling experience in both my 9 and 22 month assessments, I have always valued the honesty, constructive feedback and encouragement you gave, which were a great source of comfort throughout. And to Iain, thank you for your kind words when needed and for always encouraging me to believe in myself as a researcher. I will miss our chats about F1.

It goes without saying that I will always appreciate the members of Floor 6, past and present. Floor 6 has been somewhat of a home away from home, and many people there will remain friends for life. You are some of the funniest and most ridiculous people I have ever met. I admit that while I should have been working in the lab or office, many hours were instead spent debating important questions such as “are there more doors or wheels in the world?” Everyone: students, post-docs, and PIs alike, made it a privilege to

come into work every day. I will look back on my time on Floor 6 with laughter, friendship and a shared trauma.

To the Hoskie Group, especially JB, Kaisha, John Munnoch, Jack and Dan, thank you for enduring my poor jokes and distracting presence more than most. It has been a joy working with you all. I am beyond grateful for your support and advice on this project. You are outstanding microbiologists, but more importantly, wonderful people. A Hoskie group dinner is long overdue.

I would also like to thank Elmira. You are an amazing person, an inspiring scientist, and you supported me when I needed it most. Our Wednesday night carbonaras were a lifeline that few others can truly appreciate. On top of being a brilliant friend, thank you for helping with *Streptomyces* work, reading this thesis and always being a cheerleader for me. Truthfully, thank you.

To my parents, thank you for telling me how proud you are of me and for offering advice, love and support when I needed it. While it is true you had no idea what I was actually doing these past four years, and when asked would usually reply, “by the sounds of it, another failed experiment”, you have always been relentlessly supportive and proud nonetheless.

And above all, to my fiancée, Dr Robyn Braes, there are no words that can fully express my gratitude, but in true me style, that will not stop me trying. I cannot thank you enough for the unwavering support you have given me, especially at the times I needed it most. It’s funny how often you were the one reminding me that we suffer more in imagination than we do in reality. You wrote in your acknowledgments that without this PhD we would never have met, and that you’ll always be thankful I made the move. That feeling is more than reciprocated. You have made this PhD feel effortless because whatever happened,

you were always there at the end of it. However mundane, I will always treasure our evenings spent thesis writing together, you drawing schematics to show me how to clone vectors, or me doing bioinformatics for you. You are the best scientist and human I know, and your creativity, drive and perfectionism, both in and out of science, have been the greatest joy to share. I hope I have made you proud. Now let's go and get married!

Contents

Abstract

Antimicrobial resistance is a global challenge that requires the discovery of new antimicrobial compounds and a deeper understanding of the underlying production of current ones. *Streptomyces clavuligerus* is an industrially important bacterium used for the large-scale fermentation of the β-lactamase inhibitor clavulanic acid (CA). However, variability during fermentation undermines productivity, with limited knowledge of how population dynamics and evolutionary change contribute to this instability.

This thesis established and applied barcode sequencing (Bar-seq) to track strain-level population dynamics in industrially relevant *S. clavuligerus* fermentations. A library of 20 isogenic barcoded strains was constructed based on the vector pSET152. Validation experiments confirmed stable chromosomal integration and accurate quantification of relative abundance of barcoded strains. Bar-seq revealed that population structure in small-scale (50 mL) fermentations was stable over 48–72 hours, but scale-up to 1 L in rich media (TSB) introduced instability and batch-to-batch variation. In contrast, multistage fermentations in industrial media (S2A -CM5) were stable across scales up to 7.5 L, indicating that process-relevant conditions do not result in competitive divergence.

To assess the evolution of *S. clavuligerus* in fermentations, barcoded populations were passaged for 70 days (152 generations) in CM5. Replicate cultures differed in population dynamics, with distinct strains becoming dominant, while others declined or persisted at lower relative abundance. Whole-genome sequencing and transcriptomics linked these shifts to specific mutations and altered gene expression, though declines in CA titres were observed across all populations. Phenotypic assays showed variable colony morphology and sporulation, but these traits were unreliable indicators of any underlying population change.

Overall, this work demonstrates that population stability depends strongly on media and scale. Long-term adaptation can reduce CA production despite stable short-term population dynamics. Bar-seq provides a powerful tool for linking genotype, phenotype, and productivity in *Streptomyces*, with direct implications for industrial strain development and process monitoring.

Chapter 1 Introduction

1.1 Anti-microbial resistance (AMR)

1.1.1 Global burden of AMR

The discovery of penicillin by Sir Alexander Fleming in 1928 (Fleming, 1929) started the hunt for new antibiotic compounds which led to the 'golden age' of antimicrobial discovery. In the 1940-1960s, around 25 new antimicrobial compounds were discovered (Lewis, 2013). Resistance to salvarsan was the first recorded occurrence of AMR within the clinic (Silberstein, 1924; Stekel, 2018) which marked the beginning of awareness towards the issue. We now know bacteria, parasites, fungi and viruses have evolved to become resistant to antimicrobial compounds to an extent where AMR has become one of the most serious global burdens of the modern day. The AMR endemic is thought to be driven by the overuse of antimicrobials within clinical and agricultural settings (Llor and Bjerrum, 2014). Although it is difficult to directly contribute fatalities to AMR, in 2019 it was estimated that 4.95 million deaths globally could be directly attributed to bacterial AMR alone (Murray et al., 2022).

The 'silent pandemic' of AMR is only worsening and it is estimated that by 2050, it could become the leading cause of death worldwide due to the once reliable antimicrobial compounds being ineffective towards pathogens (O'Neill, 2016). In addition, the COVID-19 pandemic has hampered global efforts to tackle antimicrobial resistance. For example, 72% of patients hospitalised by COVID-19 received antibiotics during treatment whereas only 8% of those patients presented with bacterial or fungal co-infections (Rawson et al., 2020). Additionally, in England, dental antibiotic prescribing rose by 22% in the first year of the pandemic (Thompson et al., 2022).

There have been several efforts to tackle AMR primarily led by the United Nations. A number of initiatives including the "One Health Approach", the Global Action Plan for

managing AMR (GAP-AMR) and the Global Antimicrobial Resistance and Use Surveillance System (GLASS) all aim to tackle to burden of AMR (World Health Organisation, 2021). However, initiatives alone will not reduce the burden of AMR. This is highlighted by the lack of new antimicrobial compounds coming to market and the apparent apathy towards the issue by 'big pharma' (Prestinaci et al., 2015). For example, only eight new classes of antimicrobial compounds have been approved for use since the 1960's (Saloni, 2024) and of the 32 antibiotics currently in development, only four are designed for use against a World Health Organization (WHO) 'critical pathogen' (World Health Organisation, 2024).

Scientific, economic, and regulatory barriers have all contributed to this stagnation, despite rising global resistance (World Health Organisation, 2024). To combat AMR, we must accelerate the discovery of new antimicrobials and improve our understanding of current compounds, their mechanisms, and production.

1.1.2 Mechanisms of AMR

AMR arises when a microorganism survives exposure to an antimicrobial compound. It is microbes' rapid reproduction rate and large population sizes that allow them to develop AMR so quickly. Two mechanisms exist that facilitate the development of AMR, namely, spontaneous mutations and horizontal gene transfer (HGT). While external factors have an impact on the rate of spontaneous mutation within bacteria, that rate is around 10^{-8} mutations per base pair per generation (Martinez and Baquero, 2000; Torii et al., 2003) and may of course occur in genes contributing to AMR. Three main mechanism exists relating to HGT, those being, transformation, transduction and conjugation (Munita and Arias, 2016). Of the 12 bacterial groups published by the WHO in 2017 being listed as 'priority pathogens', 11 groups either perform or are predicted to have natural

competence for the horizontal gene transfer of antimicrobial genes (Lerminiaux and Cameron, 2018), underscoring how important HGT is for AMR acquisition.

Whether through spontaneous mutations or HGT, there are several possible mechanisms leading to AMR. Firstly, the antimicrobial compound itself may be altered. For example, through the production of β lactamases (Abraham and Chain, 1940; Baran et al., 2023). Another mechanism of AMR is through either the efflux of antimicrobials from the cell or the cessation of the compounds entering the cell in the first instance (Lorusso et al., 2022; Pagès et al., 2008). In these cases, the antimicrobial compound does not make it to the target site within the cell and therefore is rendered ineffective. A third mechanism for AMR is the alteration of the molecular structure of the target site for an antimicrobial compound. Here, while the cell's function is maintained, the binding mechanism of the antimicrobial is stopped (Munita and Arias, 2016; Weisblum, 1995). The last mechanism by which AMR may be acquired is through bacteria bypassing the affected pathway altogether. This is done through holistic cell adaptions allowing alternative enzymes and pathways to perform the affected function, this redundancy allows the cell to function even when exposed to the antimicrobial compound (Munita and Arias, 2016; Ramamurthy et al., 2022).

1.1.3 Searching for new antimicrobials and understanding current ones

Although fewer new antimicrobials are reaching clinical approval, research continues to discover novel compounds. Most of our current antibiotics are natural products isolated from Actinomycete species (van Bergeijk et al., 2020) largely isolated from terrestrial environments due to the ease of sampling (Bizuye et al., 2013; Crits-Christoph et al., 2018). As a result, it is thought that we have exhausted terrestrial environments for their natural products, and the search has gone further afield. Given the fact that many of the

natural products we use today are isolated from extremophiles such as Actinomycete species found in the Antarctic and Atacama deserts (Bull et al., 2016; Waschulin et al., 2022), it is thought looking in other extreme and non-terrestrial environments may yield further, novel natural products. For example, marine, swamp, and insect-microbe symbioses environments are currently being explored in the hope that they will yield new species of Actinomycete that produce novel antimicrobial compounds (Batey et al., 2020; Duncan et al., 2015; Zhang et al., 2023).

As well searching for novel antimicrobial compounds, it is also crucial to deepen our understanding of how existing antimicrobials function and are produced. Broadly, the improved knowledge of current antimicrobials can be split into two groups, upstream and downstream knowledge.

Upstream knowledge refers to the species that produce antimicrobials and the antimicrobials themselves. This includes the elucidation of biosynthetic gene clusters (BGCs), where genome mining has discovered many BGCs that should theoretically produce antimicrobial compounds that are not expressed under laboratory conditions (Rutledge and Challis, 2015; van Bergeijk et al., 2020). Our knowledge of BGCs can be split into four groups, known knowns, where both the cluster and its product are characterised, known unknowns, where a cluster is identifiable but its product is not, unknown knowns, where a natural product is known but the responsible cluster has not been linked, and unknown unknowns, which represent clusters and products beyond the reach of current detection methods and require new discovery approaches (Hoskisson and Seipke, 2020). Efforts are therefore directed at activating cryptic BGCs and optimising production through genetic and environmental manipulation. This

includes using co-culturing, chemical elicitors, and alteration of global regulation to induce expression (Gao et al., 2012; Moore et al., 2012; Rateb et al., 2013).

Downstream improvement of current antimicrobial compounds focuses on production processes, specifically relating to fermentation and purification. Better understanding how fermentation parameters, such as carbon source, pH, oxygen and nutrient limitation is important for achieving high yields at industrial scale (Harms et al., 2017). Aspects such as how batch or continuous culture affects antimicrobial production have also been investigated (Hoskisson and Hobbs, 2005). Strain engineering is widely used to boost antimicrobial production (Abbate et al., 2023). For example strain engineering has allowed antimicrobials to be produced by bacteria in a 'high synthesis-low growth state' through the channelling of metabolic flux away from key cellular processes to secondary metabolite production (Mannan et al., 2025). One such group of bacteria that are important for the production of secondary metabolites and are a focus of strain engineering are the Actinomycetota and *Streptomyces*.

1.2 The Actinomycetota and *Streptomyces*

1.2.1 Introduction to Actinomycetota and *Streptomyces*

Actinomycetota are a phylum of genetically diverse bacteria that have been isolated from terrestrial, marine and other, more extreme environments. They are characterised as being Gram-positive and have an unusually high G+C content, ranging from 60-70% (Barka et al., 2015; Ventura et al., 2007). Actinomycetota superficially resemble fungi due to filamentous morphology, distinctive hyphal growth leading to sporulation and the formation of aerial hyphae (Ait et al., 2015). Actinomycetota are also characterised by their ability to be both chemoautotrophic and heterotrophic with an ability to metabolise

complex polysaccharides (Flint et al., 2012; Zimmermann, 1990). Actinomycetota are also saprophytic, aerobic and morphologically diverse. The unusual morphologies of Actinomycetota and unique genetic characteristics, such as the high G+C content, lends the phylum to a wide range of biotechnological uses (Gohain et al., 2020).

The Actinomycetota phylum is one of the most populated phyla of all, containing the largest bacterial family, the Streptomycetaceae, which was first isolated in 1943 and is also the home to the genus *Streptomyces* (Waksman and Henrici, 1943). The isolation of the first *Streptomyces* species, *Streptomyces griseus* was quickly followed by the discovery that the species could produce streptomycin (Schatz et al., 1944), securing the organism's status as a significant secondary metabolite producer. Selman Waksman contributed much to the field of *Streptomyces* research while still in its early stages during the 1920s and 1930s (Waksman, 1958). David Hopwood was also instrumental in the field of *Streptomyces* research and made the discovery that *Streptomyces* was not a fungi when he discovered DNA existed freely within the cytoplasm, and not within a nuclear envelope, using electron microscopy (Glauert and Hopwood, 1959).

Streptomyces are best known for their industrial uses due to their prolific production of secondary metabolites. The genus is responsible for more than 80% of bacterial bioactive natural products, far outstripping other genera of bacteria. It also has other industrial uses aside from the production of important antibiotics such as bioenergy, herbicidal bioremediation and biocatalysts (Gomez-Escribano et al., 2021; Spasic et al., 2018), and the production of 50-55% of all antibiotics (Maiti et al., 2020). The clinical uses of *Streptomyces* secondary metabolites alone covers antimicrobials, immunosuppressives, anti-cancer agents, antifungal, anti-virulent, immunosuppressive, and anthelmintics (Chevrette et al., 2020; Gomez-Escribano et al.,

2021; Hoskisson and Seipke, 2020). It is this wide range of biotechnological and clinical uses that make *Streptomyces* so important.

1.2.2 Taxonomy, life cycle and morphology

*Streptomyce*s are phylogenetically and morphologically diverse. Numerous evolutionary events across all environments have led to great variation between species, and as such, makes it difficult to phenotypically identify species (Labeda et al., 2012). Additionally, the complex nature of *Streptomyces* genomes and the high levels of horizontal gene transfer have also made it difficult for taxonomical classification of species (Kiepas et al., 2023). The difficulty in morphological and taxonomic classification of the *Streptomyces* is not helped by the fact that the genus has the greatest number of described species across all bacterial genera.

Despite the difficulty in classification, all *Streptomyces* possess an unusual and complex lifecycle that enables them to survive and disperse in a wide range of environments. This survival and dispersion in otherwise hostile environment are largely due to the formation of spores (Ensign, 1978; Mayfield et al., 1972), under the control of a complex cascade of global secondary messengers (Bibb, 2005).

Streptomyces exist for large periods of time as spores, especially under nutrient-starved conditions (Williams et al., 1972). The transition of *Streptomyces* from vegetative growth to reproductive growth, as well as the production of secondary metabolites, are tightly controlled by nutrient limitation, facilitated through several nucleotide secondary messengers.

The *Streptomyces* life cycle begins with the germination of a dormant spore, triggered by environmental cues indicating nutrient availability. This involves swelling, polarisation, and the emergence of one or more germ tubes. These tubes grow by polar extension at

the hyphal tips, a process driven by the polarisome complex, which includes FilP, DivIVA, and Scy (Hempel et al., 2008; Holmes et al., 2013; Katsuya et al., 2013). New branches also emerge laterally along existing hyphae. The resulting mycelium forms a dense, filamentous network that penetrates the substrate, resembling root systems in soil (Jones and Elliot, 2018). Hyphal branches extend at approximately 20 μm/ hour (Jyothikumar et al., 2008). Under specific conditions, such as co-culture with fungi or nutrient stress, *Streptomyces* can undergo exploratory growth, a rapid, non-branching growth mode approximately 10 times faster than vegetative hyphal extension (Jones et al., 2017). This behaviour likely allows competitive colonisation of new niches.

As nutrients become limited, development transitions to aerial hyphae formation. This is enabled by proteins including chaplins and SapB, which lower surface tension and permit hyphae to emerge from the colony surface (Sigle et al., 2015). Chromosome segregation follows, coordinated by the ParAB system, ensuring that each spore compartment contains a complete genome. The resulting spore chains eventually mature into thick-walled, dormant spores (Figure 1.1).

This complex developmental process is tightly controlled by two major genetic systems, the *bld* (bald) and *whi* (white) genes. Mutations in *bld* genes block aerial hyphae formation, while *whi* mutations impair spore maturation (Chater, 2001; McCormick and Flärdh, 2012). *bldA* encodes a rare tRNA for the UUA codon, found in only ~2% of genes in *Streptomyces coelicolor*, and is required for both aerial growth and antibiotic production (Chater and Chandra, 2008; Hackl and Bechthold, 2015; Lawlor et al., 1987). Another key regulator, BldD, functions as a DNA binding repressor with homology to the λ repressor, and controls multiple developmental and metabolic genes (Den Hengst et al., 2010; Kim et al., 2006).

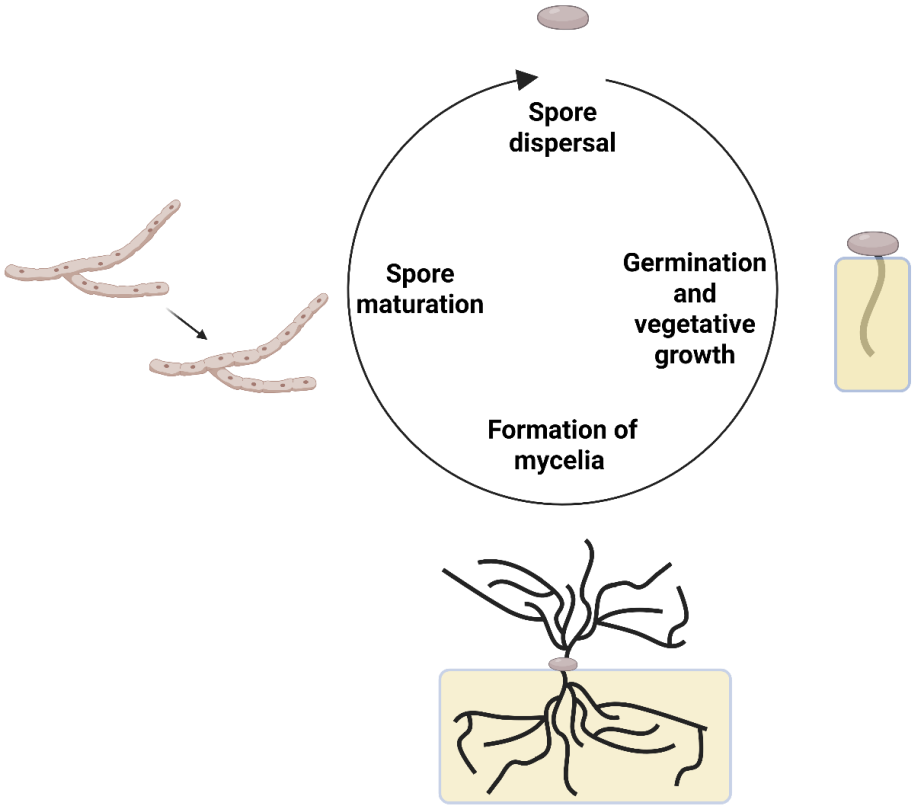


Figure 1.1. The life cycle of *Streptomyces*. A spore germinates to form branching vegetative hyphae (substrate mycelium), which grow across the surface. In response to nutrient limitation, aerial hyphae are produced that differentiate into chains of spores, completing the cycle. Created using BioRender.

Global regulation is mediated through several signalling systems. For example, phosphate limitation is sensed via the PhoR–PhoP two-component system, which activates specialised metabolite biosynthesis (Rodríguez-García et al., 2007). Regulatory overlap between development and metabolism is common. For instance, BldD indirectly controls secondary metabolism via *bldA*, while MtrAB, a two-component system, regulates both sporulation and the production of antibiotics such as chloramphenicol in *S. venezuelae* (Hoskisson and Fernández-Martínez, 2018; Som et al., 2017).

1.2.3 Life cycle within liquid culture

The life cycle of *Streptomyces* within liquid culture differs from that of *Streptomyces* in solid conditions. *Streptomyces* species grow as mycelial aggregates in liquid culture which is known to affect production of specialised metabolites (Nieminen et al., 2013a; van Dissel et al., 2014). Submerged cultures of *Streptomyces* usually exhibit three types of growth (Paul and Thomas, 1998), namely, freely dispersed mycelia, open mycelial networks, and pellets with a nutrient-deprived centre.

Factors such as agitation, aeration, temperature and batch size have also been investigated as to the efficiency of *Streptomyces* fermentation within liquid culture (Viana Marques et al., 2018). Hyphal fragmentation is common and serves as the primary means of dispersal in liquid media, replacing the sporulation-dependent dispersal of terrestrial environments. The shift from a multicellular, surface-based life cycle to a submerged, pelletised form represents a fundamental change in morphology, with implications for nutrient uptake, oxygen diffusion, and physiological heterogeneity across the culture. This fundamental change is an important consideration for antibiotic production via fermentation and must be accounted for when optimising processes.

1.3 *Streptomyces clavuligerus*

1.3.1 Introduction

S. clavuligerus is an industrially relevant streptomycete best known for its production of clavulanic acid (CA), a clinically used β-lactamase inhibitor. It was first isolated from a South American soil sample in the early 1970s and originally described for its production of two novel cephalosporin antibiotics (Higgens and Kastner, 1971). The species name, 'clavuligerus', comes from the Latin clavula, meaning little club, referring to the club-shaped appearance of its hyphae under the microscope (Reading and Cole, 1977). Like most streptomycetes, *S. clavuligerus* is a filamentous, sporulating, Gram-positive bacterium with high G+C content and a distinct fuzzy colony morphology. It is typically isolated from soil, in line with the terrestrial distribution of most *Streptomyces* species (Chater, 2016).

Unlike model species such as *S. coelicolor*, comparatively little work has been done on the core genetics of *S. clavuligerus* outside of its role in secondary metabolism. The genome consists of a linear chromosome of approximately 6.8 Mb and four linear plasmids, namely pSCL1-4 (Figure 1.2). Originally three of the four plasmids were isolated using pulse field gel electrophoresis (Netolitzky et al., 1995). In 2010 a scaffolded Sanger-based assembly with a 6.76 Mb chromosome and a single 1.8 Mb mega plasmid, was published (Medema et al., 2010). However, a different group also published a draft genome of *S. clavuligerus* in the same year, and reported four linear plasmids (Song et al., 2010b). A higher-quality assembly in 2019 resolved the

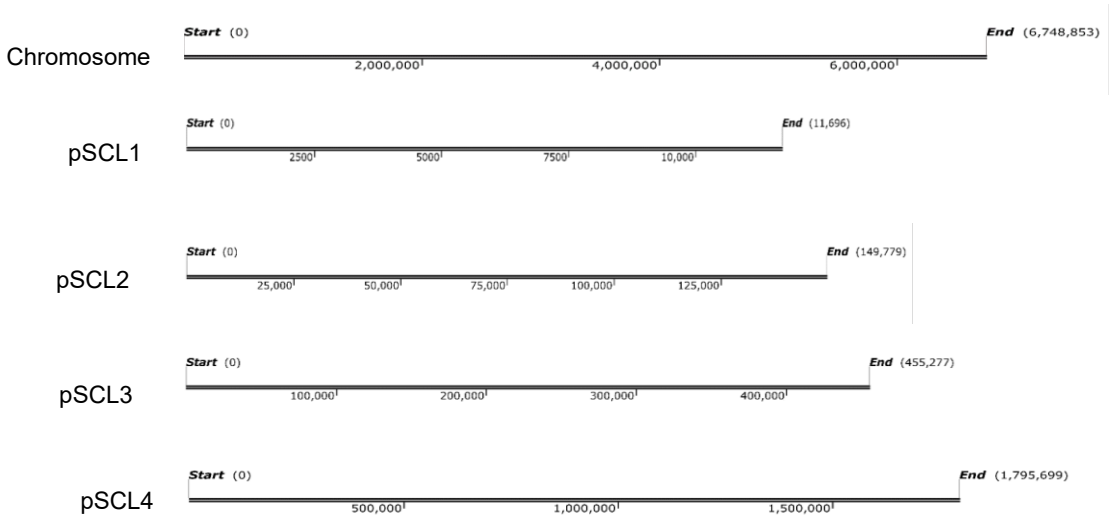


Figure 1.2. Genome arrangement of *S. clavuligerus DSM738*. *S. clavuligerus* has a 6.75 Mb chromosome alongside 4 plasmids termed pSCL1-4 ranging from 11 Kb to 1,796 Kb in size (not to scale).

chromosome and pSCL4 using PacBio and Illumina sequencing but again failed to detect pSCL1–3 (Hwang et al., 2019). More recently, a complete genome of a long-maintained ATCC 27064 isolate from the University of León was published (Gomez-Escribano et al., 2021), addressing some of these inconsistencies. Across these studies, plasmid detection has been variable, reflecting ongoing technical and biological issues with replicon stability. All of these factors contributing to the various assembly types and genomic rearrangements of the *S. clavuligerus* genome can broadly be termed genome dynamics or genome plasticity (Ríos-Fernández et al., 2024).

As of June 2025, 31 assemblies of *S. clavuligerus* are listed in the NCBI genome database. These vary in total size and plasmid content. Each assembly contains between one and four plasmids. The smallest complete assembly is 7.60 Mb (GCA_015767895.1), while the largest is 9.16 Mb (GCA_028752555.1). Most complete genomes include a chromosome of ~6.7–6.9 Mb, and in assemblies with four plasmids, the largest plasmid (pSCL4) is consistently around 1.8 Mb. These differences arise from the challenges associated with sequencing large linear chromosomes, which are complicated by the complex secondary structures formed at the telomeres of the replicons.

The arrangement of genes within the *S. clavuligerus* genome also play an important role within its genetic makeup. Core genes of *S. clavuligerus* are clustered towards the centre of the linear chromosome, with accessory genes including most BGCs, being found towards the ends of the chromosome (arms; Bentley et al., 2002; Medema et al., 2010). The central region is more likely to be stable and conserved, while the arms are more prone to deletions, rearrangements, and horizontal gene transfer (Choulet et al., 2006; Cruz-Morales et al., 2013). This is also true for other Actinomycetota such as

Micromonospora (Mark et al., 2024). BGCs on the arms are therefore less stable and more likely to be lost during strain manipulation (Romero-Rodríguez et al., 2016). This structure reflects a trade-off between maintaining essential functions and allowing flexibility in specialised metabolism.

1.3.2 Genotypic heterogeneity

The largest plasmid, pSCL4, has been shown to play a key role in maintaining chromosome integrity. Targeted deletion of pSCL4 causes loss of terminal regions and circularisation of the chromosome, consistent with a structural role in preserving chromosome linearity (Gomez-Escribano et al., 2021). In some cases, loss of pSCL4 has also been associated with the circularisation and significant genetic instability of the chromosome (Mohit, 2023). Although loss of pSCL4 can cause genetic instability, other studies show it does not affect viability, highlighting the inconsistent genetics of *S. clavuligerus* (Álvarez-Álvarez et al., 2017).

In contrast, pSCL1–3 appear dispensable for chromosomal structure and stability, deletion mutants show no major rearrangements or growth defects (Mohit, 2023). Their size and low gene content suggest they serve more accessory roles. A summary of the structure and stability impact of each plasmid is shown in Table 1.

1.3.3 Phenotypic heterogeneity

While little work has been done detailing the phenotypic heterogeneity of specifically *S. clavuligerus*, the discussed genetic plasticity can lead to phenotypic differences. Studies relating to the phenotypic heterogeneity are often contextualised to fermentation conditions due to the industrial use of the species (Lemoine et al., 2017). This also mean they often relate to submerged cultures where the life cycle differs to that on solid agars. Morphologically, smaller or fragmented mycelial forms are associated

with increased CA production, likely due to improved oxygen and nutrient diffusion (Belmar-Beiny and Thomas, 1991; Gómez-Ríos et al., 2024, 2022). In contrast, larger pellets often suffer from internal mass transfer limitations and reduced metabolic activity. Agitation plays a key role in heterogeneity, moderate shear reduces pellet size and enhances CA production, but excessive shear damages cells. Many other fermentation parameters such as pH, carbon source and nitrogen source and also affect phenotypic aspects of *S. clavuligerus* and therefore CA production (Viana Marques et al., 2018).

Table 1. Architecture and characteristics of *S. clavuligerus* DSM738 genome. Adapted from Mohit, 2023.

	Chromosome	**pSCL4**	**pSCL3**	**pSCL2**	**pSCL1**
Sequence length	6.8 Mb	1.8 Mb	455.3 kb	149.8 kb	11.7 kb
G+C%	72%	72%	72%	72%	72%
Coding sequences	5,700	1,581	528	175	12
Secondary metabolite BGCs	23	17	1	0	0

Growth and development of *Streptomyces* are similarly variable across a population (Hoskisson et al., 2024). Secondary metabolite production is accompanied by a transcriptomic change in both solid and liquid cultures (Yagüe et al., 2014). This transition is however not consistent across a cultures, while some hyphae remain vegetative, others enter the stage of their life cycle where they are producing secondary metabolites, and a subset undergoes programmed cell death (Manteca and Yagüe, 2018; Yagüe et al., 2012). The result is a mixed population at different physiological stages, even under identical conditions.

Temporal differences have also been seen in expression of key genes within CA biosynthesis. Levels of CcaR, a key regulatory gene in both CA and cephamycin C biosynthesis vary greatly between exponential and stationary growth within submerged culture (Kyung et al., 2001). Additionally, studies of *S. clavuligerus* on solid agar also support functional differentiation. For example, in the same study, CcaR was expressed in the substrate mycelium but not in aerial hyphae, indicating a developmental division even within single colonies (Kyung et al., 2001).

Differences in growth form, gene expression, and metabolite production exist both across and within colonies, and even under supposedly identical conditions. This is compounded by the fundamental shift in life cycle between solid and liquid cultures. On solid agar, sporulation allows for dispersal whereas in liquid, dispersal relies on fragmentation. These differences are not simply superficial, they affect regulation, development, and metabolic output. As a result, generalising findings from one growth condition to another is unreliable. Any study of *S. clavuligerus* must account for this.

1.4 Secondary Metabolism in *Streptomyces clavuligerus*

1.4.1 *Streptomyces* are a potent producer of secondary metabolites

S. clavuligerus, like all *Streptomyces*, is a prolific producer of secondary metabolites. *Streptomyces* are thought to produce around 100,000 antibiotic compounds (Alam et al., 2022) with each *Streptomyces* species having between 25 and 70 secondary metabolite BGCs.

The levels of secondary metabolites produced by *Streptomyces* are often related to environmental stress. Indeed the primary and secondary metabolism within *Streptomyces* has been shown to be closely linked (Schniete et al., 2018). The process by which genes responsible for sporulation are linked to those responsible for the regulation and biosynthesis of secondary metabolites is termed as pleiotropic switching (Barona-Gómez et al., 2023; Botas et al., 2018; Nieselt et al., 2010).

One such example of the close link between the regulation of primary and secondary metabolism is the key secondary messenger cyclic di-GMP (c-di-GMP). The primary role of c-di-GMP is the through the binding to the key transcriptional regulator BldD, enabling the dimerization of BldD and thus repression of genes responsible for aerial growth, leading to vegetative growth (Latoscha et al., 2020, 2019; Tschowri et al., 2014). This repression of sporulation allows for metabolic flux to be channelled towards biomass accumulation and secondary metabolism. The regulation of growth is further controlled through c-di-GMP through its binding to the sigma factor σ^{WhiG}. This binding in turn controls two transcription factors responsible for spore formation, showing the c-di-GMP has a global role in the life cycle and primary metabolism of *Streptomyces* (Gallagher et al., 2020). Further to this, c-di-GMP binds to GlgX to promote the degradation of glycogen stores within the cell during periods of sporulation

(Schumacher et al., 2022). The levels of c-di-GMP within the cell have been observed to fluctuate dynamically, being high during vegetative growth, low during developmental initiation, and high again during sporulation (Gallagher et al., 2024), highlighting the molecules role within the development of *Streptomyces* life cycle.

Additionally, studies have been performed in *S. coelicolor* where manipulation of diguanylate cyclase enzymes, responsible for the regulation of c-di-GMP itself alters the production of secondary metabolites actinorhodin and undecylprodigiosin alongside changes in morphological development (Al-Bassam et al., 2018). Broadly it can be concluded that increased levels of c-di-GMP support enhanced antibiotic production in *Streptomyces* due to the move of metabolic flux away from sporulation towards secondary metabolism (Latta and Bechthold, 2022). As a result, regulation of c-di-GMP levels within the cell would be an appropriate target for the manipulation of industrial stains of *Streptomyces* to increase the production of their antibiotic products.

1.4.2 Secondary metabolism within *S. clavuligerus*

S. clavuligerus has 58 secondary metabolite BGCs (Hwang et al., 2021), but only around a third of the proteome (2,442) of *S. clavuligerus* has been directly studied and reported, showing much research is still be completed on the species (Ferguson et al., 2016).

However, *S. clavuligerus* is well known for its important natural products clavulanic acid (CA) and cephamycin C, both of which have tightly linked regulation. Cephamycin C is a broad-spectrum β-lactam antibiotic whereas CA is a β-lactamase inhibitor, making it clear why the regulation and metabolism of these two compounds are tightly linked. CA is a secondary metabolite product produced by and named after *S. clavuligerus* (Figure 1.3). CA is a four membered β-lactam ring that is administered clinically alongside the β-lactam antibiotic amoxicillin under the brand name Augmentin™, more commonly known

as co-amoxiclav, which itself is on the World Health Organisations list of essential medicines (World Health Organisation, 2021). The clinical importance of CA lies in its ability to counteract β-lactamase mediated beta-lactam resistance.

β-lactamases hydrolyse the β-lactam ring of penicillins, cephalosporins, monobactams, and carbapenems. CA inhibits β-lactamases by acting as a suicide inhibitor. It binds to the enzyme's active site and undergoes a reaction that irreversibly inactivates the enzyme (Knowles, 1985). This inhibitory activity relies on the 3R,5R stereochemistry of its strained bicyclic nucleus.

Similar to the mode of action itself, the method of biosynthesis of the molecule is also documented (Liras and Martín, 2021; Paradkar, 2013; Saudagar et al., 2008). The molecule is produced from the precursor molecules glyceraldehyde-3-phosphate and arginine. The biosynthesis of CA is often split into late and early stages with corresponding genes. The early stage sees glyceraldehyde-3-phosphate and arginine converted to dihydroclavaminic acid while the late-stage reactions see this dihydroclavaminic acid converted to CA (Figure 1.4).

As mentioned, the genes responsible for production of CA can also be split into late and early genes. The early genes within the biosynthetic gene cluster have been entirely duplicated within the genome (Jensen et al., 2000) and located on the megaplasmid pSCL4, however there is a full set of genes responsible for the production of CA on the chromosome.

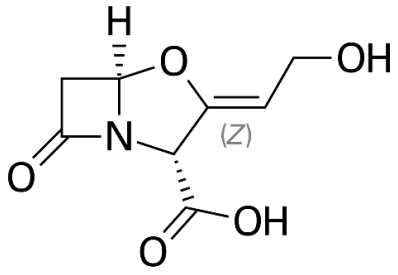


Figure 1.3. Clavulanic acid. The structure of clavulanic acid in the Z stereoisomer form.

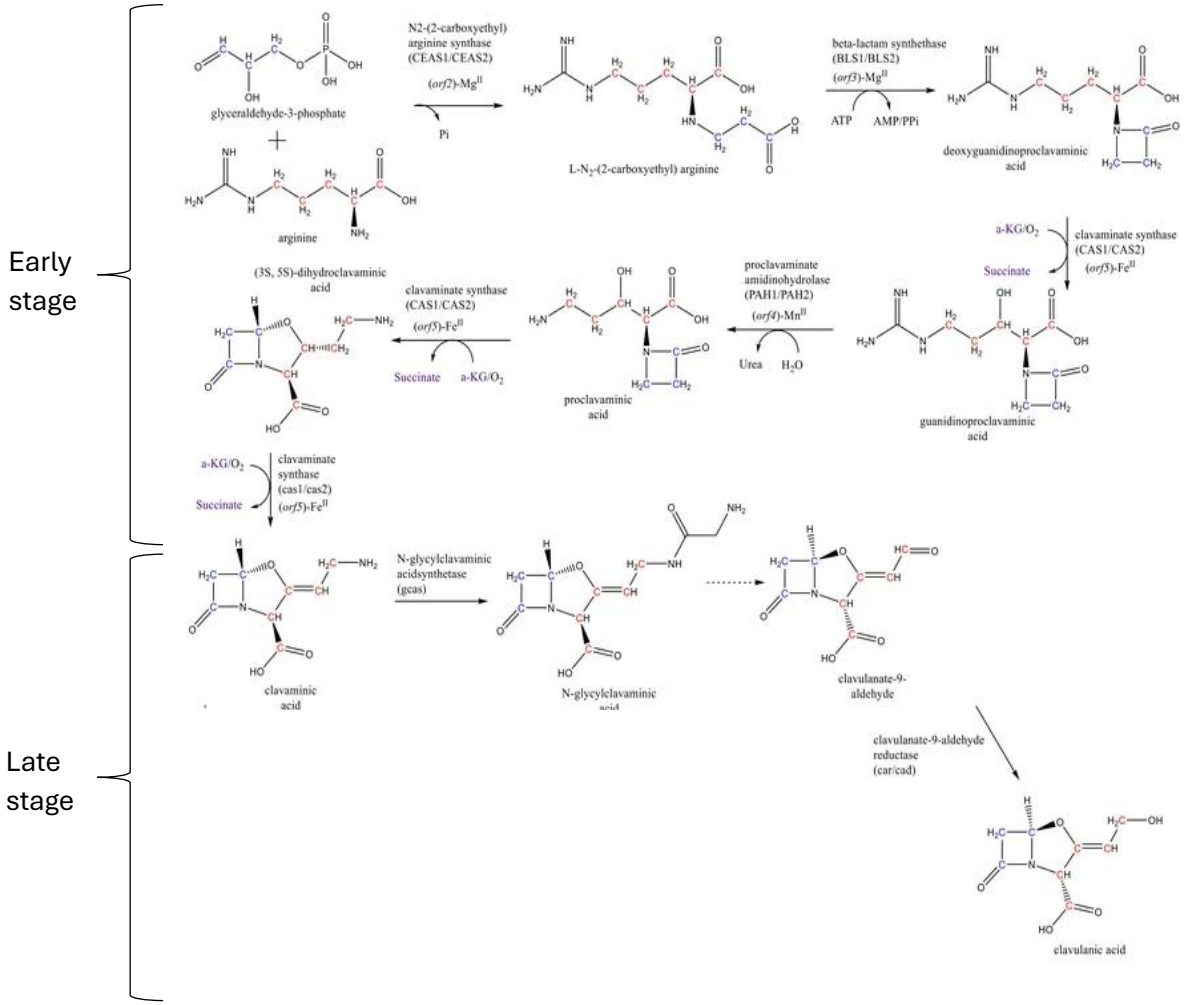


Figure 1.4. The biosynthesis of clavulanic acid. The precursor molecules of glyceraldehyde-3-phosphate and arginine go through both late and early-stage reactions for biosynthesis into clavulanic acid. Adapted from Ramirez-Malule et al., 2018.

The transcription of the CA BGC is subject to regulation by two key regulatory genes, CcaR and ClaR (Paradkar, 2013). ClaR is a LysR type transcriptional activator which positively regulates the late stages of CA biosynthesis. Loss of any late-stage genes from the BGC can cause a complete loss of CA production, while a knockout of the genes responsible for the early stages of biosynthesis results in only a conditional loss of the production of the molecule due to the paralogues in the duplicated gene set (Paradkar, 2013). A recent comparative genomics study was undertaken to elucidate the 'core genes' required for the production of CA (AbuSara et al., 2019), and found that the clavam and paralogue gene clusters are not required for the production of CA and reaffirmed the duplication of the *ceaS1, bls1*, and *pah1* early stage genes within pSCL4. Furthermore, β-lactam superclusters are present in CA producing strains such as *S. clavuligerus* and *S. jumonjinensis*, which are absent in non-CA producing strains (AbuSara et al., 2019). While progress has been made in understanding the production of CA, a complete understanding of the biosynthetic gene cluster is still some way off.

As mentioned, regulation of CA biosynthesis is initiated by the SARP family regulatory protein CcaR, located in the adjacent cephamycin C BGC (Santamarta et al., 2011; Wietzorrek and Bibb, 1997). One target of CcaR is the late-stage biosynthetic regulator, ClaR, another member of the SARP family of proteins. Knockout strains of CcaR are unable to produce both CA and cephamycin C suggesting it is essential for the regulation of both molecules (Paradkar, 2013). While CcaR is considered a cluster situated regulator, it is also suggest to be regulated by a number of global elements ubiquitous in *Streptomyces* (Bibb, 2005). CcaR is indirectly regulated by environmental stresses, nutrient availability, and growth rate (Kyung et al., 2001). It is also a target of the developmental regulator bldG, an anti-anti-sigma factor. Knockout of *bldG* abolishes the production of both CA and cephamycin C by eliminating CcaR expression (Bignell et al.,

2005). As a result of this regulation and due to the clinical relevance of CA, the BGC has been the subject of metabolic engineering efforts, including heterologous expression and regulator overexpression to enhance yields (Jnawali et al., 2010).

1.5 Industrial production of antimicrobials

1.5.1 Overview of fermentation

One key technology responsible for the production of antimicrobial compounds at an industrial level is fermentations. While humans have been using fermentation technology for over 11,000 years it is only within the past century that we have been able to truly harness production at the scale of 100,000s of litres (Perkins, 2025).

Industrial-scale antimicrobial fermentations occur at ranges from 10,000 L to over 150,000 L (Li et al., 2015). These processes use submerged liquid fermentation for bacteria and yeast, and solid-state fermentation for filamentous fungi (Green et al., 2024; Gu et al., 2024). Nutrient media typically include a carbon source (e.g., glucose), nitrogen (e.g., plant protein extracts), trace elements, and sometimes inducers. All of which are optimized for cost, efficiency and yield. Fermentation types include batch, fed-batch, continuous, and solid-state systems, each offering distinct control over nutrient supply and product formation (Hoskisson and Hobbs, 2005).

1.5.2 Multistage fermentations

Often, industrial fermentations are done over multiple stages, each promoting different parts of the cell cycle (Deindoerfer and Humphrey, 1959). These fermentations are termed multistage fermentations and are split into a first stage(s) (seed) and second stage (production). The seed stage is used to promote the growth of biomass, which is

then used as an inoculum for the production stage, which promotes the production of the desired secondary metabolite. As the seed stage is for the promotion of rapid biomass accumulation it is often run for no longer than 48 hours. The production stage of the fermentation is run for as long as the bacteria produces the desired product at an appropriate level before entering the death phase, often around 5 days (Taurino et al., 2011).

The seed and production stage are distinct because the conditions promoting each are different, such as a change in pH or temperature. However, the most common difference between the two stages are the media (Wentzel et al., 2012). Especially considering *Streptomyces* different stages of growth, having conditions that promote primary and secondary metabolism is important and an area of ongoing research.

1.5.3 Scale up challenges

Due to the industrial nature of fermentations much of the research conducted on them is proprietary, therefore, despite their importance there is relatively little published research on fermentation conditions. Within published literature, production is tightly regulated via real-time monitoring of pH, temperature, dissolved oxygen, agitation, and foam (Xia et al., 2021). Unlike lab-scale setups, which allow precise control in small flasks, scaling up introduces challenges such as poor oxygen transfer and mixing inefficiencies. Engineers apply scale-up models (e.g., oxygen transfer rate matching) and simulation tools to bridge the gap between lab and industrial conditions (Rastädter et al., 2023). Media composition must balance cost with functionality, and processes must accommodate sterilization, anti-foaming, and downstream recovery to ensure commercial viability. These constraints distinguish industrial fermentations from flexible, small-scale lab experiments (Kumar et al., 2021).

Aspects such as mixing, mass transfer, heat transfer and product concentration all alter the growth dynamics of a strain within a fermentation process (Reisman, 1993). Most fermentation vessels, commonly used in industrial processes and receive substrate from the top while being aerated from the bottom. This configuration typically shows a negative correlation between aeration and mass transfer and the height of the vessel. This is due to longer distances needing to be covered by the air and substates (Schmidt, 2005).

This dependency on mixing also means that aspects such as impeller design greatly affect the nutrient homogeneity throughout the fermentation vessel which allows for a significant variable from lab to industrial scale fermentation (Xia et al., 2021). To this end a key aspect of scale up is oxygen transfer rates within a vessel, especially in industrial medias that are often more viscous than lab-based medias. Vessels are continuously supplied with oxygen, however with an increasingly large fermentation vessel, maintaining a consistent oxygen transfer rate for all cells within the culture becomes more difficult (Garcia-Ochoa and Gomez, 2009).

As discussed, *Streptomyces* are filamentous organisms within submerged culture fermentations. This is specifically significant considering shear forces play a role within fermentation scale up and specifically on the morphology of filamentous organisms. While analysis of shear forces rely on complex fluid dynamics to assess the affects, attempts have been made to do so (Li et al., 2020). It has been suggested that while shear forces do play a role in fermentation productivity, the amount found within an industrial bioreactor is well below the limit of what cells can tolerate before becoming unviable (Pawar, 2018).

The study of the effects of shear force on a fermentation productivity is not aided by the fact that many organisms react differently to an increased level of shear force when scaling up to industrial levels. This is due to the different morphologies under which different bacterial species best produce natural products. For example *S. avermitilis* and *Aspergillus niger* both best produce avermectin and citrate, respectively, under pelleted conditions (Gómez et al., 1988; Yin et al., 2008). These would result from lower shear stress. However, *Penicillium chrysogenum* best produces penicillin as free mycelium (Veiter et al., 2018), experienced under higher shear stress conditions. Importantly, previous work has demonstrated that CA production is enhanced when *S. clavuligerus* grows as small, dispersed pellets, rather than as large aggregates or free mycelia, due to improved mass transfer and metabolic activity (Ives and Bushell, 1997). These examples highlight the difficulties in having a universal scale up of secondary metabolite production from lab scale to industrial, with case-by-case scale up solutions needing to be found for various growth dynamics of different species.

Additional shifts in metabolism have been observed with scale up of fermentation. For example, it is the additional stresses of nutrient deprivation mentioned above that may causes changes to the stress-responses within cells. This in turn causing changes to the glycolysis, citrate cycle and amino acid biosynthetic pathways within cells (Schmidt, 2005). Additionally, plasmids being passed onto daughter cells is essential for the production of some secondary metabolites. It has been found that the high stress environment caused by scaling up can negatively impact plasmid stability especially due to the increased or decreased concentration of, for example, glucose within the fermentation vessel (Neubauer et al., 2003).

While there are no published studies that specifically explore the effects of production scale up when using *S. clavuligerus* within fermentation, several examples of the effects of scale up on a strains production have been previously documented. A study in *E. coli* investigated the emergence of non-producing cells during scale-up through analysis of growth rate and product formation of mevalonic acid (Rugbjerg et al., 2018). Here, 'escape rate' is used as a term to describe the emergence of mutations that lead to non-producing strains. The impact of production load, which is defined as the sum of the effects experienced from the fitness cost of pathway-specific biosynthesis, which is dependent on the metabolic burden of production of the specific product and the accumulation of inhibitory intermediates and by-products. The authors demonstrated that increased production load increases the escape rate, which negatively affects the mevalonic acid titre. Moreover, biosynthesis of mevalonlic acid is disrupted during the fermentation process, resulting in non-producing strains. This suggests that forcing strains to produce unnaturally high levels of a metabolite is likely to increase the metabolic burden on a strain and increase escape rates (Figure 1.5). In the same study, three lineages of industrial strains were sequenced over various timepoints to detect mutations within the individual genes responsible for the production of mevalonic acid. The authors found no detectable genetic variance (i.e. SNPs in above 1% of the population) in agreement with another study that showed up to a nine-fold increase in free fatty acid production from *E. coli* despite a lack of genetic variance between various strains.

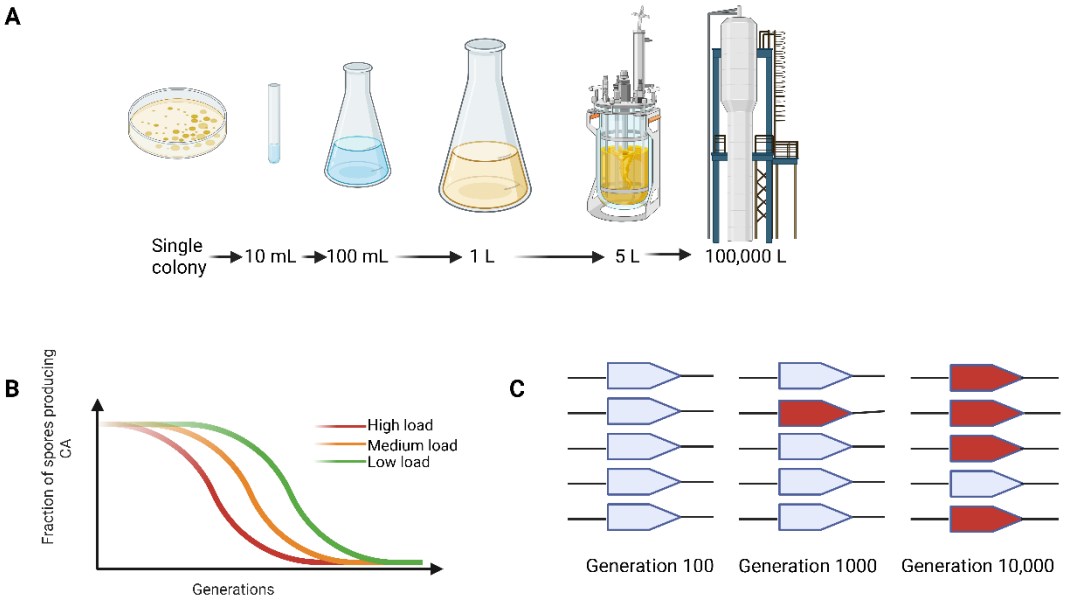


Figure 1.5. Escape rate model of various fermentation volumes given load. (A) Example of scaling up throughout an industrial fermentation, leading to increased production load. **(B)** Mathematical models can be used to predict the fraction of non-producing cells within a fermentation. Demonstrating that an increased production load also increases the escape rate. **(C)** As production load increases the escape rate increases causing more non-producing strains. Molecular microbiological methods can be used to track this heterogeneity within a population. Adapted from Delneri, 2010; Payen et al., 2016; Robinson et al., 2014.

The authors argue that these differences are due to nongenetic variances such as uneven cell division, epigenetic variation and mRNA instability (Xiao et al., 2016). Importantly, the study did find that as the number of generations in the fermentation increased, the copy number of the mevalonic acid biosynthetic pathway containing plasmid decreased (Rugbjerg et al., 2018).

1.5.4 CA fermentation

Since the advent of CA being produced on an industrial scale by *S. clavuligerus*, the fermentation process has become a critical factor in the titre and consistency of CA directly affecting profit from the product. The global production of CA can be attributed to just a few key fermentation sites such as in the UK, China and India. For example an industrial site in Mumbai has recently announced its plans to produce around 300 tonnes of CA per year (Thacker, 2024). However, this pales in comparison to GlaxoSmithKline's ambitions to produce 1,250 tonnes of CA at their Irvine industrial site, up from 900 tonnes (O'Keefe, 2016). CA fermentations commonly use complex media containing glycerol and soy-based nitrogen sources, which have been shown to enhance CA yields when paired with fed-batch strategies (López-Agudelo et al., 2021).

CA as a product has been well documented as being unstable in industrial media, synthetic buffer and pure water under neutral and alkaline pH (López-Agudelo et al., 2021), resulting in loss of product (Brethauer et al., 2008). The optimum pH for CA stability has been found to be 6.39 (Bersanetti et al., 2005; Haginaka et al., 1981), whereas the optimal growth for *S. clavuligerus* occurs in roughly neutral pH but robust growth has been observed in up to pH 8.5 (Higgens and Kastner, 1971). *Streptomyces* species also grow as mycelial aggregates in liquid culture which is known to affect production of specialised metabolites including CA (Nieminen et al., 2013a; van Dissel

et al., 2014). Pellet size and density can impact oxygen and nutrient diffusion, and smaller dispersed morphologies, as a result of higher shear stress, have been associated with improved CA titres (López-Agudelo et al., 2021).

Factors such as agitation, aeration, temperature and batch size have also been investigated as to the efficiency of *S. clavuligerus* fermentation (Viana Marques et al., 2018), which are all parameters that must be optimised for industrial fermentation of microorganisms, as previously discussed. In addition to strain engineering, precursor flux must also be carefully managed. *S. clavuligerus* diverts central carbon and nitrogen intermediates toward both CA and cephamycin C, meaning that competition between pathways can reduce CA titres unless biosynthetic balance is maintained (López-Agudelo et al., 2021).

A dynamic flux balance analysis (dFBA) in *S. clavuligerus* fermentations has been carried out to look at the relative flux of metabolites between primary and secondary metabolism (Gómez-Ríos et al., 2022). Models for the effects that various parameters had on central carbon metabolism within accurately predict the biosynthesis of CA and confirmed that fed-batch fermentation prolonged the growth phase and did not reduce or disrupt the metabolic precursors for the production of CA. More generally it has also been found that the maintenance of central carbon metabolism is essential during the industrial fermentation (at the 5L scale) of CA (Gómez-Ríos et al., 2022).

1.6 Improvement and production of antimicrobial producing strains

1.6.1 Random mutagenesis

While a number of *S. clavuligerus* strains have been published, a great many more exist as proprietary property. For decades industry has been producing new strains of *S. clavuligerus* that produce increasing levels of CA. The goal of developing strains that produce higher levels of CA (over the same amount of time and for the same or less nutrient input) is to lower the unit cost of CA for the producer and therefore for the consumer.

Strain improvement is still broadly an empirical process. This is because random mutagenesis is still one of the most common methods for the production of new industrial strains (Jeyachandran et al., 2023). Random mutagenesis involves causing genetic alterations in the DNA of industrial strains to alter its phenotype, aiming towards the production of more CA. Of course, as the nature of mutagenesis is random, many of the new strains produced by this method will not have the desired effect, as such, despite this method still being employed today within industry it can take many years to produce a strain with improved CA yields (Nielsen, 1997).

A number of mutagens are employed to alter the genetic material of industrial strains. These can be broadly classed into either chemical of physical mutagens. Physical mutagens include either UV or gamma radiation (Ikehata and Ono, 2011; Min et al., 2003) which causes cross linking between DNA strands or single or double stranded breaks, leading to structural changes. Chemical mutagens include, but are not limited to compounds such as 5- bromouracil, nitrous acid and ethyl methanesulfonate (EMS; Bökel, 2008; Hartman et al., 1994; Ma et al., 2008), which cause events such as faulty pairings and the incorrect alkylation or methylation of bases. Specifically for *S.*

clavuligerus, both chemical and physical random mutagenesis has been used for more than 50 years to achieve strains that produce increased levels of CA (Baltz, 2011).

Random mutagenesis and selection of higher producing strains mimics natural mutation to drive adaptive evolution. However, these industrial populations of *S. clavuligerus* may suffer from bottlenecks that could affect future strain development (Cisneros-Mayoral et al., 2022; Gregory, 2009). This is because random mutagenesis has given rise to strains that have increased transcriptions levels of biosynthetic genes, a greater copy number of the desirable biosynthetic gene cluster or the loss of competing biosynthetic gene clusters (Fiedurek et al., 2017; Medema et al., 2011; Yanai et al., 2006). However, off target effects may also occur within higher producing strains. This means that while industrial stains are very well adapted to produce high yields of CA within the industrial media they are specifically selected in, the unknown off target effects to their genetic material may lead to a reduced fitness landscape (Manna et al., 2011). This reduced fitness landscape allows strains to be increasingly adapted to the conditions in which they are being selected in but may shut the door to alternative pathways or methods of higher production that could be utilised by less adapted, more generalised strains. While work to better understand the effect of fitness landscapes on antimicrobial production is currently ongoing, it is thought that evolutionary outcomes depended heavily on initial conditions and chance. The work also suggest that many high fitness peaks can be reached through a large number of pathways of fitness increasing mutations, reducing predictability in the strain development process (Papkou et al., 2023).

1.6.2 Targeted approaches to strain improvement

While random mutagenesis has been used for the improvement of *S. clavuligerus* for many decades, a focus on targeted strain improvement has developed using an

increasingly complex molecular toolkit. Using targeted strain improvement is advantageous as it reduces the number of off target effects within the genome of *S. clavuligerus*. However, it comes at the cost of needing a deep understanding of the genome, transcriptome and proteome of the microorganisms in order to best engineer its metabolic pathways (Webb et al., 2022).

Although the media and laboratory conditions have a significant impact on the titre of CA, anecdotal evidence suggests that a 2004 industrial strain, created through targeted approaches, produced up to 3 g/l of CA which is around 30 x higher than that of the ATCC wild-type strain (Paradkar, 2013). Several successful efforts have been made to alter metabolism of *S. clavuligerus* to increase CA production. Disruption of glyceraldehyde-3-phosphate dehydrogenase (*gap*-1), responsible for directing the conversion of glyceraldehyde-3-phosphate (a direct precursor of CA) into glycerate-1,3-biphosphate in glycolysis, resulting around a 100% increase in CA production, presumably due to glyceraldehyde-3-phosphate being more abundant in the cell (Li and Townsend, 2006). In addition to this increase in CA production, the authors also showed that when arginine was used as a carbon source for the *gap*-1 mutants, CA production increases 200%. Other work focussed on precursor supply have indicated that glyceraldehyde-3-phosphate mediated upregulation of the *glp* operon resulted in an increase of up to 7.5 x increase in CA production when fermentations were supplemented with glycerol (Baños et al., 2009).

Previous efforts to improve CA production from *S. clavuligerus* have resulted in 2 to 20-fold increases in the production of the molecule (Song et al., 2010a). An increased understanding of *S. clavuligerus* through genome sequencing could help understand the

strain improvement process. After >30 years of industrial strain improvement, even small gains in CA titre are of value to industry.

Recent efforts to improve strains of *S. clavuligerus* for CA production have allowed key aspects of the physiology of *S. clavuligerus* to be elucidated. An *in silico* model of *S. clavuligerus* detailing 1021 genes and 1494 biochemical reactions within the species (Toro et al., 2018) utilised dFBA to show that reduced levels of phosphate and an increased ammonia within growth media are not beneficial for the biosynthesis of CA. Efforts to make a higher producing *S. clavuligerus* strains through overexpression of the *ccaR* and *claR* genes and a frameshift mutation within the *cas*1 gene via random mutagenesis increased production by ~43% (Cho et al., 2019). Similarly, overexpression of the glycerol utilization operon increased CA production by ~31% in shake flasks (Shin et al., 2021).

1.6.3 Molecular tools for targeted strain improvement

While the multi-omics profile *S. clavuligerus* is gradually better understood (Liras and Martín, 2021) the molecular tools needed to edit its genome are of course needed. These tools are reliant on homologous recombination and can be categorised into tools that perform either single crossover or double crossover homologous recombination events (Liu et al., 2018).

One such widely used tool that has been well used within the field are integrative vectors (Figure 1.6). In *Streptomyces* the integrating plasmids pMS82 and pSET152, which utilise the site specific ΦBT1 and ΦC31 integrase respectively are most commonly used. They have proved to be a reliable method to integrate plasmid DNA into a host chromosome via site-specific recombination (Kuhstoss et al., 1991). The plasmid pSET152 was first designed and used in 1992 (Bierman et al., 1992). The plasmid integrates into an *attB* site

however evidence suggests that the ΦC31 integrase inserts itself into a different *attB* site than that of ΦBT1 due to the production of site-specific recombination system with different specificities (Gregory et al., 2003). The integrating plasmid pMS82 was first published in 2003 and integrates into an *attB* site within *Streptomyces* genome which has a SCO4848 encoding a 79-amino-acid putative integral membrane protein (Gregory et al., 2003). Both integrase systems use unique 40 to 50 base pair *attP* attachment sites within the phage DNA that integrates into the *attB* site within the *Streptomyces* genome that create *attL* and *attR* sites as a result (Figure 4). The specific methods of action of the *att/int* ΦBT1 integrase and ΦC31 integrase systems are widely studied (Kormanec et al., 2019).

One barrier to the introduction of DNA to *Streptomyces* is the restriction-modification systems present that restricts methylated DNA. The restriction-modification systems exist as a defence mechanism against foreign DNA and cleaves incoming foreign DNA but can recognise self-DNA. Workarounds include the use of non-methylating *E. coli* strains (Mazodier et al., 1989). One example of a non-methylating *E. coli* strain is *E. coli* ET12567/pUZ8002 (Larcombe et al., 2024; MacNeil et al., 1992), originally developed for transformation into *S. avermitilis* the strain is both Dam and Dcm deficient allowing for direct conjugation into *Streptomyces*. The plasmid pUZ8002 carries transfer machinery for the strain. The advantages of using methylation deficient *E. coli* strains for transferring DNA to *Streptomyces* include the simplicity of *E. coli* and the lack of need to develop protoplast formation procedures. Here, restriction barriers are completely avoided or severely reduced and they have a number of possible *oriT* vectors that allow for various methods of chromosomal integration (Kieser et al., 2000). However, *E. coli* ET12567/pUZ8002 is not appropriate for long-term storage because the helper plasmid pUZ8002 is unstable without antibiotic selection and often lost over time. In addition,

the Δ*dam*, Δ*dcm* host background reduces overall genetic stability, making recovery after freezing unreliable and therefore should be used to conjugate demethylated DNA into *Streptomyces* immediately after cloning (Feeney et al., 2022).

More recent advancements in the genetic toolkit of modifications in *Streptomyces* include CRISPR based systems. Broadly, CRISPR is a gene-editing system derived from a bacterial immune mechanism that uses a guide RNA and the Cas9 enzyme to target and cut specific DNA sequences. This allows precise modifications to genes, enabling knockouts, insertions, or corrections in a wide range of organisms (Barrangou and Doudna, 2016). CRISPR has been successfully used within *Streptomyces* on a number of occasions. For example, a cumate-Based Inducible CRISPRi System has been used to alter actinorhodin production within *S. coelicolor* (Bai and van Wezel, 2023), a CRISPR-BEST system was used to inactivate the *kir*N gene within *S. collinus* (Tong et al., 2019) and a pCRISPR–dCas9 system has blocked transcription within *Streptomyces* without cleaving DNA (Tong et al., 2020). Additionally, CRISPR systems have been successfully used within *S. clavuligerus* to alter the expression of *tap* and *tpg* genes. These genes are essential to stop the chromosome from linearising due to their production of telomeric terminal proteins (Mohit, 2023). However, despite CRISPRs demonstrated use within *Streptomyces*, it is also clear that some CRIPSR systems are not suitable for certain uses within *S. clavuligerus*. For example it has been suggested that the system pCRISPomyces-2 is not suitable for use within *S. clavuligerus* despite being successfully used in other *Streptomyces* species due to the plasmid being quickly cured (Gomez-Escribano et al., 2021). The documented difficulty of working specifically with *S. clavuligerus* and CRISPR systems may arise from the fact that is already possesses a number of endogenous CRISPR-Cas systems (Ríos-Fernández et al., 2024), which may interfere with new systems being introduced for strain improvement in the future.

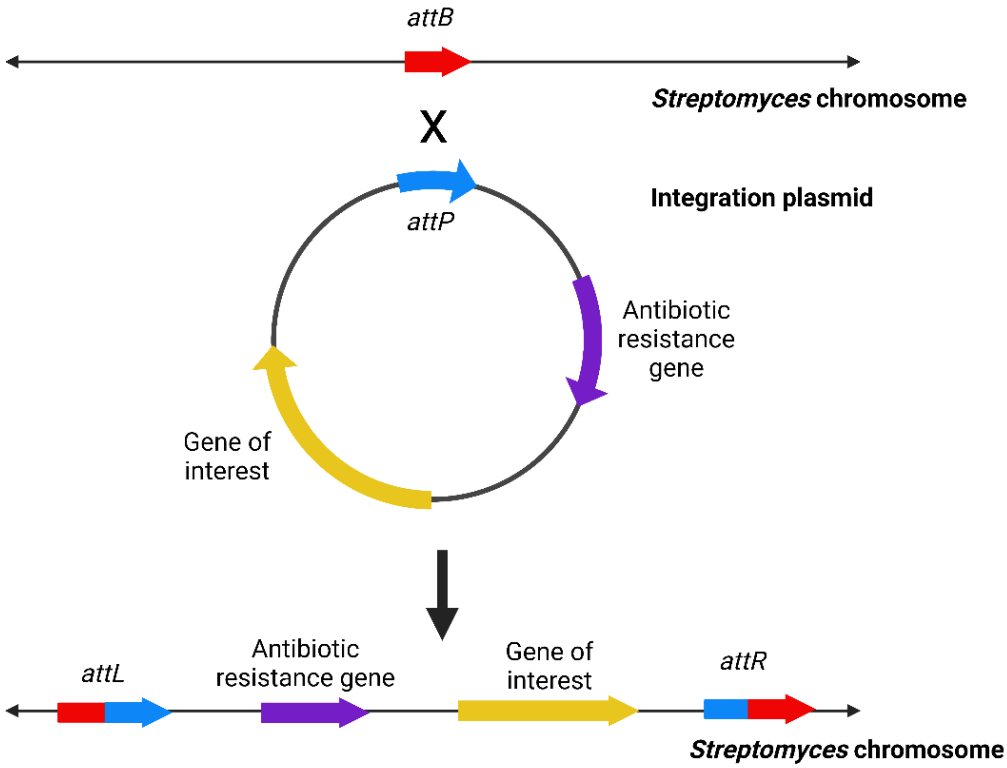


Figure 1.6. Example of an integrative plasmid inserting into a *Streptomyces* chromosome. The integrative plasmid inserts into the *attB* within the *Streptomyces* chromosome creating an *attL* and *attR* site. Adapted from Kormanec et al., 2019. Created using BioRender.

1.7 Population heterogeneity within fermentation

1.7.1 Why it's important

Understanding heterogeneity (the genotypic or phenotypic diversity within a bacterial population) within fermentation is key to interpreting and improving industrial bioprocesses. Microbial populations are often treated as uniform, but they are composed of subpopulations with different growth rates, metabolic activity, stress responses, and production levels. These differences can have major effects on yield, reproducibility, and scalability.

As discussed, phenotypic heterogeneity arises from stochastic gene expression, as well as gradients in substrate, oxygen, pH, or waste products, even in well mixed systems (Delvigne et al., 2014). These microenvironments help to drive population diversification. Some cells may be in an active production state, while other cells within the culture may be dormant.

This heterogeneity may have consequences for antimicrobial production. For example, whole population measurements such as OD or viscosity can miss subpopulation dynamics that affect process consistency. In recombinant systems for example, a minority of cells may lose plasmids or enter a low expression state due to burden or stress, reducing yield (Ackermann, 2015). Similarly, in prolonged or continuous cultures, adaptive evolution can lead to genetic divergence. Cells may acquire mutations that help them survive, but reduce productivity (Dragosits and Mattanovich, 2013). Without understanding or controlling heterogeneity, processes become harder to scale and optimise.

1.7.2 Tools to study heterogeneity

Studying heterogeneity in microbial populations requires tools that resolve sub population phenotypes and genotypes. Bulk measurements average out individual differences, masking rare or transient phenotypes that can cause a drop in antimicrobial production. An increasing number of tools allow for sub population analysis, but all come with limitations.

Flow cytometry is used to assess variability in size, granularity, and fluorescence markers at high throughput. It's effective for monitoring expression heterogeneity, stress responses, or differentiation (Delvigne and Goffin, 2014). However, it relies on appropriate fluorescent markers and provides no spatial or temporal information. For filamentous or adherent organisms like *Streptomyces*, detachment may skew results.

The lack of spatial or temporal data may be gained using fluorescence microscopy, especially time lapse imaging, which may give an insight into heterogenetic gene expression. It's useful for tracking cell cycle progress (Locke and Elowitz, 2009) but the technique is low throughput and labour-intensive.

Single-cell RNA sequencing (scRNA-seq) offers transcriptome wide profiling at single-cell resolution. It's powerful for identifying rare states and regulatory variation. Newer methods like PETRI-seq adapt this to bacteria (Blattman et al., 2020), but challenges remain. For example, low RNA yield (often seen from industrial media), incomplete lysis, and sparse coverage as well as the cost and complexity also limit its practical use in fermentation monitoring.

A common issue across all methods is scale. Most experiments are done under ideal lab conditions, which do not consider factors such as gradients, shear forces, and stress present in industrial fermenters. Many tools also require reporter constructs or staining,

which can affect cell behaviour. Sampling is another weak point, removing cells from their environment often distorts the very thing being measured. However, these sub-population tools help identify causes of inconsistent yields and guide strain and process optimisation. But interpretation needs to be grounded, detecting heterogeneity doesn't mean it matters. Being able to contextualise sub population behaviour to process outcomes is key. The question of how we can monitor population dynamics within a fermentation and how those dynamics may affect the production of antimicrobials remain.

1.7.3 Barcode sequencing: monitoring the emergence of population heterogeneity

A key part of understanding fermentations is being able to track which strains are present and how their frequencies change over time, even though only one, homogenous, strain should be present within the population. Tools for this are limited. Most previous work has relied on random transposon tagging, which isn't always practical or scalable. Barcode sequencing (Bar-seq) has emerged as a better alternative for tracking population dynamics, particularly in complex or competitive environments (Wetmore et al., 2015).

Bar-seq uses high-throughput sequencing to detect short, unique DNA barcodes inserted into different strains. It's a statistically powerful method for quantifying strain abundance within mixed populations. This approach has been used extensively in *Saccharomyces cerevisiae* mutant libraries to monitor strain fitness under a range of conditions (Robinson et al., 2014), and has shown clear potential in linking genotype to phenotype (Ferrari et al., 2021). Bar-seq has also been used to study glucose triggering oxidative stress resistance in *Candida albicans* (Larcombe et al., 2023).

Unlike approaches such as transcriptomics or proteomics which provide useful but static snapshots of population activity, Bar-seq enables temporal tracking of strain abundance across time points. It is only through these long term studies that allow us to get a key insight into the evolution of microbes (Lenski et al., 1991; Stroud and Ratcliff, 2025). Bar-seq can also be used alongside omics approaches to better interpret phenotypic changes. In *S. cerevisiae*, Bar-seq has helped link stress responses and drug mechanisms to specific mutant strains (Delneri, 2010; Payen et al., 2016), offering insight into how fitness traits emerge under different growth conditions.

Bar-seq has also been applied in combination with transposon sequencing to improve genome annotations. For example, one study identified 11,779 protein coding genes that were previously unannotated (Price et al., 2018). It's also been used in *Pseudomonas putida* to eliminate unnecessary metabolic pathways and improve the production of industrially relevant compounds like levulinic acid and valerolactam (Rand et al., 2017; Thompson et al., 2019). In another case, Bar-seq was used to assess the impact of scale-up on mutant populations of *S. cerevisiae*, comparing strain fitness between shake flasks and bioreactors (Wehrs et al., 2020). This study showed that selection pressures differed across scale and identified several mutants with growth advantages, including those with mutations in glycogen synthase kinase 3 (Gsk-3). The same study also tracked population diversity during fed-batch fermentation.

Despite these applications, Bar-seq hasn't been used to analyse *Streptomyces* populations. Given the genome plasticity of *Streptomyces* and their role in secondary metabolite production, Bar-seq could be a useful tool to study mutation rates and inter strain competition during fermentations. It would also help capture subpopulations that may not be obvious using bulk measurements. The lack of Bar-seq studies in

Streptomyces, or even in industrial bacterial fermentations more broadly, is surprising considering the demonstrated success of the technique in yeast.

For Bar-seq to be reliably used, the sequencing technology powering it must be accurate enough to detect and quantify barcode reads correctly. Historically, Sanger sequencing was the standard, but its low throughput, high cost, and tendency to introduce bias when barcodes are present at high amplicon concentrations make it suboptimal for this purpose (Polz and Cavanaugh, 1998; Smith et al., 2012).

Modern sequencing platforms like Illumina and Oxford Nanopore offers much better options. Illumina provides high read accuracy (~99%) and is widely used for barcode quantification (Stefan et al., 2022). Nanopore offers much longer reads, which can be useful for linking barcodes to nearby genomic context or identifying structural variants, but with lower raw accuracy, that has however increased with its new chemistry 10 kits (~95%) (Stevens et al., 2023). Hybrid approaches that combine Illumina's accuracy with Nanopore's read length can yield consensus sequences with >99% identity (McNaughton et al., 2019; Shokralla et al., 2014), although this may not be necessary for short barcode reads alone.

Applying Bar-seq to *Streptomyces* fermentations could help answer questions around strain stability, genome dynamics, and the emergence of fitter subpopulations under industrial conditions. When combined with transcriptomic or phenotypic data, Bar-seq could help build a more complete picture of how *Streptomyces* populations behave over time and how this links to production variability. Given the importance of *Streptomyces* in industrial biotechnology, and the challenges of working with a genetically unstable organism like *S. clavuligerus*, integrating Bar-seq into fermentation studies seems overdue.

1.8 Hypothesis and aims

1.8.1 Rationale

The emergence of poor or non-producing mutants in industrial *Streptomyces* fermentations reduces the efficiency and yields of commercial antibiotic production processes. It is hypothesised that this is the result of mutations and genome plasticity of *Streptomyces* species. The rate and nature of the mutations that occur are currently unknown. This work will develop methods to monitor changes in the population structure in the industrial strain *Streptomyces clavuligerus* during fermentations and identify mutations that result in the emergence of lineages that alter the bacterial population in the fermentation.

1.8.2 Specific aims of the project

- Design and validate a barcode sequencing (Bar-seq) technique that can identify and track the relative abundance of isogeneic strains of industrial *S. clavuligerus* during a fermentation
- Use Bar-seq to investigate population frequency changes in *S. clavuligerus* fermentations in shake flasks
- Use Bar-seq to investigate population frequency changes in *S. clavuligerus* fermentations in bioreactors
- Identify mutant phenotypes that arise during shake flask and fermenter level reactions of *S. clavuligerus*
- Identify mutations that emerge in *S. clavuligerus* fermentations and determine if there is parallelism in the mutations that emerge

Chapter 2 Materials and Methods

Throughout this project, all chemicals used were of analytical grade and were purchased from Fisher Scientific, Invitrogen, Qiagen or Sigma-Aldrich, unless stated otherwise.

2.1 Microbiological methods

Growth and maintenance of bacteria and media

Unless otherwise stated, all media was prepared in 500 mL Duran bottles and sterilised by autoclaving at 121 °C at 100 kPa. The media used throughout this study is listed in Table 2.1.

Streptomyces clavuligerus was grown at 26 °C on L3M9 medium to prepare single colonies of strains or for creating lawns of bacteria for preparation of spores. Liquid cultures of *Streptomyces clavuligerus* were grown at 26 °C in a shaking incubator at 250 rpm in 250 mL Erlenmeyer flasks filled to 50 mL of Tryptone Soy Broth (TSB). Spore stocks of *S. clavuligerus* were made from a streaking an existing spore stock for a single colony, which was picked and used to create confluent lawns of growth on L3M9 agar. Aliquots (2 mL) of sterile 25% glycerol (v/v) was added to the plate after ~10 days of growth. Spores were disrupted using a sterile cotton bud. The resulting mixture was passed through a sterilised cotton wool syringe to remove cell debris/hyphae according to (Feeney et al., 2022). Spores were stored in cryotubes at -20 °C or -80 °C. *Escherichia coli* strains were streaked on to Lysogeny Broth (LB) agar to prepare single colonies for overnight cultures at 37 °C. Liquid cultures of *E. coli* were grown in at 37 °C in a shaking incubator at 250 rpm. Stocks of *E. coli* were made from 10 mL of LB broth inoculated with a single colony of the relevant strain and grown for ~18 hours with appropriate

antibiotics. The resulting culture was then centrifuged at 20,000 *x g* for 10 minutes and the resulting pellet was resuspended in 1 mL of sterile 20% glycerol (0.5 mL of LB mixed with 0.5 mL of 40% glycerol). All *E. coli* stocks were stored at -20 °C or -80 °C.

Spore stock colony counts

Serial dilutions of *Streptomyces* stocks were prepared with a titre ranging from 10^{-1} to 10^{-8} colony forming units/ mL. An aliquot (90 µL) of these dilutions were plated onto LB agar and incubated at 37 °C for ~4 days. A plate with between 30 – 100 colonies was selected, and the number of colonies were counted. The concentration of the original stock was calculated from this dilution using the equation CFU/mL (original) = (N / V) x D, where N = number of colonies on the plate, V = volume plated in mL (0.09 mL) and D = dilution factor.

Table 2.1. Media and recipes used throughout this study

Media	Recipe	pH
L3M9 (Collis et al., 2021)	0.3 g Dextrin 10 g α-Trehalose 0.5 g Di-potassium hydrogen orthophosphate 1 g NaCl 1 g MgSO4 0.5 g CaCl 2 g Casamino acids 11 g MOPS powder 1 mL Trace salts* 30 g Roko agar 1 L dH_20 *0.5 g $FeSO_4$ $7H_20$ 0.5 g $ZnSO_4$ 0.5 g $MnSO_4$ Made up to 50 mL with distilled H_20	6.8
Tryptone Soya Broth (TSB)	30 g Tryptone soya broth (casein soya bean digest broth), premade mixture. Purchased from Oxoid Ltd., Basingstoke 1 L dH_20	6.8
Lysogeny Broth (LB) agar (Sambrook et al., 1989)	10 g Tryptone 5 g Yeast extract 5 g NaCl 15 g Agar Made up to 1000 mL with distilled H_20	N/A
Lysogeny Broth (LB) (Sambrook et al., 1989)	10 g Tryptone 5 g Yeast extract 5 g NaCl Made up to 1000 mL with distilled H_20	N/A
Complete Medium 5 (CM5). Modified from Collis et al., 2021.	40 g Rapeseed oil 18 g 70% Soy protein concentrate flour 11.5 g Maltodextrin spray dried 3 g Potassium dihydrogen orthophosphate 10.5 g MOPS powder 1.6 g MgSO4 . $7H_2O$ 0.3 g Ammonium sulphate 10 mL UFU Shakeflask trace salts * 1 L tap water *26 g MgSO4 . $7H_2O$ 3 g $FeSO_4$. $7H_2O$ 0.5 g $ZnCl_2$ 0.5g $CuCl_2$. $2H_2O$ 0.5 g $MnSO_4$. H_2O 1 L deionized H_2O	6.8

Seed Medium (S2A). Modified from Collis et al., 2021, Huckle, Personal Communication.	1 g Rape seed oil 25 g 70% Soy protein concentrate flour 20 g Dextrin 1 g Potassium dihydrogen orthophosphate 2.5 g $MgSO_4 . 7H_2O$ 20 mL STS trace salts* 1 L tap water *10 g $CaCl_2 . 2H_2O$ 10 g $MgCl_2 . 6H_2O$ 10 g NaCl 5 g $FeCl_3 . 6H_2O$ 0.5 g $ZnCl_2$ 0.5 g $CuCl_2 . H_2O$ 0.4 g $MnSO_4 . H_2O$ 1 L deionized H_2O	6.8
Soft nutrient agar (Kieser et al., 2000)	4 g Difco Nutrient Broth Powder 5 g Agar 1 L dH_2O	6.8
GYM agar (Kieser et al., 2000)	4 g Yeast extract 10 g Malt extract 2 g $CaCO_3$ 12 g Agar 8 mL 50% glucose (added after autoclaving) 1 L dH_2O	6.8
Minimal medium agar (Kieser et al., 2000)	0.5 g L-asparagine 0.5 g K_2HPO_4 0.2 g $MgSO_4 . 7H_2O$ 0.01 g $FeSO_4 . 7H_2O$ 20 mL 50% glucose (added after autoclaving)* 10 g Agar 1 L dH_2O *Glucose can be replaced by alternative carbon sources (e.g. 20 mL 50% sucrose (added after autoclaving)	6.8
Maltose Yeast Extract Medium (MYM) agar (Kieser et al., 2000)	4 g Maltose 4 g Yeast extract 10 g Malt extract 10 g Agar 500 mL tap water 500 mL dH_2O	6.8
Tryptone Soya Broth (TSB) agar	30 g Tryptone soya broth (casein soya bean digest broth), premade mixture. Purchased from Oxoid Ltd., Basingstoke. 15 g Agar 1 L dH_2O	6.8

Modified Yeast Extract Malt Extract agar (Hoskisson et al., 2000)	3 g Yeast extract 3 g Malt extract 5 g Peptone 10 g Glucose 1 L dH_2O 15 g Agar 2 mL $MgCl_2$ $6H_2O$ (2.5M) (added after autoclaving)	6.8
2 x Yeast Tryptone (YT) agar (Kieser et al., 2000)	16 g Tryptone 10 g Yeast extract 5 g NaCl 1 L dH_2O 15 g Agar	6.8
International *Streptomyces* Project (ISP) 3 agar (Shirling and Gottlieb, 1966)	20 g White Oats 18 g Agar 1 mL Trace salts solution* 1 L dH_2O *0.1 g $FeSO_4$ $7H_2O$ 0.1 g $MnCl_2$ $4H_2O$ 0.1 g $ZnSO_4$ $7H_2O$ 100 mL dH_2O	6.8
International *Streptomyces* Project (ISP) 4 agar (Shirling and Gottlieb, 1966)	10 g Soluble Starch 1 g $MgSO_4$ $7H_2O$ 1 g NaCl 2 g $(NH_4)_2SO_4$ 2 g $CaCO_3$ 1 mL Trace Salts solution* 1 L dH_2O *0.1 g $FeSO_4$ $7H_2O$ 0.1 g $MnCl_2$ $4H_2O$ 0.1 g $ZnSO_4$ $7H_2O$ 100 mL dH_2O 20 g agar	6.8
International *Streptomyces* Project (ISP) 7 agar (Shirling and Gottlieb, 1966).	15 g Glycerol 0.5 g Tyrosine 1 g L-asparagine 0.5g K_2HPO_4 0.5 g $MgSO_4$. $7H_2O$ 0.5 g NaCl 0.01 g $FeSO_4$. $7H_2O$ 1 L dH_2O 20 g Agar	6.8
Mannitol Soya Flour agar (MS) (Hobbs et al., 1989)	20 g Mannitol 20 g Soya protein flour 20 g Agar 1 L dH_2O	6.8
Cornmeal agar	2 g Corn Meal Agar (Purchased from Oxoid Ltd., Basingstoke) 15 g Agar 1 L dH2O	6.0

Starch agar (Feeney et al., 2022)	10 g soluble starch 2 g K_2HPO_4 2 g KNO_3 0.3 g casein 0.05 g $MgSO_4$ $7H_2O$ 0.02 g $CaCO_3$ 0.01 g $FeSO_4$ $7H_2O$ 15 g agar	6.8
Milk agar (Feeney et al., 2022)	50 g Dry milk powder 5 g Tryptone 2.5 g Yeast Extract 12.5 g Agar 5 g Glucose (added after autoclaving) 1 L dH2O	6.8
Reasoner's 2A (R2A) agar (Reasoner and Geldreich, 1985)	0.5 g Proteose peptone 0.5 g Casamino acids 0.5 g Yeast extract 0.5 g Dextrose 0.5 g Soluble starch 0.3 g Dipotassium phosphate 0.05 g Magnesium sulphate 0.3 g Sodium pyruvate 15 g Agar 1 L dH_2O	6.8
Sabouraud agar (Sabouraud, 1896)	40 g dextrose 10 g peptone 20 g agar 1 L dH_2O	5.6
Instant mash agar (Feeney et al., 2022)	20g Instant mashed potato 20g Agar 1 L Tap water	6.8

Table 2.2. Antibiotics and their working concentrations used throughout the project.

Antibiotic	Class	Working concentration	Storage Solvent
Apramycin	Aminoglycoside	50 µg/ mL	H_2O
Chloramphenicol	N-dichloroacylphenylpropanoid	25 µg/ mL	~98% Ethanol
Hygromycin B	Substitute aminoglycoside	100 µg/ mL	H_2O
Kanamycin	Aminoglycoside	50 µg/ mL	H_2O
Nalidixic acid	Quinolone	25 µg/ mL	H_2O
Ampicillin	Aminopenicillin	50 µg/ mL	H_2O
Thiostrepton	Thiopeptide	50 µg/ mL	DMSO

Cell growth quantifier (CGQ) and dry weight growth curves for *S. clavuligerus*

Aliquots (50 mL) of TSB were inoculated with 1 x 10^8 *Streptomyces clavuligerus* spores. For CGQ readings the aquilabiolabs (Baesweiler) Cell Growth Quantifier was used as per the manufacturer's guidelines with backscatter (AU) being used as the final measurement. The dry weight of cells contained in 1 mL of a culture sample was determined by washing the cells twice before resuspending the pellet in 1 mL of distilled water. Cells were applied to a GF/A grade micro glass fibre filter (purchased from Camlab, 1.6 µm pore size) that had dried to a constant mass. The biomass was filtered through a Buchner Funnel (KIF Laboport and NALGENE 180 PVC metric tube) at reduced pressure. The filter was rinsed with 2 mL of distilled water. The filter containing the biomass was dried to a constant mass. The biomass was determined gravimetrically (g/L).

Cell wet weight

Wet weights were calculated by spinning 1 mL of cell culture at 20,000 x *g* for 10 minutes in a 1.5 mL Eppendorf tube. The supernatant was removed and the difference in weight of the Eppendorf was calculated.

2.2 Molecular biology methods

Strains and plasmids

Table 2.3: Strains used in this study with a detailed description, genotype and source.

Strain	Description	Genotype	Reference/ source
E. coli DH5α	*E. coli* K12 derivative	F- ,Φ80, lacZΔM15, Δ(lacZYA-argF), U169, *recA1, endA1, hsdR17, (rk-, mk+), phoA, supE44, thi-1, gyrA96, relA1*	(Grant et al., 1990) Invitrogen
E. coli ET12567/pUZ8002	Methylation deficient strain allowing for direct conjugation into *S. clavuligerus*	*dam-13::Tn9, dcm-6, hsdM, hsdR recF143, zij201::*Tn10*, galK2, galT22, ara14, lacYI, xylS, leuB6, thi-1, tonA31, rpsL136, hisG4, tsx78, mtli, glnV44, F*	(Larcombe et al., 2024)
S. clavuligerus DSM 738	Progenitor strain of *S. clavuligerus* industrial strains	Wildtype	DSMZ
S. clavuligerus SC2	GSK production strain	GSK Production strain	GSK culture collection

S. clavuligerus SC6	GSK production strain	GSK production strain	GSK culture collection
Micrococcus luteus ATCC4698	Wild type *M. luteus* strain	Wildtype	(Fleming and Wright, 1997)

Table 2.4. Plasmids used in this study with detailed antibiotic selection and sources.

Plasmid name	Description	Resistance marker	Reference
pSET152	ΦC31 integrase containing plasmid	Apramycin	(Bierman et al., 1992)
pMS82	ΦBT1 integrase containing plasmid	Hygromycin	(Gregory et al., 2003)
pUC19	Commonly used cloning vector	Ampicillin	(Yanisch-Perron et al., 1985)
pSET152 JCb1-20	20 x pSET152 with unique 25bp barcodes	Apramycin	This study

Primer list

Table 2.5. Primers used in this study, designed on SnapGene™ made by IDT.

Primer	Sequence	Tm (°C)	Description
JTM105 pMS82 Fwd	GCAACAGTGCCGTTGATCGTGCTATG	64	For testing cloning inserts between HindIII and KpnI restriction sites
JTM106 pMS82 Rev	GCCAGTGGTATTTATGTCAACACCGCC	63	
JTM115 pSET152 Fwd	GCTGCGCCGATGGTTTCTACAAAGATCG	65	To amplify *lacZa* fragment of pSET152
JTM116 pSET152 Rev	GAGCGGATAACAATTTCACACAGGAAACAGCTATGAC	65	
JC pSET152 barseq Fwd	TTTCTGTTGGTGCTGATATTGCGTAAAACGACGGCCAGTGCCAAG	65	To amplify *lacZa* fragment of pSET152 with barcodes within the plasmid for the purposes of Bar-seq. With ONT ends for multiplexing
JC pSET152 barseq Rev	ACTTGCCTGTCGCTCTATCTTCCAGGAAACAGCTATGACATGATTACGAA	65	
JCb10 check fwd	GGGAATCGGGTAGGTCTTCCCTGTA	61	To detect the presence or absence of JCb10 within a colony
JCb10 check rev	CGCGTAATCTGCTGCTTGCAAAC		

PCR

Diagnostic and confirmatory PCRs were performed with REDTaq DNA Polymerase (Sigma) according to the manufacturer's instructions. Thermal cycling conditions are shown in Table 2.6 and do not alter unless otherwise stated. Unless otherwise stated PCR primers were used at a concentration of 100 µM.

Gel electrophoresis

Gels containing 1 % agarose were used for agarose gel electrophoresis. Agarose was dissolved in TAE buffer (40 mM Tris-acetate, 1 mM EDTA) using a microwave. The gel mixture was then cooled before ethidium bromide was added to a final concentration of 100 µg/ mL and poured into a casting tray, sealed with rubber stoppers. DNA samples were mixed with 6x blue/orange loading dye (Promega), and relevant molecular weight markers were chosen according to the expected DNA fragment size (Promega). Throughout this project, agarose gels were run using Bio-Rad power pack systems, with 1 x TAE used as the buffer. Unless otherwise stated, gels were run at 90 V for ~60 minutes. Gels were visualised and imaged under UV excitation using Syngene Bioimaging Ingenius trans-illuminator.

Plasmid DNA extraction

Plasmid DNA was extracted using the Wizard SV Minipreps Kit (Promega, 2010) according to the manufacturer's protocol (available at https://www.promega.co.uk/-/media/files/resources/protocols/technical-bulletins/0/wizard-plus-sv-minipreps-dna-purification-system-protocol.pdf?la=en).

Table 2.6. Red-taq thermal cycling conditions for PCRs run throughout this study.

* Primer annealing temperature was adjusted as required

Step	Temperature (°C)	Time
1.	95	1 minute
2.	95	15 seconds
3.	55 - 65 *	15 seconds
4.	72	20 seconds / kb
5. Repeat steps 2-4 29 times		

Table 2.7. Q5 high fidelity polymerase thermal cycling conditions for PCRs run throughout this study. * Primer annealing temperature was adjusted as required

Step	Temperature (°C)	Time
1.	98	30 seconds
2.	98	10 seconds
3.	50 - 72 *	30 seconds
4.	72	30 seconds / kb
5. Repeat steps 2-4 29 times		
6.	72	2 minutes

Electrotransformation of *Escherichia coli* ET12567/pUZ8002

Aliquots (100 µL) of electrocompetent *E. coli* ET12567/pUZ8002 were mixed with plasmid DNA (10 ng) in an ice-cold electroporation cuvette. The sample was electroporated with BioRad GenePulser II set to 200 Ω, 25 µF and 2.5Kv with an expected time constant of 0.4 – 0.6 seconds. 1 mL of ice-cold LB was added to the cuvette and mixed via inversion. The entire sample was then transferred to an Eppendorf and allowed to recover at 37 °C and shaking at 250 rpm for three hours. Following recovery, cells were plated on LB agar containing the appropriate antibiotics and incubated overnight at 37 °C.

Transformation of *E. coli* DH5α using heat shock

Commercial *E. coli* DH5α (50 µL) were thawed on ice for 10 minutes. Aliquots (10 ng) of plasmid DNA was gently mixed with the thawed cells. This mixture was incubated on ice for 30 minutes. The transformation mixture was then heat shocked at 42 °C for 1 minute, before being placed on ice for 2 minutes. 1 mL of LB broth was then added to the mixture and the cells were allowed to recover for 1 hour at 37 °C, shaking at 250 rpm. These cells were then plated on LB agar with appropriate antibiotics and incubated overnight at 37 °C.

Conjugation of pMS82 and pSET152 into *Streptomyces clavuligerus*

Overnight cultures of *E. coli* ET12567/pUZ8002 were prepared in LB broth (see above). Aliquots (400 µL) of the culture were inoculated into 10 mL of LB broth with appropriate antibiotics and grown to an OD_{600} 0.6. This sample was centrifuged at 20,000 x *g* and

washed three times with LB broth with a final volume of 500 µL. Freshly collected *S. clavuligerus* spore stocks (1 x 10^8 spores) were centrifuged at 20,000 x *g* and resuspended in 500 µL of LB broth. The 500 µL aliquots of spores and *E. coli* cells were mixed in an Eppendorf and centrifuged at 16,000 x *g* in a microfuge and the supernatant was discarded. The pellet was resuspended in 100 µL of LB broth which was plated onto L3M9 agar containing 10 mM of $MgCl_2$ solution. After 18 hours of growth at 26 °C, the plates containing *S. clavuligerus* and *E. coli* ET12567/pUZ8002 were overlayed with 1 mL of sterile antibiotic solution (containing 25 µL of each appropriate antibiotic and 25 µL of nalidixic acid, made up to 1 mL with sterile H_2O). Primary exconjugants were selected after seven days of growth and streaked on to fresh plates of L3M9 agar containing the appropriate antibiotics.

2.3 Sequencing and bioinformatic methods

Genomic DNA extraction from *S. clavuligerus*

Genomic DNA was isolated via salt extraction. *S. clavuligerus* culture grown for ~48 hours in TSB at 26 °C (as above). The culture was centrifuged at 20,000 x *g* for 30 minutes with the supernatant being discarded. The pellet was resuspended in 3 mL of TES buffer (75 mM NaCl, 25 mM EDTA pH 8, 20 mM Tris-HCl pH 7.5). After, 150 µL of 20 mg/mL lysozyme was added and incubated for at 250 rpm, 37 °C for ~90 minutes. Proteinase K was added (84 µL of 20 mg/mL) alongside 360 µL of 10% SDS solution, mixed via inversion and incubated at 55 °C for 2 hours, inverting occasionally. Then, 1.2 mL of 5M NaCl was then added to the mixture and mixed via inversion, the sample was allowed to cool, before the addition of 3 mL of chloroform, mixed via inversion and left to incubate at room temperature for 30 minutes. The sample was then centrifuged at 20,000 x *g* for 15 minutes. The upper phase of the sample was then transferred to a sterile falcon tube.

An equal volume of ice-cold isopropanol was added and the precipitated DNA was spooled out using a sterile, sealed, glass Pasteur pipette and stored in an Eppendorf tube. The DNA was washed with 1 mL of 70% ethanol and the sample was centrifuged at 20,000 x *g* for 10 minutes, before the supernatant was removed and the residual ethanol allowed to evaporate. The DNA was solubilised in 100 µL of sterile H_2O for storage.

Amplicon library preparation and next-generation sequencing

After amplicons containing unique barcodes were obtained by PCR from the genomic DNA from fermentations, amplicon concentration of purified DNA was determined by Qubit BR (Termo Fischer). If multiplexing within a single run, multiple samples were pooled in equimolar concentrations. A library of 200 fmol of amplicons were prepared using the Nanopore ligation sequencing kit (SQK-LSK114, Oxford Nanopore Technologies) as per the manufacturer's guidelines. The prepared library was then sequenced on a Flongle flowcell (FLO-FLG114, Oxford Nanopore Technologies), basecalled and demultiplexed if required using MinKNOW software (version 22.12.7, Oxford Nanopore Technologies) as per the manufacturer's instructions.

Barcode frequency calculations

To calculate the relative abundance of each barcode, the FASTQ files produced by Oxford Nanopore Technologies basecalling were used as raw data. Pycharm (community edition 2022.2.1) was used to run the code (that can be found at https://github.com/jamescroxford/Bar-Seq.git), to count the forward and reverse complements of each barcode accounting for a Levenshtein distance of 1 (lines 52 and

59 of code respectively). Barcode relative abundance was calculated by dividing the count of each barcode by the total number of detected barcodes. Barcode relative abundance was used as a proxy for the relative abundance of each barcoded strain of *S. clavuligerus*.

Log_2 fold change of barcoded strain relative abundance

The log_2 fold change of a barcode comparing the start and end relative abundance was calculated. The relative abundance at the final timepoint was divided by the relative abundance at the start timepoint. The log() function on Microsoft Excel was used to find the log_2 of this fold change value.

S. clavuligerus SC2 and SC6 reference genome assemblies

The Streptomyces clavuligerus SC2 and SC6 reference genomes used throughout this thesis were assembled in-house by John Munnoch (University of Strathclyde) using Illumina short-read sequencing for both strains and PacBio long-read sequencing for SC2 and SC6. Illumina reads were generated on the Illumina MiniSeq (SC2) and NovaSeq 6000 (SC6) following manufacturer protocols. PacBio long-read sequencing for SC6 was performed as described in Gomez-Escribano et al., (2021). PacBio long-read sequencing for SC2 was performed on a PacBio Sequel instrument (Nu-omics, University of Northumbria, UK) using SMRTbell library preparation and 10-hour movie capture.

PacBio data were assembled using HGAP4 (Chin et al., 2013). Illumina reads were assembled de novo using SPAdes (Bankevich et al., 2012) and used for polishing/validation against the long-read assemblies. Where required, assemblies

were manually curated and structural variants were confirmed by PCR and Sanger sequencing. Raw sequencing data are available under BioSamples SAMN42008322–SAMN42008326 and SAMN51091877 (NCBI), and the re-annotated assemblies (generated using Prokka; Seemann, 2014) are available under BioProject PRJNA1127551.

Alignment of Illumina reads using Bowtie2

Illumina sequencing of the *S. clavuligerus* genome was performed by Novogene Europe on an NovaSeq X Plus25B Flow cell, where 150bp paired end reads were created. Bowtie2 (Version 2.5.2; Langmead and Salzberg, 2012) were used to map the reads against the *S. clavuligerus* SC6 reference genome (BioProject PRJNA1127551) with output files being in binary alignments (BAM) format with a binary alignments index. Mapped reads were visualised using JBrowse2 (Diesh et al., 2023).

Breseq analysis (Deatherage and Barrick, 2014)

Whole-genome sequencing (WGS) data generated by Novogene Europe from genomic DNA (Illumina NovaSeq X Plus, 25B flow cell, 150 bp paired-end) were analysed using breseq (version 0.39.0; Deatherage and Barrick, 2014). Reads were aligned to the *Streptomyces clavuligerus* SC6 reference genome (BioProject PRJNA1127551) and variants were called using the default pipeline. Breseq was run on a Xubuntu virtual machine (version 22.04).

To control for background variation, the parental SC6 genome was first analysed with breseq and mutations identified in this stock strain were treated as baseline signatures.

These background mutations were excluded from subsequent comparisons of fermentation-derived strains, ensuring that only novel mutations relative to SC6 were reported.

RNA extraction from *S. clavuligerus*

Streptomyces clavuligerus cultures were grown in 250 mL Erlenmeyer flasks containing 50 mL of TSB at 26 °C and 250 rpm. Samples were collected during exponential and stationary phases, with three biological replicates per condition. For each replicate, 1 mL of culture was transferred into a 2 mL screw-cap tube containing an equal volume of RNAlater (Qiagen) and immediately mixed by inversion. Samples were stored at −80 °C until extraction.

Cell pellets were obtained by centrifugation of thawed samples at 20,000 × *g* for 10 minutes at 4 °C. The pellet was resuspended in 100 µL TE buffer (50 mM Tris-HCl, 20 mM EDTA, pH 8.0), followed by the addition of 100 µL lysozyme (100 mg/mL) and incubation at 30 °C for 30 minutes. Cells were lysed by adding 50 µL of 10% SDS and 85 µL of 5 M NaCl, after which 500 µL TRIzol reagent (Thermo Fisher Scientific) was added. Samples were mixed by inversion before addition of 300 µL chloroform, vortexed for 15 seconds, and centrifuged at 20,000 × *g* for 10 minutes at 4 °C.

The aqueous phase was transferred to a new tube and applied to RNeasy spin columns (Qiagen) according to the manufacturer's instructions. An on-column DNase digestion was performed by adding 10 µL RNase-free DNase (Qiagen) with 70 µL RDD buffer, incubating for 60 minutes at room temperature. An additional off-column DNase digestion was carried out for all samples (RNeasy MinElute Cleanup Kit, Qiagen)

according to the manufacturer's protocol. RNA was eluted in 50 µL RNase-free water at 50 °C, and the eluate was reapplied to the same column for a second elution in 50 µL

RNA concentration and purity were measured using a Nanodrop 2000 spectrophotometer (Thermo Fisher Scientific). RNA integrity was assessed using the Agilent 2100 Bioanalyzer with RNA 6000 Nano reagents (Agilent Technologies), with RNA integrity number (RIN) values used to confirm sample quality. DNA contamination was checked by 260/280 absorbance ratios. High-quality RNA samples were used for downstream transcriptomic analysis.

Transcriptomic analysis

Following RNA quality control, prokaryotic directional mRNA libraries were prepared using rRNA removal (Novogene Europe) and sequenced on an Illumina NovaSeq X Plus platform on the 25B flow cell, to produce 150 bp paired-end reads, generating ~2 Gb of raw data per sample. Three biological replicates were sequenced per condition. FASTQ files were returned for downstream analysis.

Transcript abundance was quantified using Salmon (Patro et al., 2017) against an annotated *S. clavuligerus* SC2 genome (BioProject PRJNA1127551). This reference was selected as it provides more complete annotation than the SC6 genome and contains all genes identified in SC6. A tx2gene table was constructed from the annotation and used with the tximport package to summarise transcript counts at the gene level. DESeq2 (Love et al., 2014) was then used to normalise count data and test for differential expression between conditions as required.

Principal component analysis (PCA) was performed on rlog-transformed counts to assess replicate clustering. Differentially expressed genes (DEGs) were identified using a false discovery rate (padj) ≤ 0.05 and $\log_2$fold change > 1. Volcano plots were generated using the EnhancedVolcano package (Blighe, 2025) to visualise DEGs.

All analyses were carried out in R (version 4.3.0) within RStudio (version 2023.3.1) using tximport (version 1.30.0) and DESeq2 (version 1.42.1). Transcript quantification with Salmon (version 1.10.2) was performed on a Xubuntu virtual machine (version 22.04).

2.4 Biochemical assays

Clavulanic acid (CA) assay (Bird et al., 1982)

CA concentrations in culture supernatant were determined by 8 µL of supernatant that had previously been cleared by centrifugation for 5 minutes at 16,000 x *g* in a microfuge being added to 200 µL of imidazole (prepared by dissolving 10 g imidazole (VWR Chemicals) in 7.2 mL 10M HCl and made up to a final volume of 100 mL with distilled water). CA standard solutions of the following concentrations were assayed along with the samples: 400 mg/mL, 350 mg/mL, 300 mg/mL, 250 mg/mL, 200 mg/mL, 100 mg/mL, for which the appropriate amount of potassium clavulanate (purchased from Merck) was dissolved in distilled water and filter sterilised. The standard solutions were stored at -80 °C. Samples (technical triplicates) were analysed by measuring absorbance at 324 nm following a 30-minute incubation at room temperature on a Microplate Reader System (BMG Labtech). A calibration curve was constructed from the CA standards, which was used to determine CA concentrations in the culture supernatants.

Total protein concentration assay

Total protein concentration within a solution was determined by pelleting 1 mL cell culture and removing the supernatant. The cell pellet was resuspended in BugBuster (Qiagen) (5 mL of BugBuster was used per 1 mL cell wet weight) and incubated at room temperature for 30 minutes at 100 rpm. The solution was then spun at 20,000 x *g* for 10 minutes. 10 µL of supernatant was added to 200 µL of Bradford reagent (BioRad) that had been diluted 1:4 with dH_2O. Protein standards were made using 0 mg/mL, 0.2 mg/mL, 0.4 mg/mL, 0.6 mg/mL, 0.8 mg/mL and 1 mg/mL BSA. The solution was left for 5 minutes at room temperature. Protein concentration was determined using a Microplate Reader System (BMG Labtech) reading at 595 nm with a standard curve being used to calculate the protein concentration within the 1 mL of cell culture.

Agar plug bioassay against *Micrococcus luteus*

M. luteus was cultured overnight from a single colony isolation at 37 °C at 250 rpm in 10 mL of LB. Aliquots (1 mL) of overnight culture was used to inoculate 10 mL of fresh LB and was grown at 37 °C at 250 rpm for 6 hours. Once grown, 15 mL of soft nutrient agar was inoculated with this *M. luteus* to a final OD_{600} of 0.01 and poured onto petri dishes containing 20 mL of LB agar. Agar plugs (6 mm diameter) from L3M9 culture plates of the 10 randomly chosen end-point *S. clavuligerus* SC6 pSET152 JCbx fermentation strains were placed onto the *M. luteus* containing plates in triplicate. Negative controls were an L3M9 media plug and positive controls were co-amoxiclav paper discs (30 µg). Plates were incubated for 24 hrs at 37 °C then zones of inhibition (mm) were measured with a ruler.

Glucose-6-phosphate dehydrogenase assay

Streptomyces clavuligerus was grown in Erlenmeyer flasks for 48 hours and then centrifuged at 20,000 *x g* for 15 minutes (as above). The supernatant was removed, and the remaining pellet was resuspended in 3 mL of ice-cold PBS. Then, 150 µL of 20 mg/mL lysozyme was added and incubated for at 250 rpm, 37 °C for 90 minutes. The sample was then centrifuged at 20,000 *x g* for 15 minutes with the supernatant being transferred into a fresh 15 mL falcon tube. The Glucose 6 Phosphate Dehydrogenase Activity Assay Kit – colorimetric (Abcam, ab102529) was then used to test the level and activity of glucose 6 phosphate dehydrogenase within the supernatant according to the manufacturer's guidelines. Activity was normalised to total protein in solution measured using a total protein concentration assay.

Phase contrast microscopy

Streptomyces clavuligerus was cultured in 250 mL Erlenmeyer flasks containing 50 mL of TSB for 48 hours at 26 °C and 250 rpm. A 10 µL aliquot of culture was applied directly onto a sterile glass slide for imaging. Samples were visualised using a TE-2000 inverted widefield epi-fluorescence microscope (Nikon, Japan) equipped with a ×20/0.50 numerical aperture objective lens. Phase contrast images were acquired with a Hamamatsu ORCA-100 CCD camera under tungsten-halide illumination. Images were collected using NIS Elements software and processed in FIJI (Schindelin et al., 2012).

Images were scaled at 2360 pixels/mm using a 1 mm graticule (Analyse/Set Scale). The area of cell aggregates was measured by manually outlining pellets with the freehand drawing tool, followed by Analyse/Measure. Processed images were flattened (Image/Overlay/Flatten) and exported for further analysis.

ImageJ analysis

Colony size was quantified from plate images using ImageJ (Schindelin et al., 2012). Images were first calibrated by setting the scale with a 1 cm graticule, repeated for each plate. Colonies were detected using the Analyse Particles function following manual thresholding against the background. Both colony area and diameter were recorded. Output data were exported for downstream analysis.

2.5 Graphs, plots and statistical tests

All data was visualised using the ggplot2 package within Rstudio (version 2023.3.1) unless otherwise stated. All statistical analysis was carried out using Rstudio (version 2023.3.1) running R (version 4.3.0). All Linux softwares were run using a Xubuntu (version 22.04) virtual machine.

Chapter 3 Establishing and validating Bar-seq in *S. clavuligerus*

3.1 Introduction

Understanding and controlling strain-level heterogeneity during *Streptomyces* fermentations is essential for the optimisation of industrial antimicrobial production. Even small changes in population composition can significantly impact yields and reproducibility (Harms et al., 2017; Hoskisson and Hobbs, 2005). Within a mixed population individual strains may vary in their metabolic output or competitive fitness, yet current methodologies to study population dynamics in filamentous bacteria remain inadequate. Approaches such as colony morphology screening or endpoint plate counts offer little temporal resolution and cannot robustly link genotypic identity to phenotypic outcome. More advanced methods, including genome sequencing of isolates, are labour-intensive and not suited for routine monitoring of large populations over time. A new, reliable method to monitor population dynamics while preserving strain identity is therefore needed.

Barcode sequencing (Bar-seq) is a high-throughput method that quantifies the relative abundance of uniquely tagged strains in a population by amplifying and sequencing strain-specific DNA barcodes. It has been successfully applied in yeast to investigate fitness of mutants and competitive interactions (Larcombe et al., 2023). However, its utility in bacteria, and particularly in *Streptomyces*, has not been established. For Bar-seq to be adopted in this context, its accuracy, reproducibility, and suitability for tracking populations in fermentation-like conditions must be validated. An additional challenge is ensuring that the integrated barcodes themselves do not perturb strain fitness or introduce artefactual biases in abundance estimates.

To address these knowledge gaps, the aim of this chapter is to explore the design, construction, and validation of a Bar-seq platform in *Streptomyces clavuligerus* based on the barcodes previously used in *Candida albicans* (Larcombe et al., 2023).

Following construction of the barcode plasmids, the barcode sequences will be evaluated *in silico* and *in vitro* to confirm their distinguishability and stability. This will lay the foundation for applying Bar-seq to industrially relevant fermentation experiments, where linking genotype to phenotype is essential for understanding population stability and optimising production.

All experimental work in this chapter, including strain handling, barcode amplification, sequencing, downstream analysis, and figure preparation, was performed by James Croxford. Barcode sequences were provided by Daniel Larcombe, and barcode constructs were synthesised and assembled by GenScript (Oxford, United Kingdom) as described.

3.2 Proposed barcode sequencing (Bar-seq) methodology

Bar-seq has not previously been used to study bacterial population dynamics. To establish its use in *S. clavuligerus*, a Bar-seq method was proposed (Figure 3.1). The barcode sequences used here were provided by Daniel Larcombe and have previously been validated in *Candida albicans* (Larcombe et al., 2023).

In this approach, a set of otherwise isogenic *S. clavuligerus* strains will be constructed, each carrying a unique barcode integrated into the chromosome via standard integrative vectors such as pMS82 or pSET152. *S. clavuligerus* SC6 was chosen because it is the industrial strain of interest. Each barcode will be stably inserted so that it is maintained over multiple generations.

To track relative strain abundance over time, barcoded strains will be grown individually and mixed in equal proportions to generate a pooled inoculum. This mixture will be used to start a fermentation under defined conditions. Genomic DNA will be extracted at multiple timepoints (e.g. start and end of fermentation), and the barcode containing

region will be PCR amplified from the populations genomic DNA. This pool of amplicons will then be sequenced.

The number of reads corresponding to each barcode will be counted and used to infer the relative abundance of each strain at each timepoint. As each strain is isogenic apart from the barcode, the barcode read count acts as a direct proxy for strain abundance. This allows us to assess how the relative abundance of each strain changes over time and to investigate stability and dynamics in mixed strain populations.

This method can also be used to link phenotype to genotype. For example, if an individual colony isolated from a fermentation plate shows a distinct phenotype, such as altered morphology or sporulation, a single colony can be picked, its DNA extracted, and the barcode amplified using the same PCR reaction used for the pooled samples. Sequencing the barcode then reveals which strain this colony corresponds to. This enables retrospective tracking of where that barcode (and therefore the phenotype) was present and at what abundance throughout the fermentation.

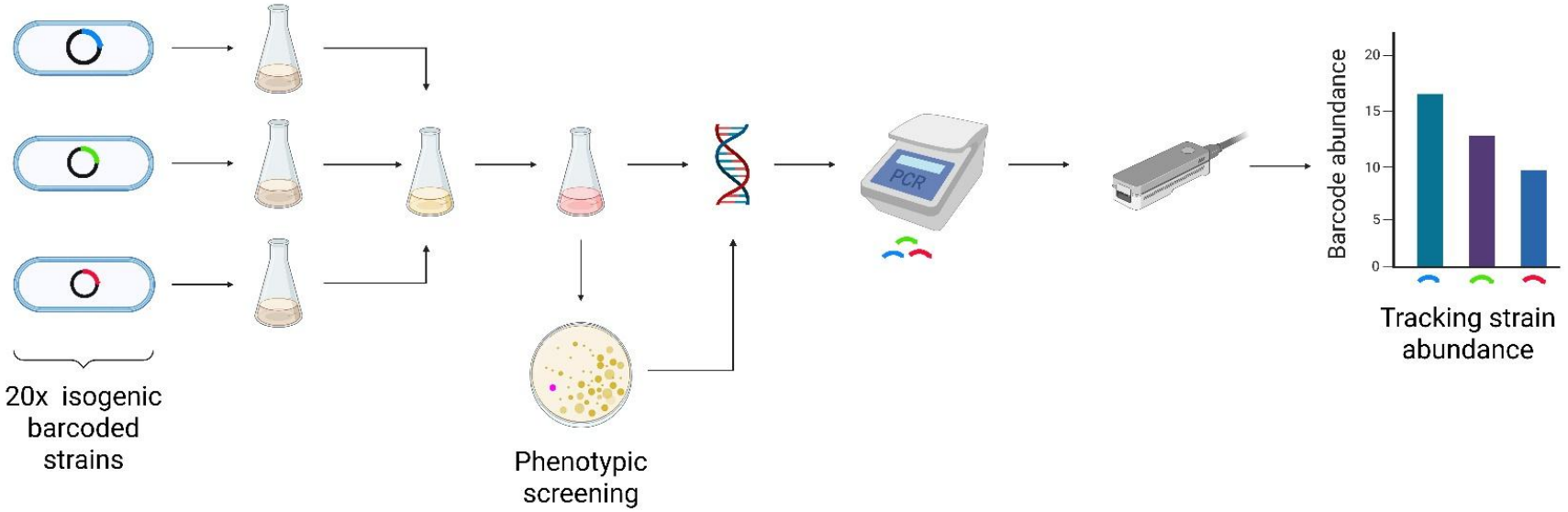


Figure 3.1. Proposed barcode sequencing methodology. Twenty isogenic strains of *S. clavuligerus* SC6, each carrying a unique chromosomally integrated barcode, will be grown separately and then mixed in equal proportions. This mixture will inoculate a fermentation. Genomic DNA will be extracted at selected timepoints, the barcode region amplified, and amplicons sequenced. The relative abundance of each barcode will reflect the abundance of the corresponding strain in the population over time.

3.3 Selection of integrative plasmids used within this study

After confirming the sequences of the barcodes used to tag *S. clavuligerus* SC6, a method was needed to integrate the barcode sequences into its chromosome. The integrative plasmids pMS82 and pSET152 are widely used to insert genetic material into *Streptomyces*. While much validation has previously been done on the two vectors, no work has been published analysing potential effects of the integrative vectors specifically on *S. clavuligerus* SC2 and SC6. The effects of both empty vectors of *S. clavuligerus* SC2 and SC6 was investigated to ensure they would not affect growth of the strain within a fermentation. Both vectors were introduced into *S. clavuligerus* SC2 and SC6 via conjugation to assess if either had unwanted off target/growth effects in the *S. clavuligerus* SC2 or SC6 strains. PCRs were carried out to validate the successful conjugation and integration of the plasmids into the genome. Primers JTM105 pMS82 Fwd, JTM106 pMS82 Rev, JTM115 pSET152 Fwd and JTM116 pSET152 Rev were used for pMS82 and pSET152 respectively. Bands at 230 bp and 410 bp respectively for pMS82 and pSET152 were obtained from PCR of the relevant genomic DNA confirming successful conjugation into both *S. clavuligerus* SC2 and SC6 (Figures 3.2 and 3.3).

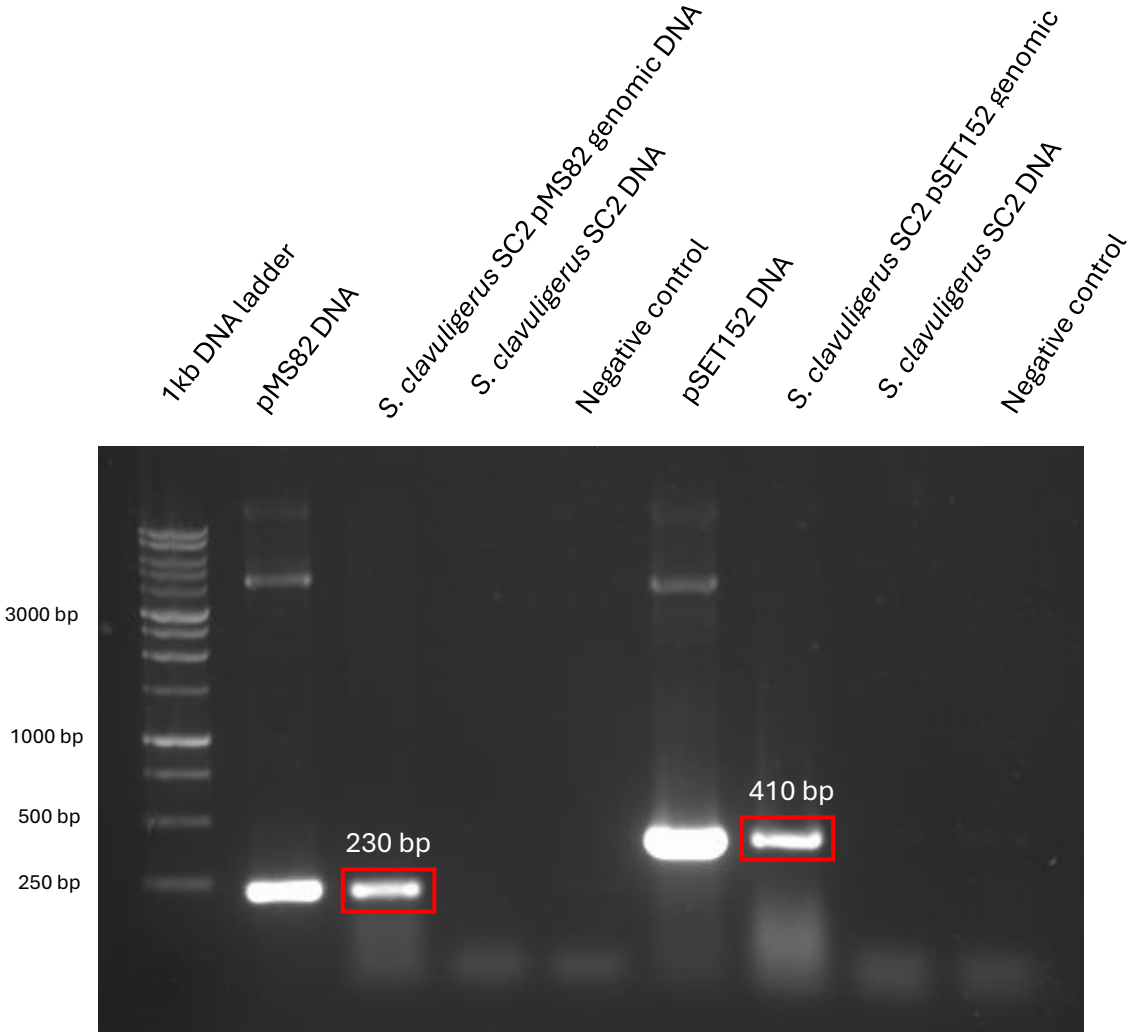


Figure 3.2. Successful amplification of pMS82 and pSET152 DNA from *S. clavuligerus SC2* genomic DNA. Agarose gel of amplified pMS82 and pSET152 from *S. clavuligerus SC2* using primers JTM105 pMS82 forward and reverse and JTM115 pSET152 forward and reverse, red boxes indicate the relevant plasmids are present at approximately 230bp and 410bp respectively. Positive control is pMS82 and pSET152 DNA, negative control is nuclease free water.

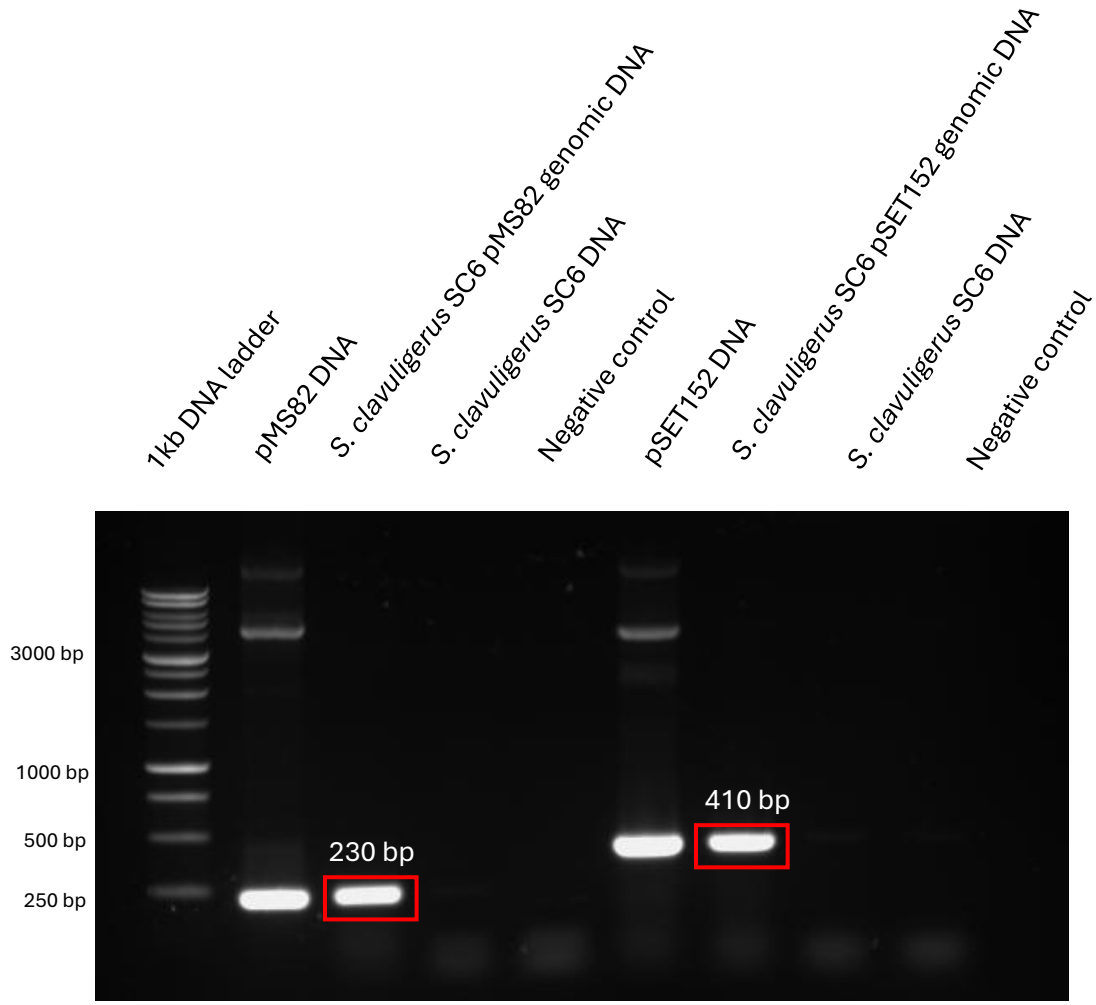


Figure 3.3. Successful amplification of pMS82 and pSET152 DNA from *S. clavuligerus SC6* genomic DNA. Agarose gel of amplified pMS82 and pSET152 from *S. clavuligerus SC6* using primers JTM105 pMS82 forward and reverse and JTM115 pSET152 forward and reverse, red boxes indicate the relevant plasmids are present at approximately 230bp and 410bp respectively. Positive control is pMS82 and pSET152 DNA, negative control is nuclease free water.

After demonstrating the successful conjugation of pMS82 and pSET152 in *S. clavuligerus* SC2 and SC6, the effect of both vectors on the growth of *S. clavuligerus* SC6 was assessed as this was the most relevant industrial strain within fermentation. To achieve this, *S. clavuligerus* SC6 and *S. clavuligerus* SC6 pMS82 and pSET152 were grown in 50 mL TSB for 48 hours, with backscatter of the culture being measured by the aquilabiolabs Cell Growth Quantifier (Figure 3.4). Based on these growth curves, the specific growth rate for each strain was calculated by using backscatter values from the start and end of the exponential growth phase for each strain (Table 3.1). The following equation was used to calculate the specific growth rate $\mu = ((\log_{10} N - \log_{10} N_0)\ 2.303) / (t - t_0)$ where μ is the specific growth rate, N is the backscatter value and t is the time corresponding to that value.

It was observed that pSET152 has a negligible impact on the specific growth rate (a 1.22% increase compared to 'wild type'), however, pMS82 reduces the specific growth rate by 1.502×10^{-3} μ (17.6%). Based on these specific growth rates it was determined that both vectors could be appropriate for use of integration of barcode sequences into the chromosome, however pSET152 would be preferable.

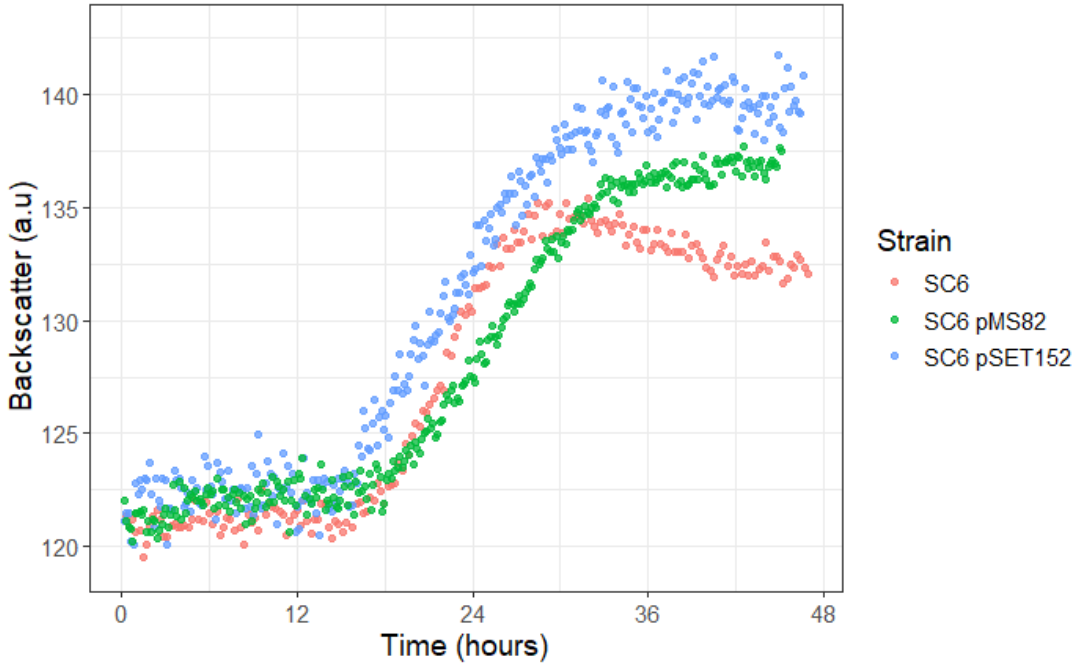


Figure 3.4. Analysis of the effects of pMS82 and pSET152 on the growth of *S. clavuligerus* SC6. *S. clavuligerus SC6, S. clavuligerus SC6 pMS82 and S. clavuligerus SC6 pSET152* were grown for 48 hours in 50 mL TSB. Backscatter was quantified using the cell growth quantifier.

Table 3.1. The plasmids pMS82 and pSET152 do not significantly adversely affect the specific growth rate of *S. clavuligerus SC6*. The specific growth rates for all three strains were calculated from the data in Figure 3.4. While pSET152 has a negligible impact on the specific growth rate, pMS82 reduces the specific growth rate by 1.502 x 10^{-3}.

Strain	Specific growth rate ($hour^{-1}$)	% change compared to 'wild type'
S. clavuligerus SC6	8.539×10^{-3}	Wild type
S. clavuligerus SC6 pMS82	7.037×10^{-3}	-17.6%
S. clavuligerus SC6 pSET152	8.643×10^{-3}	+1.22%

To further test if empty vector pMS82 or pSET152 affect *S. clavuligerus* SC2 or SC6 within fermentation, CA assays were carried out on *S. clavuligerus* SC2 and SC6 with pSET152 and pMS82, to assess if either vector affected secondary metabolite production. All strains were grown for 48 hours in 50 mL TSB and triplicate samples of supernatant were taken and CA yields were determined. Results were normalised to cell dry weight (Figure 3.5). The average CA titre produced by *S. clavuligerus* SC2 and SC6 was 41.3 and 14.8 µg / mg cell dry weight respectively. These values differ from those previously reported (Birke, 2020) in which, after 24 hours growth in TSB clavulanic acid yield in *S. clavuligerus* SC2 and SC6 were 45 and 112 µg / mg cell dry weight respectively. In this study CA yield then dropped to 3 and 6 µg / mg cell dry weight respectively after 65 hours.

In this study, the lower yield observed for *S. clavuligerus* SC6 compared to SC2 is expected, as SC6 has been adapted for enhanced CA production in industrial media. In contrast, SC2 performs better in TSB, likely due to its closer genetic similarity to the parental strain DSM738, which also grows more effectively in TSB than in CM5. Furthermore, CA production is more consistent among the three SC6 strains than among the SC2 strains.

An ANOVA test was carried out on clavulanic acid yields finding that there was no difference between the CA production of *S. clavuligerus* SC6 and its respective ex-conjugants. The p value was 0.616 and 0.980 when comparing the means of *S. clavuligerus* SC6 to *S. clavuligerus* SC6 pMS82 and pSET152 respectively. However, there was a significant difference between the CA production of *S. clavuligerus* SC2 and *S. clavuligerus* SC2 pMS82 and pSET152. The p value was 1.35×10^{-5} and 1.24×10^{-3} when comparing the means of *S. clavuligerus* SC2 to *S. clavuligerus* SC2 pMS82 and pSET152 respectively. This correlates to a 1.91 and 1.27 fold decrease in production for *S. clavuligerus* SC2 pMS82 and pSET152 respectively compared to 'wild type'.

As a result, the variability of CA yields within *S. clavuligerus* SC2 pMS82 and pSET152 meant pSET152 was chosen to integrate the barcode sequences into *S. clavuligerus* due to the fact that pSET152 had a smaller negative effect on the CA yield of SC2 and a smaller negative effect on SGR.

Additionally, while integration sites for both pMS82 and pSET152 are well documented, pSET152 carries apramycin resistance, whereas pMS82 carries hygromycin resistance. Hygromycin is salt sensitive (Yoshikawa et al., 1995) which may pose problem working with L3M9 agar which contains salt. This further gave reasoning to pSET152 being taken forward for use within this study.

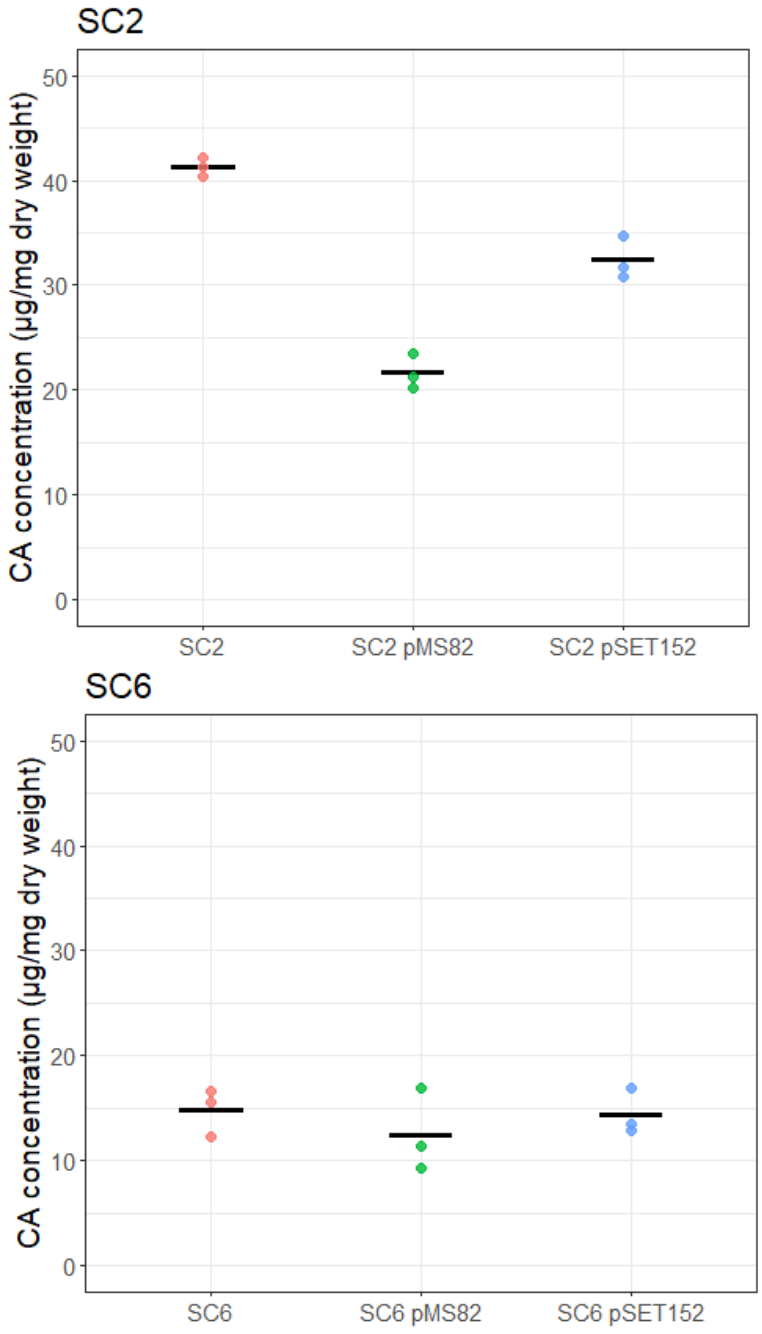


Figure 3.5. Analysis of the effect of pMS82 and pSET152 on the titre of clavulanic acid produced by *S. clavuligerus* SC2 *and S. clavuligerus* SC6. *S. clavuligerus* SC6, *S. clavuligerus* SC6 pMS82 *and S. clavuligerus* SC6 pSET152 were grown in TSB for 48 hours and an imidazole assay was performed on the supernatant to determine the concentration of clavulanic acid. Data was normalised to the cell dry weight. The p value for an ANOVA test between clavulanic acid concentrations from SC2 and SC6 are 1.75 x 10^{-5} and 0.610 respectively.

3.4 Further validation of *S. clavuligerus* SC6 pSET152

After selection of pSET152 to integrate barcode sequences into the chromosome of *S. clavuligerus*, further validation was carried out on ex-conjugants of *S. clavuligerus* SC6 pSET152, as the dominant strain the be used in this study, to assess the vectors effects on colony morphology. Observing how the vector affects colony morphology and cell life cycles was vital, as these factors are closely linked to *Streptomyces* antibiotic production.

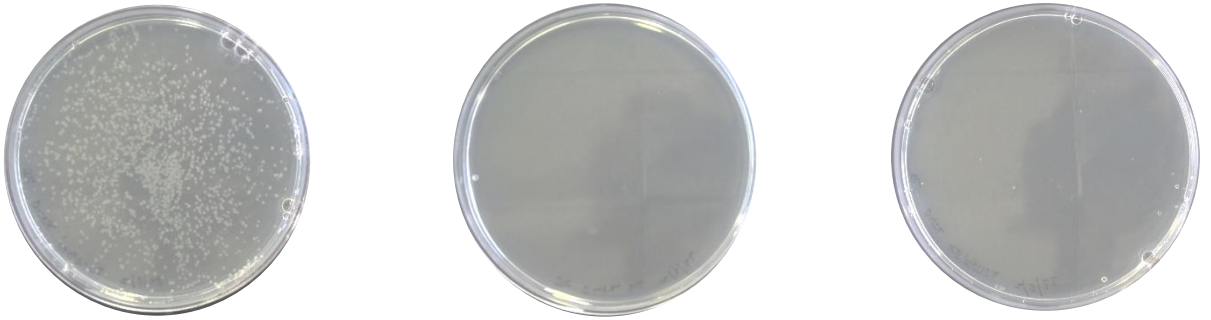
Colonies of *S. clavuligerus* SC6 pSET152 were grown directly from a spore stock to obtain single colonies on L3M9 agar. Differences in morphology within a single strain were observed (Figure 3.6). Multiple morphologies were seen over two plates ranging from white colonies (Figure 3.6 A), non-sporulating colonies, 'sick' looking colonies that produce no arial hyphae and appear in a halo shape (Figure 3.6 C) and sporulating, green colonies as per a 'wild type' morphology (Figure 3.6 B). Despite differences being observed between colonies of a single stock it was determined these levels of variation are normal within *S. clavuligerus* (Ríos-Fernández et al., 2024). Additionally, these results allow us to observe sporulation and hyphal growth with *S. clavuligerus* SC6 pSET152 and as such confirm life cycle progression, confirming the suitability of the vector for use within this study.

.

Figure 3.6. Colony morphology of *S. clavuligerus* SC6 pSET152. Two plates of *S. clavuligerus* SC6 pSET152 were grown directly from a single 20% glycerol spore stock. Multiple morphologies can be seen over the two plates including white colonies producing arial hyphae (colony A), green colonies producing arial hyphae (colony B) and 'sick' looking colonies that produce some green pigment but do not appear to have arial hyphae, with smaller satellite colonies around them (colony C).

As the stable integration of pSET152 into the *attB* site is essential for the reliable detection of barcoded strains the permanence of integration of the plasmid into the chromosome was investigated. If the plasmid excised itself from the chromosome, it would also remove the barcode from the chromosome, stopping us from accurately tracking a barcoded stains relative abundance.

To test the stability of integration of pSET152, genomic DNA of *S. clavuligerus* SC6 pSET152 was recovered from a single ex-conjugant via genomic DNA extraction. This recovered DNA was then transformed directly into DH5α via heat shock. If pSET152 DNA existed outside of the chromosome it would be able to transform DH5α. These transformants were plated and grown on LB agar plates containing apramycin. Similarly, *E. coli* DH5α was transformed with pSET152 DNA and plated on apramycin LB plates as well as *E. coli* DH5α. The plates with *E. coli* DH5α and *E. coli* DH5α transformed with the 'recovered' pSET152 DNA did not produce any colonies indicating that free pSET152 was not recovered from *S. clavuligerus* SC6 pSET152. As such it was concluded pSET152 was stably integrated into the chromosome and unlikely to be in its circular form outside of the chromosome (Figure 3.7).



DH5α + pSET152

DH5α

DH5α + ‘recovered’ pSET152

Figure 3.7. Growth of *E. coli* DH5α with ‘recovered’ pSET152 DNA from *S. clavuligerus*. *E. coli* DH5α is apramycin sensitive and exhibits resistance when transformed with pSET152. DNA extracted from *S. clavuligerus* SC6 pSET152 ex-conjugants does not yield *E. coli* DH5α transformants.

3.5 Selection of barcodes to use within this study

Competing strains have previously been tracked using barcode sequencing within *Candida albicans* (Larcombe et al., 2023). However, within the work by Larcombe et al. several barcodes were seen to be artificially overrepresented within the data (Larcombe, personal communication). As a result, barcoded strains were validated to ensure their reliability within this work. Firstly, 20 barcode sequences from the study by Larcombe et al. were taken for use in the current study (Table 3.2). Barcodes were randomly assigned the names JCb1-20 (James Croxford barcode 1-20). Barcode sequences contain a unique 25 bp middle region, flanked by identical 17 and 18 bp sequence before and after the barcode respectively.

Table 3.2. Unique barcodes used within the study. Barcodes were named JCb1-20. The unique region of each barcode is in bold. The identical flanking regions can be seen adjacent to each barcode.

Barcode	Sequence
JCb1	CGGTGTCGGTCTCGTAG**TACAATAAAGGGCAGGTGCAACCAT**AGAGACCTCGTGGACATC
JCb2	CGGTGTCGGTCTCGTAG**TGACCAAGCCTGAATAGCGTCATAA**AGAGACCTCGTGGACATC
JCb3	CGGTGTCGGTCTCGTAG**CGTGTATTAGAGTAATCGCATCTAG**AGAGACCTCGTGGACATC
JCb4	CGGTGTCGGTCTCGTAG**TAGCGTTAGCGGCCACGGAACATCA**AGAGACCTCGTGGACATC
JCb5	CGGTGTCGGTCTCGTAG**TCCTCTCGGGAGGCCAATAGAAACG**AGAGACCTCGTGGACATC
JCb6	CGGTGTCGGTCTCGTAG**TACGCTTGACCCGTGTAGCTGTCAG**AGAGACCTCGTGGACATC
JCb7	CGGTGTCGGTCTCGTAG**TGGTTAACACGGAACCTGAGATCCC**AGAGACCTCGTGGACATC
JCb8	CGGTGTCGGTCTCGTAG**TAGCACACCTAGCGTGAACCGAGCC**AGAGACCTCGTGGACATC
JCb9	CGGTGTCGGTCTCGTAG**TTACCCATTAAGAGCTCAGCGCAAC**AGAGACCTCGTGGACATC
JCb10	CGGTGTCGGTCTCGTAG**TACAGGGAAGACCTACCCGATTCCC**AGAGACCTCGTGGACATC
JCb11	CGGTGTCGGTCTCGTAG**TAAGTCAAGTGTCCGGTTCTCCCCG**AGAGACCTCGTGGACATC
JCb12	CGGTGTCGGTCTCGTAG**TCCGTGTTGTACGACGTTAGACAAC**AGAGACCTCGTGGACATC
JCb13	CGGTGTCGGTCTCGTAG**TAGTCTACCTCGCATGGCTAAGCGG**AGAGACCTCGTGGACATC
JCb14	CGGTGTCGGTCTCGTAG**TATCCAGTCGGCGATAAGTATGCAA**AGAGACCTCGTGGACATC
JCb15	CGGTGTCGGTCTCGTAG**TCACAGGGACAGGTCATATCGTTCT**AGAGACCTCGTGGACATC
JCb16	CGGTGTCGGTCTCGTAG**TGACCCCGTAGTGCCGTTAAGTATC**AGAGACCTCGTGGACATC
JCb17	CGGTGTCGGTCTCGTAG**TCAATAAGCGAGCCGTAAGTTGATC**AGAGACCTCGTGGACATC
JCb18	CGGTGTCGGTCTCGTAG**TCAAAGATGCCCCGACTAGGTAATA**AGAGACCTCGTGGACATC
JCb19	CGGTGTCGGTCTCGTAG**TCTGCACATCCAGGCGACCGAATAT**AGAGACCTCGTGGACATC
JCb20	CGGTGTCGGTCTCGTAG**TCGACGATCATCGGTGAAGCGTTTG**AGAGACCTCGTGGACATC

Analysis of the barcode sequences was carried out to ensure that correct barcode identification could be achieved from sequencing data. To assess if the 20 barcodes were sufficiently different from each other for accurate barcode detection, Clustal Omega (Sievers and Higgins, 2014) was used to align the whole barcode sequence to assess their similarity, with the alignments were visualised using Jalview (Figure 3.8). The position with the highest sequence similarity was base position 29 where 55% of all barcodes contained a guanine. For all remaining positions, less than half of all barcodes shared a base.

To further check for sequence similarity between barcodes the unique 25 bp section of each barcode was aligned against each other using the pairwise2 module form Biopython. Both forward and reverse complement orientations of each sequence were tested, as both strands may be present during sequencing. A local alignment was performed for each pair using a match score of +2, mismatch penalty of –1, gap open penalty of –2, and gap extension penalty of –0.5. A similarity threshold was defined as a minimum alignment length of 25 bp and ≥70% identity.

Out of 190 pairwise comparisons, only one met these criteria: JCb4 and JCb16 showed a 27 bp alignment with 66.7% identity on opposite strands (Table 3.3). Although the match length slightly exceeds the unique 25 bp barcode region, the proportion of identical bases remains below the 70% identity threshold. Based on this result, all 20 barcodes were considered sufficiently distinct for confident identification in sequencing data, and all were retained for downstream use.

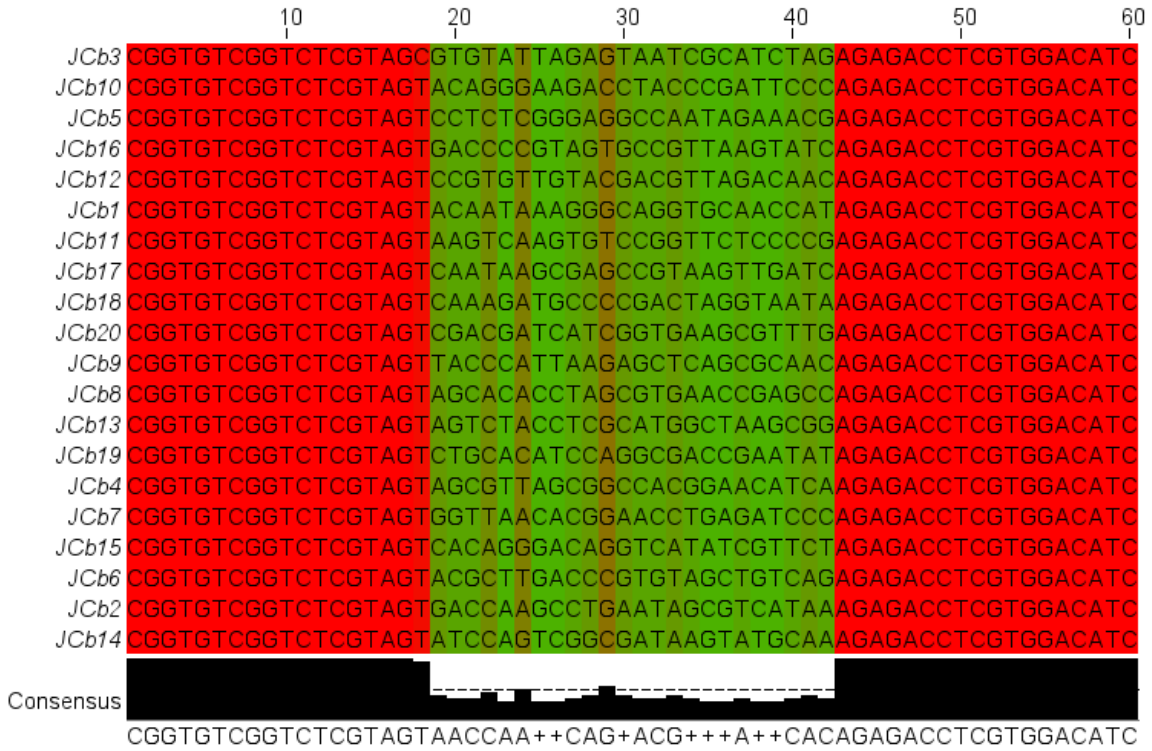


Figure 3.8. Alignment of 20 barcodes used in this study including flanking regions shows low consensus between sequences. The 20 barcodes used within this study were aligned using ClustalOmega and visualised on JalView. Red shows high consensus and green shows low consensus. The identical flanking regions can be seen in red. Barcodes are in order of best alignment.

Table 3.3. Two barcodes show significant levels of identity. The unique region of all 20 barcodes were aligned against each other using pairwise2. One combination of barcodes was found to be similar.

Similar barcodes	Identity (bp)	Match length (bp)	Identity (%)	Match % of barcode
4 and 16	18	27	66.7	108

3.6 Synthesis and conjugation of barcoded strains

After validation of pSET152 for use with *S. clavuligerus* SC6, the fully cloned vectors containing the 20 barcode sequences were synthesised by GenScript (Oxford, United Kingdom) which were inserted into the *lacZ* gene (Figure 3.9). *E. coli* ET12567/pUZ8002 was transformed with each of these plasmids which was then conjugated into *S. clavuligerus* SC6. An isogenic stock of *S. clavuligerus* SC6 was used to conjugate into to ensure the barcoded strains of *S. clavuligerus* SC6 were identical except for their barcodes.

Genomic DNA was extracted from each barcoded strain after growth in 50 mL TSB for 48 hours. The primers JC pSET152 barseq Fwd and JC pSET152 barseq Rev were used to amplify the barcoded region from each strains genomic DNA (Figure 3.10). After amplification, PCR products were preprepared for sequencing using the kit SQK-LSK114 sequenced on the Flongle flowcell (FLG114) using a MinION mk1B. The resulting POD5 files were base called, demultiplexed and the code available at https://github.com/jamescroxford/Bar-seq was used to detect barcodes within the FASTQ files for each strain. It was determined that the genomic DNA of each barcoded strain contained 100% of the expected barcode demonstrating the 'purity' of each stock.

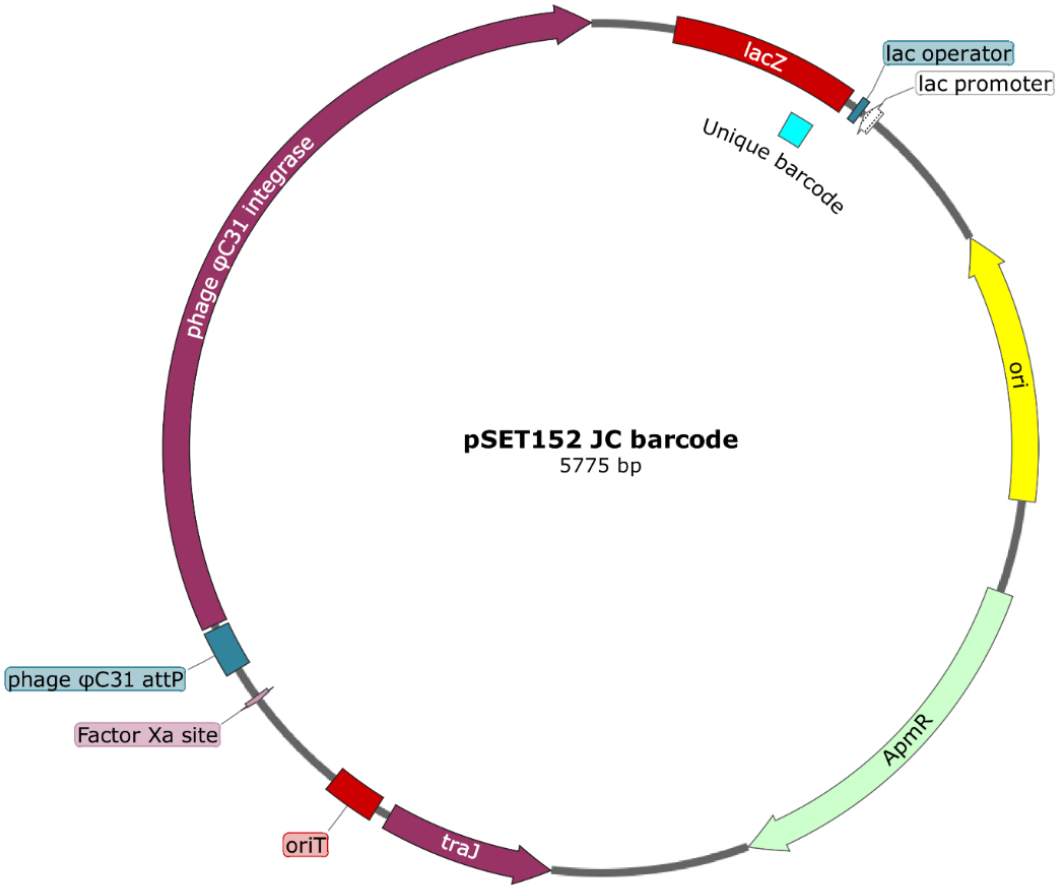


Figure 3.9. Plasmid map of pSET152 carrying the unique barcodes. The unique barcodes are cloned into the lacZ gene within the vector (cyan).

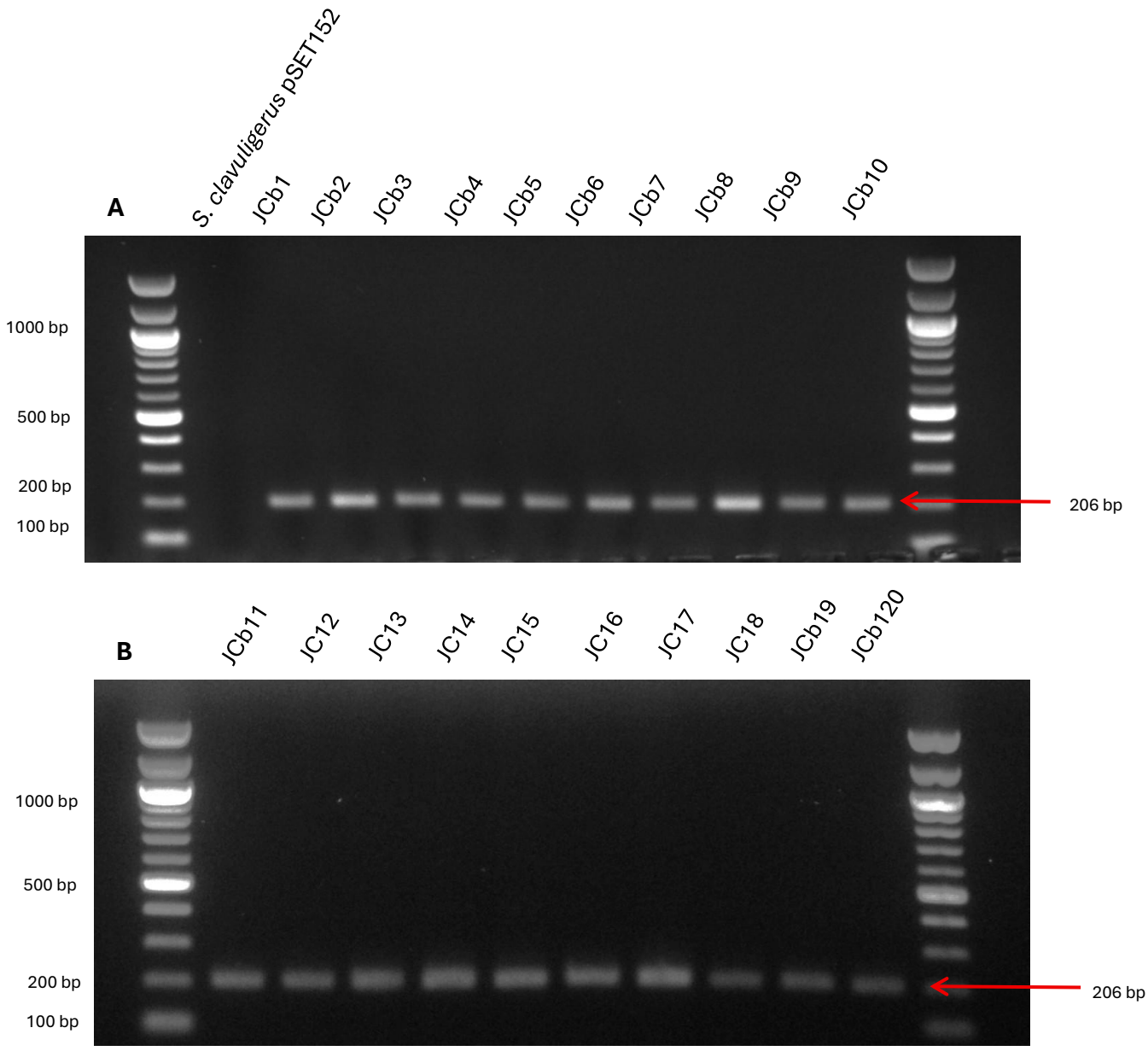


Figure 3.10. Amplification of barcodes from genomic DNA of *S. clavuligerus* SC6 pSET152 JCb1–20. Genomic DNA was extracted from 20 barcoded strains of *S. clavuligerus* SC6 pSET152 JCb1–20 and barcode regions were amplified using primers JC pSET152 Barseq Fwd and JC pSET152 Barseq Rev. **(A)** PCR products from strains JCb1–10 are shown. **(B)** PCR products from strains JCb11–20 are shown. All reactions yielded a band of the expected size, consistent with successful amplification of barcodes from genomic DNA. Q5 polymerase was used with an annealing temperature of 52 °C for the PCR reaction (Table 2.7).

Although pSET152 has previously been shown not to affect the growth of *S. clavuligerus* SC6 (Table 3.1), further validation was done to ensure that none of the unique barcodes within the synthesised vectors affected the specific growth rates of *S. clavuligerus* SC6. To test this, each of the 20 strains of *S. clavuligerus* SC6 pSET152 JCb1-20 was grown within 50 mL TSB for 48 hours with growth monitored on the cell growth quantifier. Exponential growth was determined from this data and specific growth rate was calculated (Figure 3.11). The average specific growth rate was 7.78×10^{-3}, consistent with previous result of 8.643×10^{-3} as the specific growth rate of *S. clavuligerus* SC6 pSET152. A Grubbs test was performed on the specific growth rate to detect if one barcoded sequence affected the growth rate of *S. clavuligerus* SC6 more than any other. This returned a p value of 0.423 indicating there are no outliers within the data. The statistical inference of this experiment should however be treated cautiously as this experiment was performed once (single biological replicate) and is presented as a validation/feasibility check.

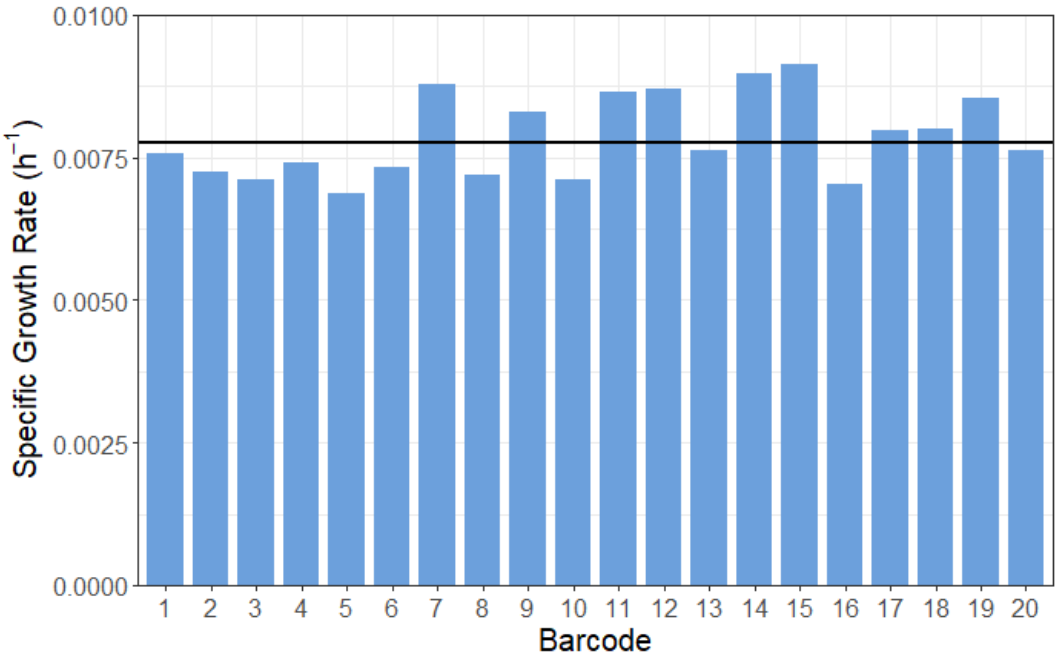


Figure 3.11. Unique barcodes do not significantly affect the specific growth rate. Each of the 20 barcoded strains of *S. clavuligerus* SC6 pSET152 JCb1–20 was grown in 50 mL TSB for 48 hours, and growth was monitored using the cell growth quantifier. Specific growth rates were calculated from the exponential growth phase. The average specific growth rate across all strains was 7.78×10^{-3} h^{-1} (black horizontal line). A Grubbs test ($p = 0.4226$) detected no statistical outliers, indicating that no individual barcode significantly altered strain growth. Only a single replicate was performed for each sample and therefore inference should be treated cautiously.

3.7 Validation of barcode identification once sequenced

After the 20 barcoded strains were confirmed as being appropriate for use within this study, validation of identification of the unique barcode after sequencing was carried out. Firstly, the ability of the code to correctly identify barcode sequences was tested. All strains of *S. clavuligerus* SC6 pSET152 JCb1-20 were grown from stocks, and the genomic DNA was extracted. The barcoded region of each strain was amplified from the genomic DNA using primers JC pSET152 barseq Fwd and JC pSET152 barseq Rev. After amplification PCR products were sequenced on the Flongle flowcell (FLG114) using a MinION mk1B. The resulting POD5 files were base called, demultiplexed and code available at https://github.com/jamescroxford/Bar-seq was used to detect barcodes within the FASTQ files for each strain.

It is important to consider the possibly that the barcode sequence itself may mutate or be incorrectly base-called, therefore losing the ability of the barcode to be identified within the reads. Natural mutation rates have been observed to be 2.84×10^{-8} per nucleotide per cycle (Dagva et al., 2024). Additionally, while Oxford Nanopore Technologies claim a read accuracy up to 99.99% for chemistry 10, this value is more likely to be closer to 98% when using amplicon sequencing due to read instability at the start and end of amplicons (Luo et al., 2022).

Levenshtein distances (a string metric for measuring the difference between two sequences) are used within the code to allow for detection of barcode sequences with potential mutations occurring within the barcode (Berger et al., 2021). However, it is crucial an increased Levenshtein distance does not lead to a drop in specificity of barcode identification. As a result, lines 52 and 59 of the code were used to sequentially increase the Levenshtein distance from 0 to 6 using the max_l_dist function for the forward and reverse complements of the barcode sequences respectively. This process

was carried out for all strains. Levenshtein distance was used to detect barcodes from the previously sequenced single stock of each strain. The results of *S. clavuligerus* SC6 pSET152 JCb2 was used as an example (Figure 3.12). At a Levenshtein distance of zero to four, 100% of identified barcodes were assigned to JCb2, as expected. This demonstrates a robust ability of the code to correctly identify barcode sequences. At a Levenshtein distance of five 99.67% of identified barcodes were correctly assigned to JCb2 but this drops to 71.09% at a Levenshtein distance of six. A Levenshtein distance of one is considered enough to account for mutations within barcodes in future experiments. These results demonstrate that barcodes assignment will be 100% specific under these conditions. Results were similar for all other barcodes.

Barcode abundances were also assessed in a short fermentation time course, confirming that using a Levenshtein distance of 0 or 1 yielded identical results. As these data are part of the core fermentation findings, they are shown later (Figure 3.16) and are not repeated here.

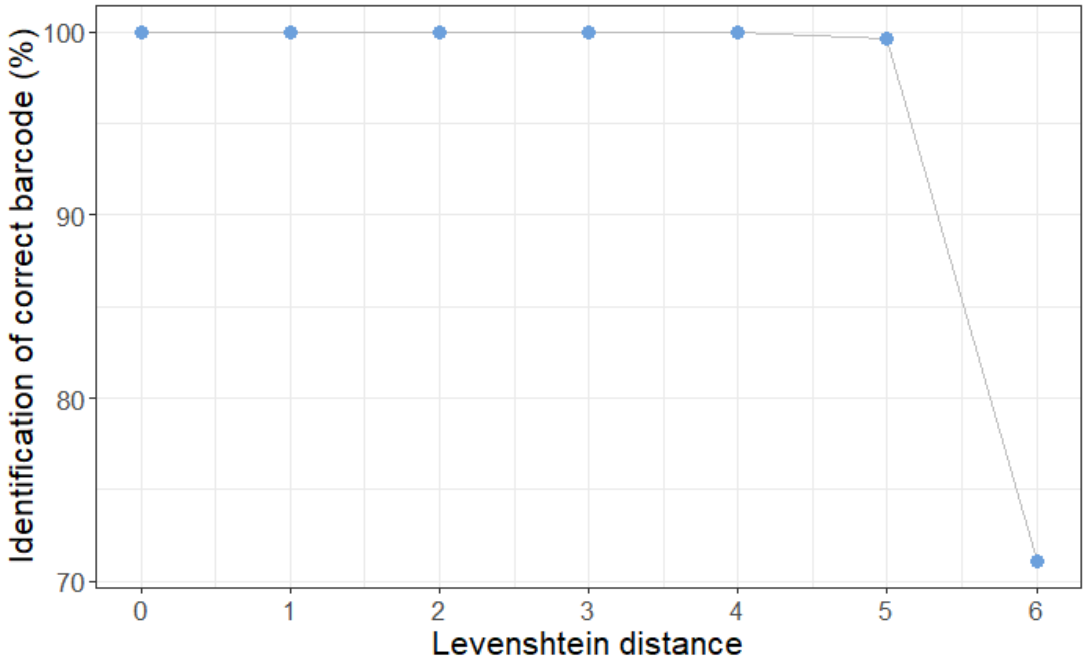


Figure 3.12. Accuracy of identification of correct barcode as a function of Levenshtein distance for JCb2. At distances up to 4, the identification accuracy remains 100%, indicating high robustness in correct barcode assignment. At a distance of 5 the identification accuracy drops to 99.6%. However, a notable decline in accuracy is observed at distance 6, suggesting a threshold for reliable barcode identification must be below this value. An example subset of reads from one barcoded strain was used with the fuzzysearch python package for calculating identification at different distances.

Once correct identification of barcodes by the code had been validated, the effect of read depth on the consistency of barcode relative abundance estimates was investigated. Sequencing data from several fermentations containing all 20 barcoded strains, presented later in this chapter and in Chapter 4 (Figures 3.16, 4.1, and 4.2), were used for this analysis. Various read depths were tested by subsetting FASTQ files using the seqtk sample function (Figure 3.13). Only timepoints with over 35,000 reads were included for subsetting. It was found that higher read depths led to more accurate barcode abundance predictions, with estimates converging more closely to the original values. For example, at a read depth of 100 per timepoint, barcode abundance prediction ranged from 0 to 198% of the original relative abundance predictions. At a read depth of 2,500 the barcode prediction accuracy compared to the original abundance prediction ranged from 79 to 123%, at a depth of 7,500 the range was 93 to 112%. However as read depth increased from 7,500, only marginal gains were seen in the closeness of the barcode abundance prediction to it's original of 100%. At 10,000 reads per timepoint the average deviation from 100% abundance prediction was 6.5% whereas at 20,000 reads that value was 4.5%. However, at 22,500 reads the value increased to 5.2% showing accuracy of abundance prediction increases up to but not beyond 20,000 reads per timepoint, using 20 barcodes. This data suggests that beyond 10,000 reads per timepoint gains in relative abundance prediction accuracy are negligible. As a result, it was determined that a minimum of 10,000 reads were needed per timepoint going forward, any sequencing runs yielding fewer reads than this would either be repeated of discarded.

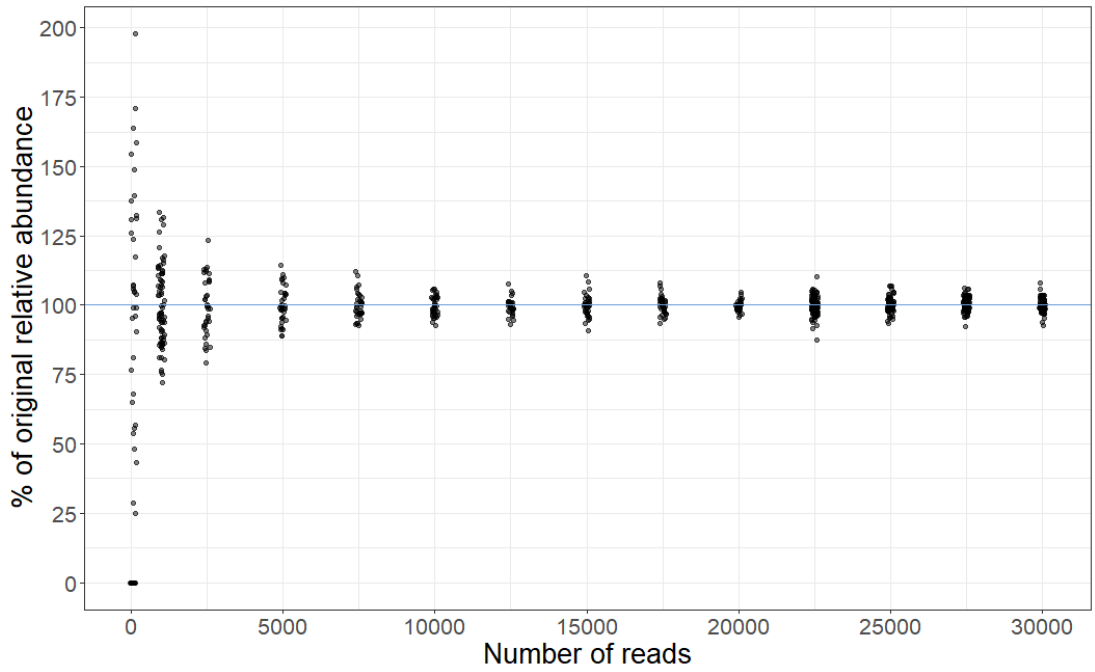


Figure 3.13. Variation in correct barcode assignment at various read depths. Reads from different fermentation time points were subsetted to estimate the relative abundance of each barcoded strain. The values presented represent each barcode's estimated abundance as a percentage of its original relative abundance in the complete dataset. All time points used originally had over 35,000 reads. The blue line indicates the original abundance estimate using all reads for each time point. Accuracy improves up to approximately 10,000 reads, beyond which no further increase in accuracy is observed.

Having confirmed that sequencing conditions for correct barcode identification and read depth could be met, validation of whether barcodes could be detected at an expected relative abundance was carried out. All 20 strains of *S. clavuligerus* SC6 pSET152 JCb1-20 were grown, genomic DNA was extracted and barcoded regions were amplified. A qubit fluorometer was used to quantify the DNA concentration of each barcode after amplification. Equal DNA concentrations of all barcodes were mixed totalling 1µg DNA (50 ng of each barcode). This pool of barcodes was then sequenced, basecalled and the relative abundance of each barcode calculated. If mixed together and sequenced in equal proportions, we would expect the relative abundance of each barcode to be exactly 5%. The average relative abundance of all 20 barcodes is 5.00% (Figure 3.14). The smallest percentage abundance was 3.03% (JCb11) and the greatest relative abundance was 7.62% (JCb6). A Wilcoxon ranked test performed of this data has a P value of 0.867, demostrating little variation from this mean of 5.00. The statistaical inference of this experiment should however be treated cautiously as this experiment was performed once (single biological replicate) and is presented as a validation/feasibility check.

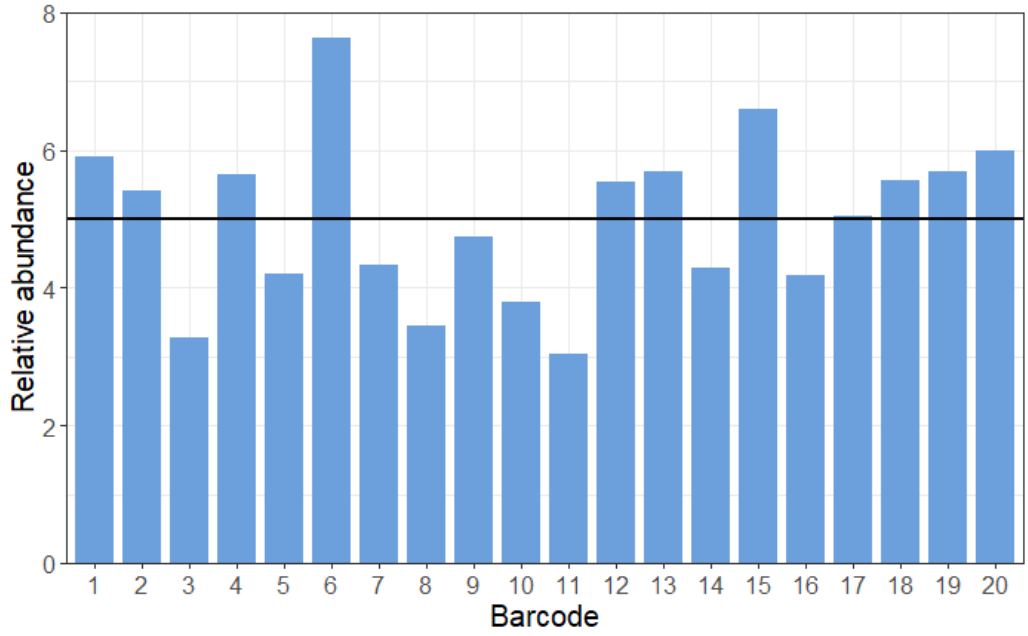


Figure 3.14. Relative abundance of barcoded amplicons in a mixed DNA sample. Stocks of *Streptomyces clavuligerus* SC6 pSET152 JCb1-20 were cultured, and barcoded amplicons were generated using the primers JC pSET152 barseq Fwd and Rev. The amplicons were mixed in equal concentrations, sequenced, and barcode abundance calculated. The relative abundance of each barcode is shown, with the target abundance set at 5% (indicated by the horizontal black line). While the mean relative abundance across all barcodes was 5.00%, variations are observed. A Wilcoxon ranked test performed of this data has a p value of 0.867, demostrating little variation from the mean of 5. Only a single replicate was performed for each sample and therefore inference should be treated cautiously.

Once it had been determined that amplicons of barcode DNA could be mixed and sequenced in equal proportion, we sought to ensure that we could also mix strains of *S. clavuligerus* SC6 pSET152 JCb1-20 in equal proportions which would then be reflected by sequencing. To achieve this, stocks of barcoded strains were grown individually for 48 hours in 50 mL TSB. After 40 hours the OD_{600} of each stock was recorded, and strains were mixed in equal proportion based on these OD_{600} measurements making a total pooled mixture of 9.23 mL. This pooled mixture was taken, and the genomic DNA was extracted. Primers JC pSET152 barseq Fwd and JC pSET152 barseq Rev were used to amplify out the barcoded region from the population. The resulting amplicons were sequenced, and the relative abundance of each barcode strain was calculated (Figure 3.14).

The average relative abundance of all barcodes was 5.00%. A Wilcoxon ranked test performed of this data to test if any barcode deviated from the expected relative abundance of 5%. This returned a P value of 0.674, demonstrating little variation. The statistical inference of this experiment should however be treated cautiously as this experiment was performed once (single biological replicate) and is presented as a validation/feasibility check. The barcode with the lowest relative abundance was JCb10 at 2.98% and the barcode with the highest relative abundance was JCb5 with 8.41% of the total detected barcodes.

Despite the minimal variability these results suggest we can mix barcodes strains in equal proportions to make an inoculum for a fermentation. These results also demonstrate that one barcode does not seem to be preferentially sequenced (in agreement with the data shown with Figure 3.15).

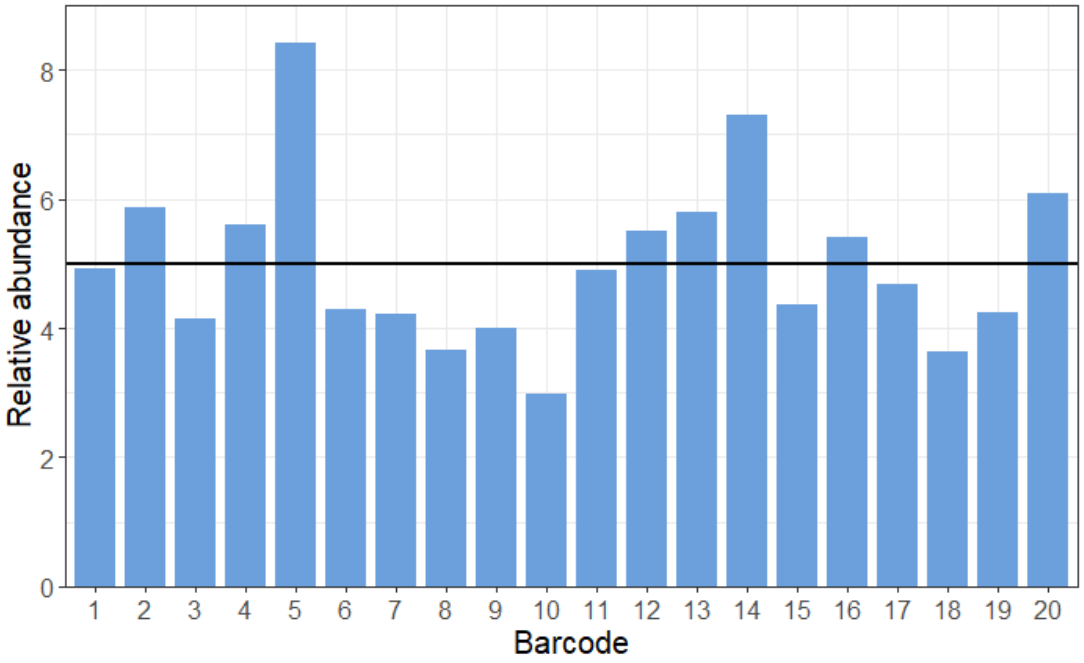


Figure 3.15. Relative abundance of barcodes in a sample of barcoded *Streptomyces clavuligerus* SC6 strains mixed by OD_{600}. Stocks of *Streptomyces clavuligerus* SC6 pSET152 JCb1-20 were grown for 48 hours and mixed in equal proportions based on OD_{600} measurements. Genomic DNA was extracted, and amplicons were generated using the primers JC pSET152 barseq Fwd and Rev. The amplicons were then sequenced, and barcode abundance was calculated. The relative abundance of each barcode is shown, with the target abundance set at 5% (indicated by the horizontal black line). While the mean relative abundance across all barcodes was 5.00%, variations are observed. A Wilcoxon ranked test performed of this data has a p value of 0.674, demostrating little variation from the mean of 5. Only a single replicate was performed for each sample and therefore inference should be treated cautiously.

After demonstrating that barcodes could be detected in equal abundance when mixing both DNA and barcoded strains (Figures 3.12 and 3.13), further validation to demonstrate PCR bias does not influence barcode relative abundance prediction was performed. This was done as two PCR stages are involved in barcode relative abundance estimation (firstly in the amplification of the barcoded region from genomic DNA and secondly as part of the library preparation using the SQK-LSK114 sequencing kit). PCR bias is a well-documented phenomenon and may impact the accuracy of predicted barcoded strain abundances due to preferential amplification of certain barcodes over others (Gohl et al., 2016; Polz and Cavanaugh, 1998).

Genomic DNA was extracted in triplicate of a test fermentation in 50 mL TSB (not shown) at a random timepoint. The average relative abundance of each barcoded strain across the triplicates was then plotted (Figure 3.16 A). The deviation of replicates from the mean value for that barcode was also plotted as a bar and whisker plot (Figure 3.16 B). The average deviation from the mean predicted abundance over the 20 barcoded strains was 0.19% relative abundance. The largest average deviation from a mean predicted abundance was JCb13 deviating from the mean at an average of 0.46%. Across all barcoded strains triplicates deviated both above and below the mean broadly equally, meaning there was not a tendency to over or under predict the relative abundance of barcoded strains according to PCR bias. It was determined PCR bias does not affect the abundance prediction of barcoded strains.

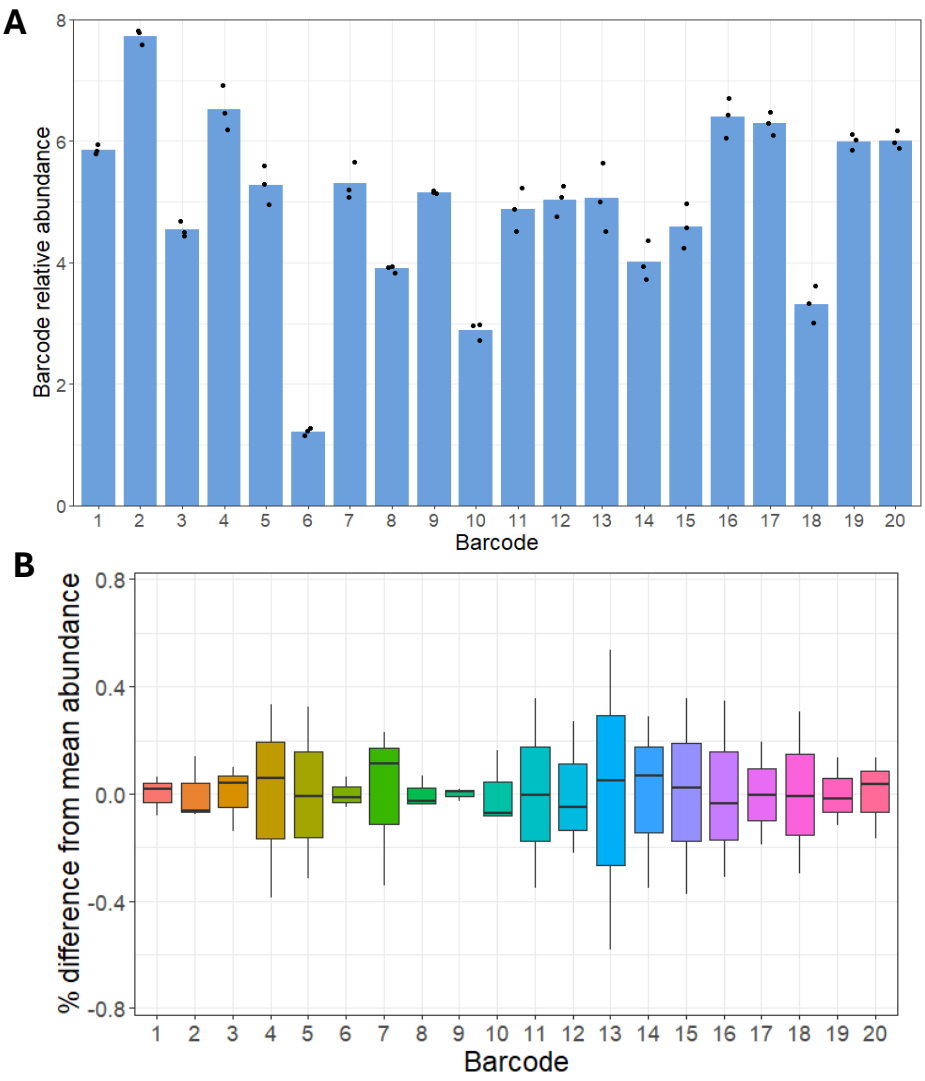


Figure 3.16. PCR bias does not impact barcode relative abundance during Nanopore library preparation. Genomic DNA was extracted from a fermentation timepoint containing the 20 barcoded strains of *Streptomyces clavuligerus* SC6 pSET152 JCb1-20. Amplicons were generated using barcode forward and reverse and subsequently used for library preparation and sequencing, all conducted in triplicate. **(A)** Average relative abundance across triplicates, with individual values for each replicate shown as black dots. **(B)** A box and whisker plot illustrating the deviation of each replicate's calculated barcode abundance from the mean abundance. The average deviation from the mean across all 60 data points is 1.22×10^{-16}, indicating minimal PCR bias in barcode representation despite two rounds of PCR in the generation of amplicons and the sequencing library preparation. The average rage between the highest and lowest vale across the 20 barcodes is 0.447%.

After validating the use of the code for barcode detection, demonstrating that barcodes could be mixed in equal abundance and detected at expected rates and analysing the effects of PCR bias, a 50 mL fermentation in TSB of the barcoded strains was carried out to demonstrate that Bar-seq can be used to track the relative abundance of barcoded strains within a fermentation. Barcoded strains of *Streptomyces clavuligerus* SC6 pSET152 JCb1-20 were cultured for 48 hours and mixed in equal proportions based on OD_{600} measurements, as previously demonstrated. This mixture was used to inoculate 50 mL TSB and grown for 48 hours. In total approximately 5×10^7 CFUs were added to this fermentation as a starting inoculum. Strain JCb6 was not included in this study as the original 48-hour culture was contaminated. Genomic DNA was extracted at the start and end of the fermentation, and barcodes were amplified using the primers JC pSET152 barseq Fwd and Rev. Amplicons were sequenced, and the relative abundance of each barcode was calculated for both time points. A muller plot was created to show the difference in strain abundance over time (Figure 3.17 A). The ability to track barcode relative abundance throughout a fermentation demonstrates that Bar-seq can explore population dynamics within a fermentation. These results also show that this 50 mL fermentation appears to be stable, with no barcoded strain becoming dominant within the fermentation.

To evaluate the changes in the relative abundance of barcoded strains and their differences at the start and endpoints during fermentation, the $\log_2$ fold change was used to quantify these differences (Figure 3.17 B).

Figure 3.15 A and B both show that strains can be barcode relative abundance, and therefore the use of Bar-seq can be taken forward into several different fermentation environments.

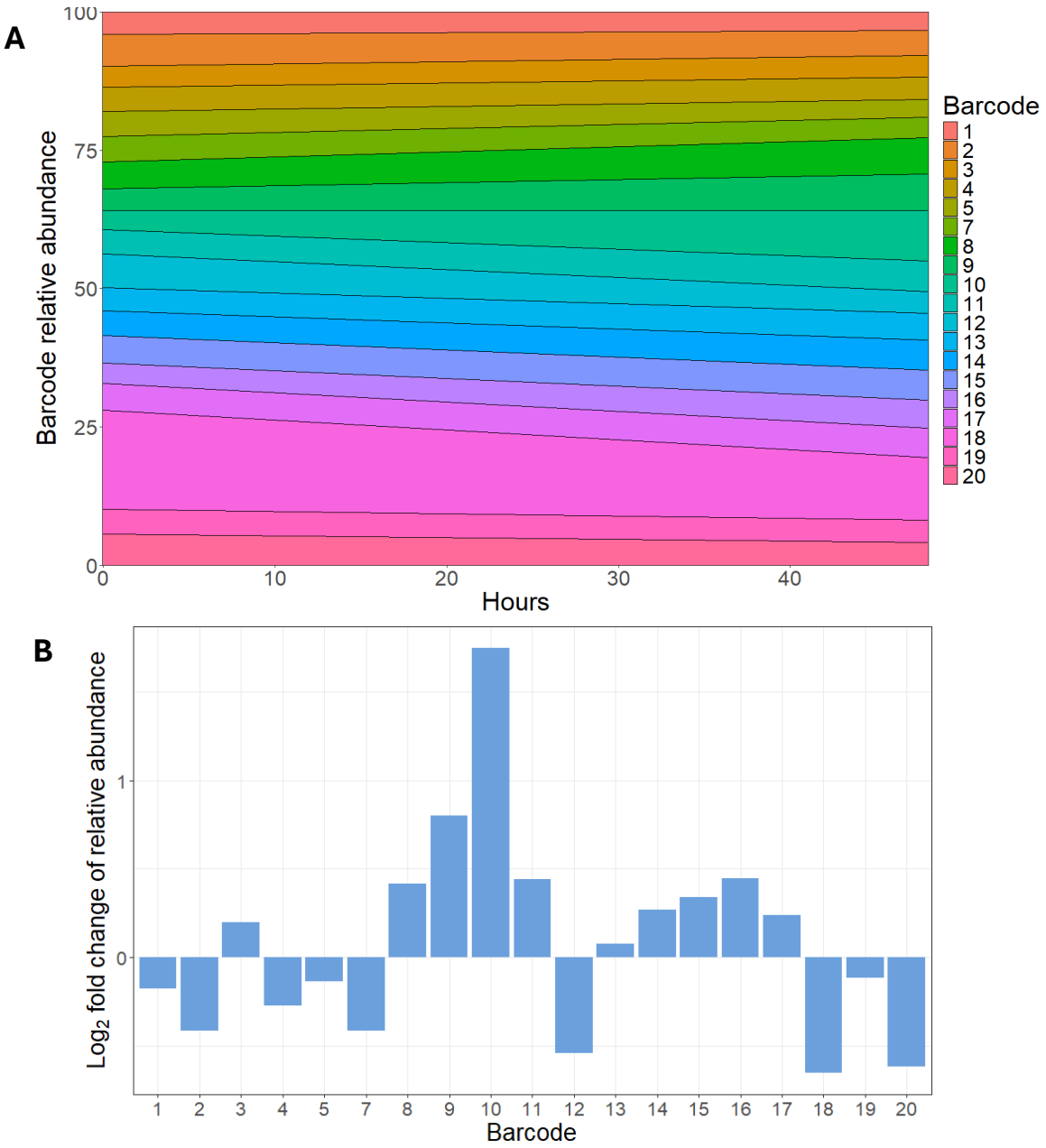


Figure 3.17. Bar-seq can be used to track barcodes within a fermentation and determine the change in their relative abundance. (A) The barcoded strains of *Streptomyces clavuligerus* SC6 pSET152 JCb1-20 were cultured for 48 hours and mixed in equal proportions based on OD_{600} measurements. This mixture was used to inoculate 50 mL TSB and grown for 48 hours. Genomic DNA was extracted at the start and end of the fermentation, and barcodes were amplified using the primers barcode forward and reverse. Amplicons were sequenced, and the relative abundance of each barcode was calculated for both time points. **(B)** The log_2 fold change in each barcode's relative abundance between the two time points was calculated and plotted, demonstrating that Bar-seq can effectively track changes in barcode representation over time and facilitate time-point comparisons.

3.8 Summary

In this chapter Bar-seq methodology was proposed and validated for the purpose of studying population dynamics within industrial *S. clavuligerus* during a fermentation. Barcoded strains were mixed equally, fermented, DNA extracted, the barcoded were PCR amplified, sequenced, and quantified to assess relative abundance (Figure 3.17).

Using the barcode sequences previously used by Larcombe et al., 2023, the similarity between sequences was explored to ensure correct barcode identification would be possible. ClustalOmega and Jalview (Figure 3.8) alignments showed low consensus, confirming sequence uniqueness. Pairwise2 local alignments detected only one instance of moderate similarity between barcode sequences (JCb4 and JCb16), which did not exceed defined similarity thresholds (Table 3.3).

After the selection of barcode sequencing the plasmids pMS82 and pSET152 were explored for their use of inserting the barcode sequences into *S. clavuligerus* SC6. After conjugation, growth curves were completed of *S. clavuligerus* SC6 with pMS82 and pSET152 (Figure 3.4 and Table 3.1). It was found that while pMS82 reduces growth rate (–17.6%), pSET152 has negligible effect. Additionally, a CA assay was performed on *S. clavuligerus* SC2 and SC6 with pMS82 and pSET152 to detect if either vector affected the secondary metabolism of the strains (Figure 3.5). pSET152 minimally affects CA production in SC6. However, pMS82 has greater negative impact, especially in SC2. Given these findings and pMS82 carrying hygromycin resistance, pSET152 was selected to integrate the barcoded into the chromosome of *S. clavuligerus* within this study.

Further validation of *S. clavuligerus* SC6 pSET152 was carried out demonstrating that colony morphology does not appear to be affected and that pSET152 stably integrates into the chromosome (Figure 3.6 and Figure 3.7).

The 20 barcoded strains were then created and the PCR reaction amplifying the barcode region from the genomic DNA of each strain was conformed (Figure 3.10). Further growth curves demonstrated that barcoded strains all have a similar growth rate (Figure 3.11).

The ability to amplify out barcoded regions from genomic DNA and identify the expected barcodes at expected abundance was then validated. The code used to identify barcode sequences within a FASTQ was presented. This code was then used to explore the effect of Levenshtein distances on barcode identification to allow for mutations within the barcode while maintaining identification stringency. Up to a distance of 4, identification was 100% accurate. This accuracy drops at distances of 5 and 6 (Figure 3.12).

Subsets of reads were also tested to define the minimum amount of reads accurately predict barcode relative abundance per timepoint. It was concluded that 10,000 reads or more were needed per timepoint to ensure reliability (Figure 3.13). Additionally, mixing both barcode DNA and barcoded strains in equal abundance resulted in barcodes being detected at expected ratios (Figures 3.14 and 3.15). The potential of PCR bias to skew the results of barcode relative abundance prediction was also explored given the fact that two PCR steps are included in the proposed methodology. PCR bias was negligible (Figure 3.16) with an average deviation from mean abundance of 0.19% and no systematic over or under-estimation being observed.

Given the validation of the proposed Bar-seq methodology demonstrating that it can accurately detect barcodes, a test fermentation was run in 50 mL TSB over 48 hours (Figure 3.17). This demonstrated that Bar-seq was able to successfully track the relative abundances of 20 barcoded strains of *S. clavuligerus* SC6 throughout a fermentation. Additionally, we could analyse the log fold change of the relative abundance of barcodes at the end compared to the start of a fermentation.

No barcoded strain became dominant within this fermentation. This work is foundational to the remainder of the thesis and demonstrates that we can now use Bar-seq to study population dynamics within a fermentation of *Streptomyces*. Yielding temporal data with an ability to attempt to link genotype to phenotype in future experiments.

Chapter 4 Population dynamics across fermentation scales

4.1 Introduction

Scale-up of industrial fermentations requires a significant investment, with strain selection representing a significant risk (Crater and Lievense, 2018). Scaling up laboratory fermentations to industrial production volumes introduces significant variability in microbial population dynamics, which can undermine yield, reproducibility, and process control (Rugbjerg and Sommer, 2019). Small-scale flasks grown in rich laboratory media, such as TSB, do not adequately reproduce the complex selective pressures imposed by industrial fermentations, which often involve large culture volumes, multistage seed-to-production transfers, stirring and nutrient-limited production media designed to maximise secondary metabolite synthesis (Hoskisson and Seipke, 2020). These changes in environmental conditions can destabilise the intended strain composition, favouring subpopulations with undesired phenotypes and reducing overall process consistency. Heterogeneity and batch-to-batch variability remain significant barriers to scaling *Streptomyces* fermentations effectively (Gomez-Escribano et al., 2021).

The extent to which scale and media composition influence strain-level population dynamics is not well understood, largely due to the lack of high-resolution methods to monitor populations over time. While the previous chapter established that Bar-seq is a suitable method to track mixed populations of barcoded *S. clavuligerus* strains, its application to more industrially relevant conditions is needed to test its robustness and to provide insights into how population structure responds to environmental changes. Importantly, industrial fermentations typically involve a multistage process, where biomass is first generated in nutrient-rich seed medium and is then transferred into production medium, which is optimised for secondary metabolite production. This

transition introduces an additional layer of selection, potentially amplifying heterogeneity in unpredictable ways.

This chapter aims to apply the Bar-seq methodology validated in the previous chapter to investigate how population dynamics change under extended fermentation timescales, increased culture volumes, and industrially relevant media and processes. Using barcoded *S. clavuligerus* populations, fermentations under multiple media conditions and timescales will be investigated. These insights will help to highlight the importance of monitoring and managing heterogeneity in industrial fermentations to ensure consistent performance.

James Croxford designed and performed the experimental work, sequencing, analysis, and figure preparation for this chapter, including all shake-flask and 1 L experiments and associated sample processing. For stirred-tank bioreactor fermentations (i.e. >1 L) conducted at GSK, reactor inoculation, operation, and sample collection were performed by Wei Li Thong and Alistair Middlemiss (GSK), while James Croxford prepared the inoculum and performed all downstream processing, sequencing, and data analysis.

4.2 Time-dependent dynamics in small-scale fermentations

After confirming that Bar-seq was able to track strain abundance throughout the course of a 48-hour fermentation in 50 mL of TSB, to move towards mimicking an industrial fermentation, Bar-seq was used to track strains of *S. clavuligerus* within a 50 mL TSB fermentation over 72 hours. This was done to see if the increase in timescale influenced the change in relative abundance of barcoded strains.

Using the same methodology from the previous chapter, where barcoded strains of *S. clavuligerus* SC6 pSET152 JCb1-20 were cultured for 48 hours and mixed in equal proportions based on OD_{600} measurements the barcoded strains were instead grown for 72 hours. For this fermentation and all fermentation going forward, ~ 1 x 10^7 CFUs were used as a starting inoculum. Strains JCb6 and JCb9 were not included in this study as the original 48-hour culture was contaminated. Genomic DNA was extracted at the start and end of the fermentation, and barcodes were amplified using the primers JC pSET152 barseq Fwd and Rev. Amplicons were sequenced, and the relative abundance of each barcode was calculated for both time points. A muller plot was created to show the difference in strain abundance over time (Figure 4.1).

In this case the average log_2 change in absolute barcode relative abundance (i.e. whether positive or negative) was 0.35, this compares with an average of 0.44 within the 48-hour fermentation. Additionally, the greatest absolute change in log_2 relative abundance for the 72-hour fermentation was 0.98 (JCb20) compared to 1.75 (JCb10) of the 48-hour fermentation. These results show the 72-hour fermentation is more stable than the 48-hour fermentation due to the decreased changes in barcoded strain relative abundance.

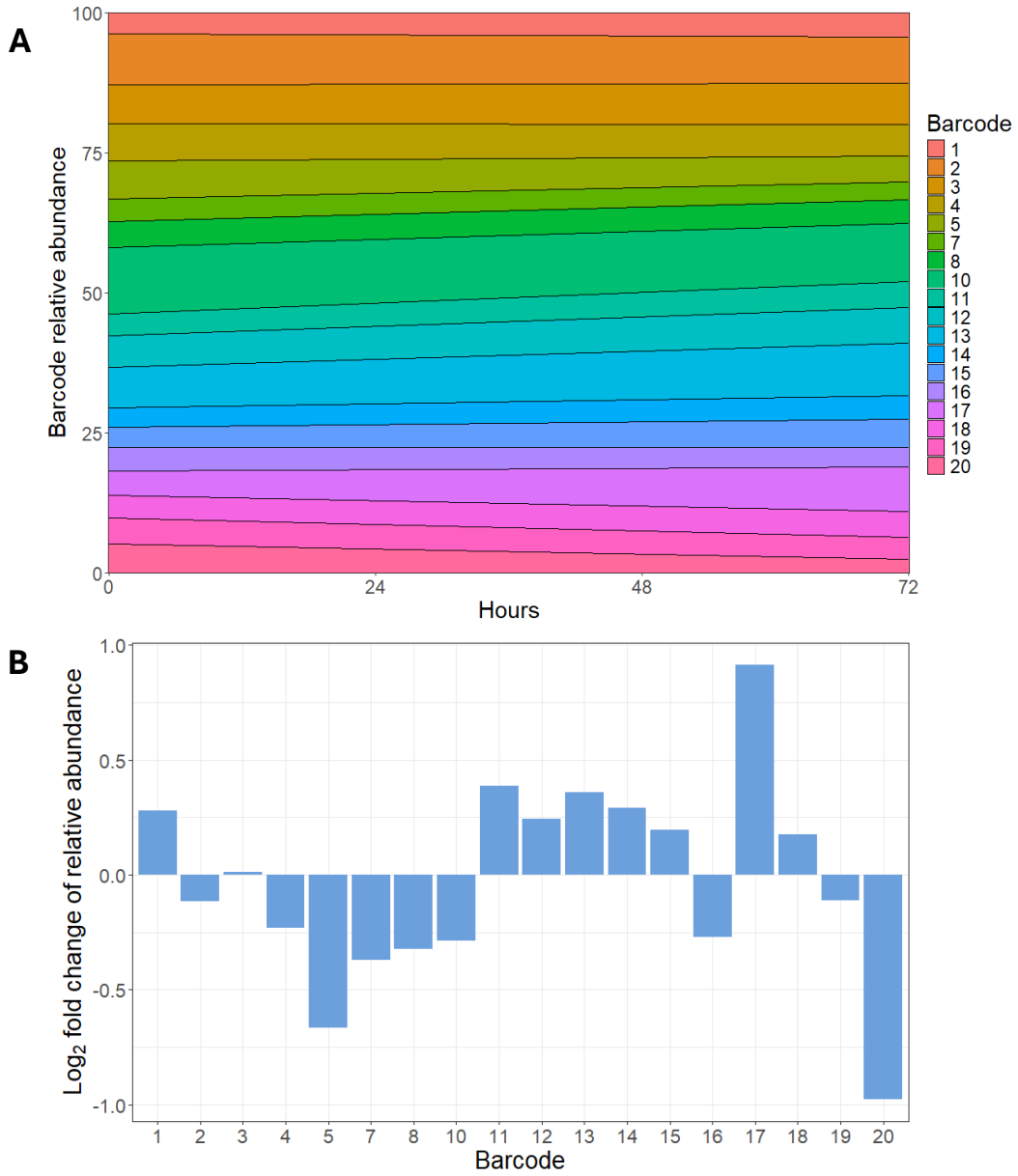


Figure 4.1. Minimal variation is observed in barcode abundance for flask fermentations of *S. clavuligerus* over 72 hours. (A) Fermentation in 50 mL TSB tracked over 72 hours shows similar variation to those tracked over 48 hours, indicating that Bar-seq is effective for monitoring changes across three days. The additional 24 hours does not significantly alter the observed variation compared to 48 hours. **(B)** Log_2 fold changes in barcode relative abundance between start and end time points demonstrate the capability of Bar-seq to track barcode representation over time and compare different time points.

4.3 Scale-up effects in TSB fermentations

To further move towards mimicking industrial conditions, fermentations using the 20 barcoded strains of *S. clavuligerus* SC6 pSET152 JCb1-20 were run in 1 L TSB over the course of 5 days (120 hours), in triplicate. As before, the 20 barcoded strains of *Streptomyces clavuligerus* SC6 pSET152 JCb1-20 were cultured for 48 hours and mixed in equal proportions based on OD_{600} measurements. This mixture was used to inoculate 1 L TSB within a 5 L Erlenmeyer flask (20% vessel volume) and grown for 7 days. Genomic DNA was extracted at the start and end of the fermentation, and barcodes were amplified using the primers JC pSET152 barseq Fwd and Rev. Amplicons were sequenced, and the relative abundance of each barcode was calculated for both time points. A muller plot was created to show the difference in strain abundance over time (Figure 4.2, A1, A2 and A3).

Compared to the two-fermentations run at a 50 mL scale it is clear there are greater changes in barcode relative abundance across the three replicates. Fermentation A1 has 8 strains that decrease in abundance over the course of 120 hours. At the start of the fermentation, cumulatively these strains have a relative abundance of 62.9%, whereas at the end of the fermentation they occupy just 7.76% of the relative abundance. Seemingly this is to the benefit of all other strains, specifically of JCb12, going from 6.30 to 20.4% of the strains within the fermentation. Similarly, JCb10, JCb16 and JC19 all increase in abundance within fermentation A2, seemingly at the cost of most other strains. Within fermentation A3 it is clear three strains become dominant at the end of the fermentation. These three strains occupy 41.7% of the final relative abundance. It is therefore clear that strains can significantly increase or decrease in relative abundance within a fermentation at this scale. Indeed, over the course of the 120 hours the barcode

with the highest absolute change in log_2 was 4.18 (JCb9 in A3). The average log_2 change across the three was 1.98 which compares to 0.40 of the two 50 mL fermentations completed previously. It is however worth noting that despite the large increases in instability, in none of the three replicates does one strain become dominant (>50% of the relative abundance) within the fermentation.

It is also worth noting that batch to batch variability, a common issue within industrial fermentation, is clear even at this scale across the three replicates. For each replicate the average absolute log_2 change was 2.05, 1.38 and 2.53 respectively. Additionally, across all three replicates a different number of strains increased or decreased in abundance with significantly different cumulative relative abundance of dominant strains at the end of each fermentation.

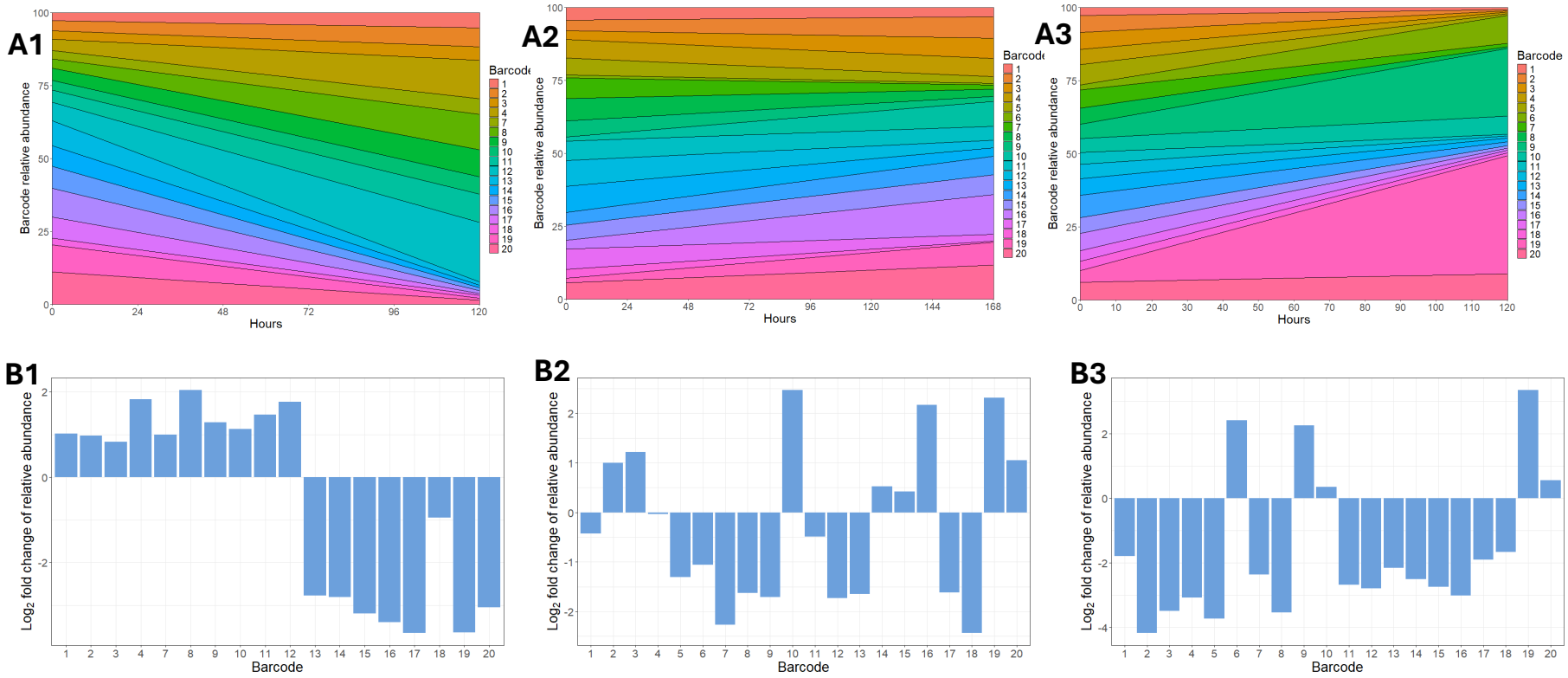


Figure 4.2. Scale-up of fermentation increases barcode variability in 1L fermentations of *S. clavuligerus* but not at consistent rates. **(A1, A2, A3)** Barcode relative abundance was tracked within a 1 L fermentation in TSB over 120 hours. **(B1, B2, B3)** The maximum $\log_2$ fold change of a barcode was 2.08, 1.48 and 1.14 times higher than that observed in a 50 mL TSB fermentation from B1 to B3 repsectively, demonstrating inconsistent levels of variability during scale-up.

4.4 Bar-seq in multistage fermentations using industrial media

To more closely mimic the industrial fermentation conditions Bar-seq was explored within a multistage fermentation. This would include using a seed media (S2A), used to promote rapid and dense bacterial growth, which would be used as an inoculum for production media, used to promote antibiotic production. This would better replicate barcoded strain dynamics within industrial fermentation conditions for CA production. Barcode relative abundance would be used to track the relative abundance of barcoded strains at each of these stages of the fermentation. In the case of this model the 20 barcoded strains of *S. clavuligerus* SC6 pSET152 JCb1-20 were cultured for 48 hours and mixed in equal proportions based on OD_{600} measurements. This mixture was used to inoculate 50 mL of S2A seed media which was then grown for 48 hours. This seed media was then used as a 10% inoculum for 50 mL CM5 complete production media which was grown for a further 120 hours (7 days growth in total). Genomic DNA was extracted and barcode relative abundance determined at 0, 48 and 168 hours. A muller plot was created to show the changes in barcode relative abundance (Figure 4.3).

Based on the plotted abundances Bar-seq can track relative abundance throughout a multistage fermentation. Using the data gathered it is possible to look at the log_2 fold changes in relative abundance between both 0 to 48 hours (seed growth), 48 to 168 hours (complete medium growth), and 0 to 168 hours (total growth).

In this case when comparing the relative abundances at hour 0 compared to hour 168 the average log_2 change was 0.57 with the greatest change being 1.75 (JCb6). The average change of 0.57 compares to an average of 0.40 and 1.98 for 50mL (48 hours) and 1L (120 hours) TSB fermentations respectively. No strain became dominant within the fermentation.

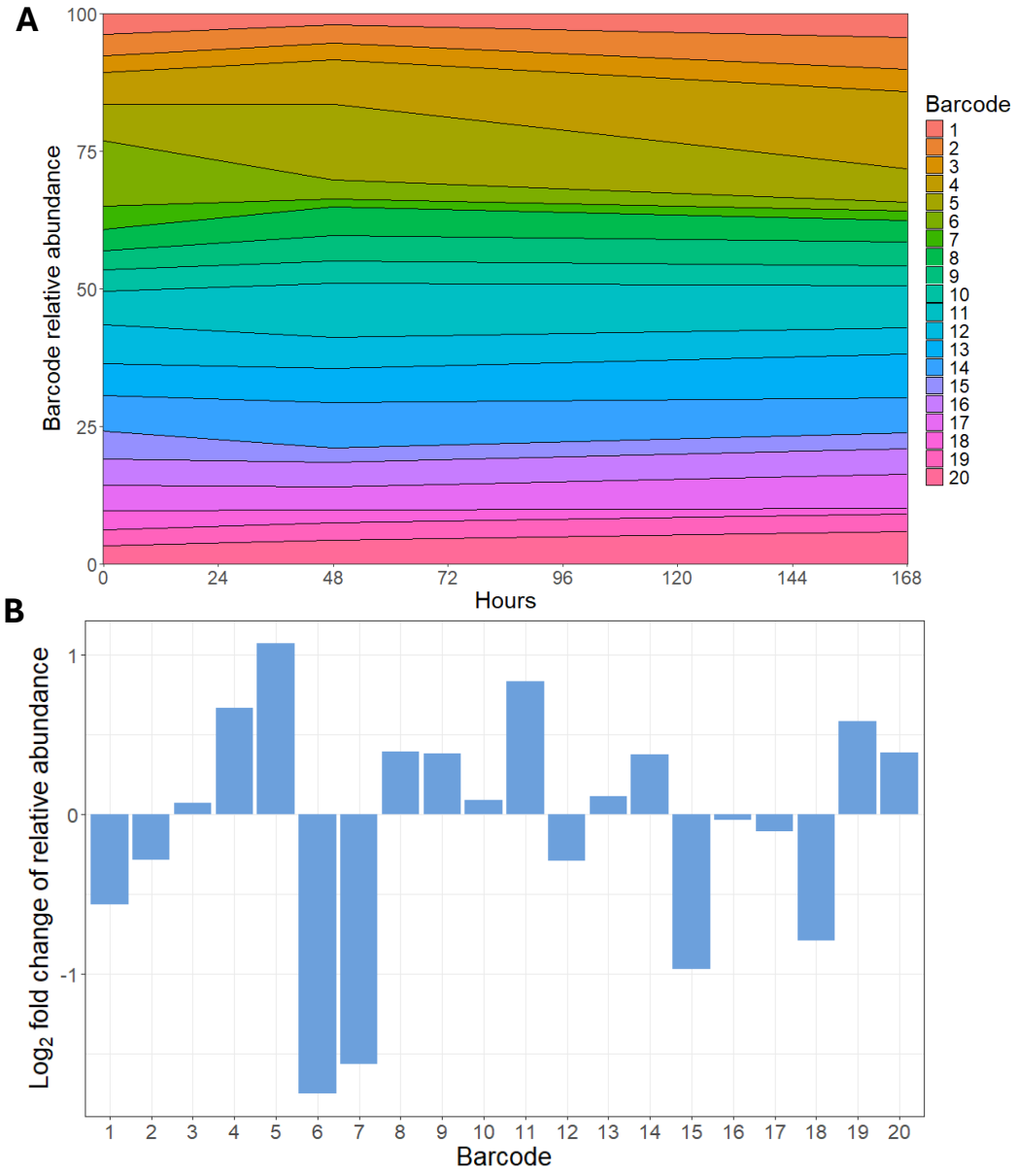


Figure 4.3. Bar-seq can be used to track multistage fermentations of *S. clavuligerus*. **(A)** A 50 mL fermentation in S2A media was grown for 48 hours and used as a 10% inoculum for 50 mL CM5 media, where it was grown for an additional 120 hours, mimicking industrial multistage fermentations. **(B)** Log_2 fold change in barcode relative abundance between hours 0 and 168. The transition from S2A to CM5 media does not result in increased variation compared to TSB fermentations.

Once Bar-seq had been shown to be suitable for use within a multistage fermentation, efforts were made to further replicate industrial fermentation conditions. To do this a multistage fermentation was carried out in 1 L of industrial media. The 20 barcoded strains of *S. clavuligerus* SC6 pSET152 JCb1-20 were cultured for 48 hours and mixed in equal proportions based on OD_{600} measurements. This mixture was used to inoculate 50 mL of S2A seed media which was then grown for 48 hours. This seed media was then used as a 5% inoculum for 1 L CM5 complete production media which was grown for a further 120 hours (7 days growth in total) in a 5 L Erlenmeyer flask.

Given that we are now moving closer to mimicking an industrial fermentation, it is important that we validate the growth of our *S. clavuligerus* strains within the multi-stage fermentations using industrial media as this media is not commonly used within the lab. Traditional methods to measure bacterial growth such as spectrophotometry or dry weight analysis are not possible when using industrial media to the opaque nature of the media, its high viscosity and high particulate count. Furthermore, the high levels of rapeseed oil and soy protein concentrate flour within the CM5 media make any visual analysis of growth within the media difficult.

While a thickening of the media and minimal change in colour do qualitatively indicate growth, further validation is needed to validate growth within the media. Additionally, while it is possible to extract nucleic acids from the media, using the same method as that used for DNA extraction from TSB media, the extracted DNA is often lower in concentration and higher in salt contamination, allowing it to be suitable for basic PCR methods but difficult using methods such as qPCR to measure growth.

Such is the difficulty to measure growth within this industrial media, a viscometer is often used in industrial settings to measure growth. This is because as the biomass

increases as does the thickness of the media, it may also be due to the emulsification of oil within the media. Furthermore, while it is possible to isolate colonies form this media, the high levels of viscosity make it difficult to reproducibly and reliably take homogenous samples. For example, when plating out 1 mL of culture onto an agar plate high variations in CFU were observed (data not shown).

In lieu of commonly used growth quantification methods, a total protein assay, cell wet weight and CA acid assay were completed over the course of the fermentation. These methods were used to confirm the growth of *S. clavuligerus* within the media and predict the growth rates at different stages.

The total protein of the culture was measured through a Bradford assay in triplicate samples taken at 24-hour intervals. When the total protein of the multistage fermentation was plotted using a log scale it resembled a traditional bacterial growth curve (Figure 4.4). The growth appeared to plateau after 72 hours, suggesting that growth was no longer exponential. This includes 24 hours of growth within production media, this aligns with the purpose of production media not being for the rapid growth of cells but instead the production of antibiotic. The drop after 48 hours within the total protein levels represents the inoculation of S2A into CM5 media.

To calculate cell wet weight over the same time course 1 mL of cell culture was spun at 20,000 x *g* for 10 minutes in a 1.5 mL Eppendorf tube. The supernatant was removed and the difference in weight of the Eppendorf was calculated. A similar growth trend is seen here compared to the total protein assay of growth (Figure 4.5). There is a drop in 48 hours due to the transfer into production media but broadly growth can be seen that plateaus out, suggesting the end of exponential growth. However, compared to the total protein growth curve this plateau occurs after 120 hours. Nonetheless, these results both indicate

growth within both S2A and CM5 media. This coupled with the fact that *S. clavuligerus* can be isolated from the media after 7 days of growth confirms that growth does reliably occur within this multistage fermentation.

To also confirm the production of CA within the production media, a CA assay was conducted in samples over the same period using an CA imidazole assay. Similar to the growth curves observed when using total protein and wet weight assays a drop in production can be seen between the transfer of S2A and CM5 as expected (Figure 4.6). The production of CA however quickly increases and remains stable at a constant level of around 40 µg/mL after transfer into CM5 media. Unfortunately, due to the previously discussed viscosity and high particulate nature of the S2A and CM5 media, it is not possible to normalise the levels of CA produced to the cell dry weight as done in previous experiments.

Genomic DNA was also extracted from this fermentation and barcode relative abundance determined at 0, 48 and 168 hours (Figure 4.7). It is again clear that, like the multistage fermentation carried out within 50 mL, there is little variation in barcode relative abundance. Additionally similar to the 50 mL fermentation one strain does not become dominant within the fermentation. Two barcoded strains do significantly increase in abundance (JCb10 and JCb17) over the course of the 168 hours; however these changes appear to occur within the seed medium growth and then relative abundance remains stable within the production media, an aspect that will be explored later. It is however worth again noting that although the strains JCb10 and JCb17 increase in abundance, they in no way become domain within the fermentation, occupying 26.2% of the total barcode relative abundance.

The average absolute change in $\log_2$ fold change across 168 hours for all 20 barcoded strains was 0.57 with the largest absolute change being 1.86 (JCb10). This average value of 0.57 compares to the 50 mL (168 hours) multistage fermentation of 0.57 and the 1 L (120 hours) TSB fermentation of 1.99.

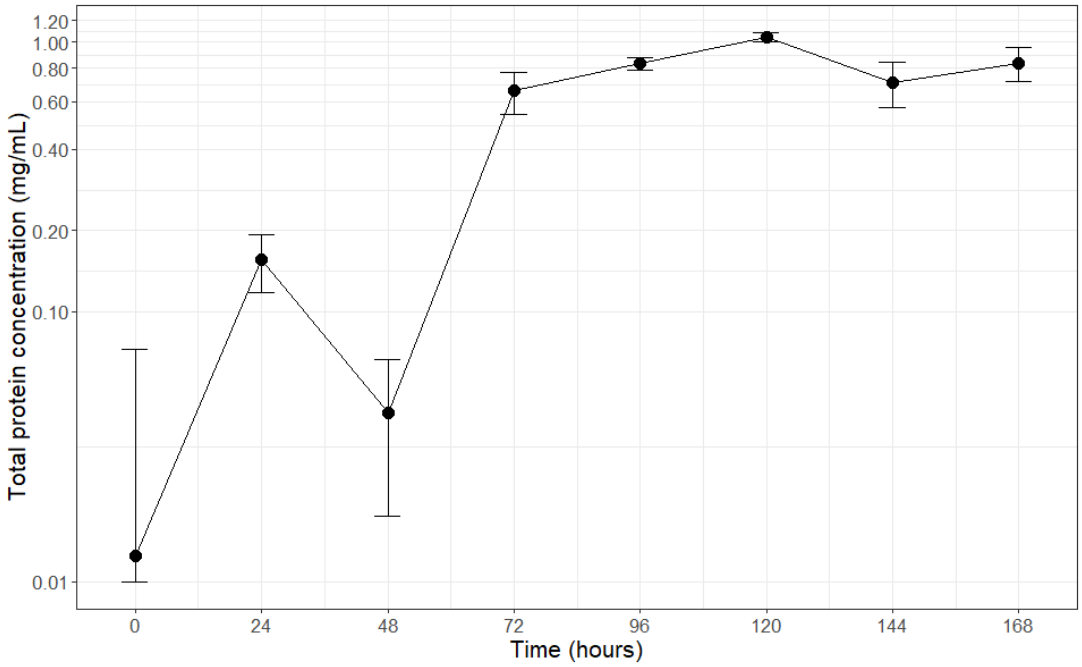


Figure 4.4. Total protein concentration can be used to measure growth within a multistage fermentation of *S. clavuligerus*. A Bradford assay was used to measure total protein concentration during a 1 L S2A-CM5 fermentation. This method substitutes for dry weight and OD_{600} growth curves, which are impractical due to the high viscosity of the media. The drop in total protein observed at 48 hours corresponds to the transfer between S2A and CM5 media. Error bars correspond to SDs (n = 3).

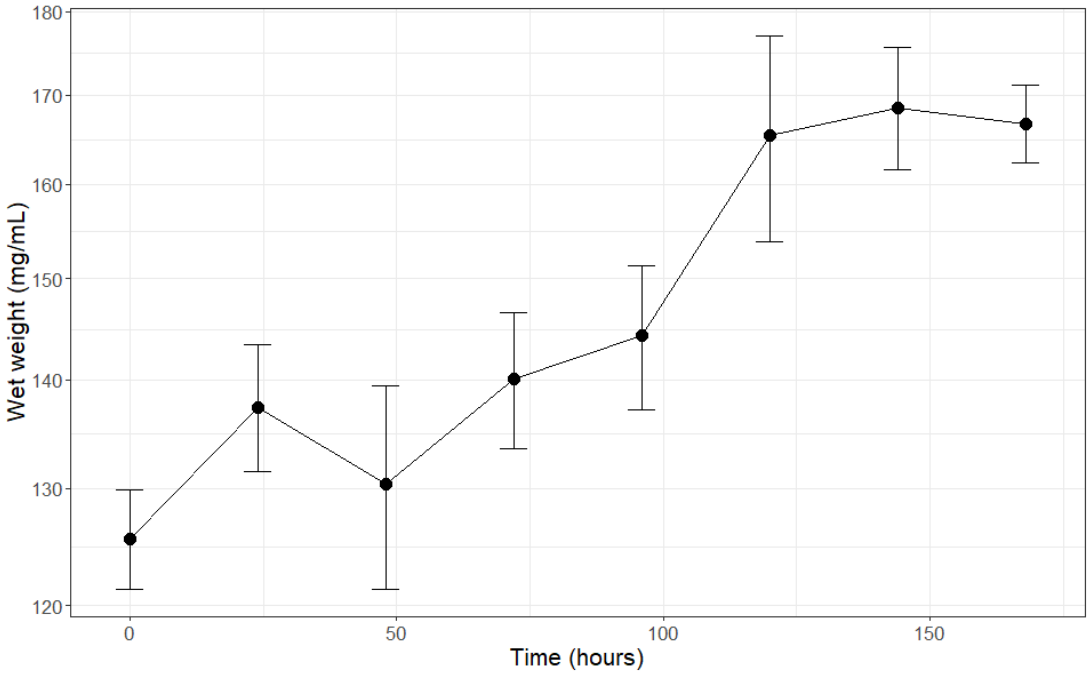


Figure 4.5. Wet weight can be used to measure growth within a multistage fermentation of *S. clavuligerus*. At each timepoint, 500 µL of media was collected, centrifuged at top speed, and the supernatant removed. This method substitutes for dry weight and OD_{600} growth curves, which are impractical due to the high viscosity of the media. The drop in wet weight observed at 48 hours corresponds to the transfer from S2A to CM5 media. Error bars correspond to SDs (n = 3).

.

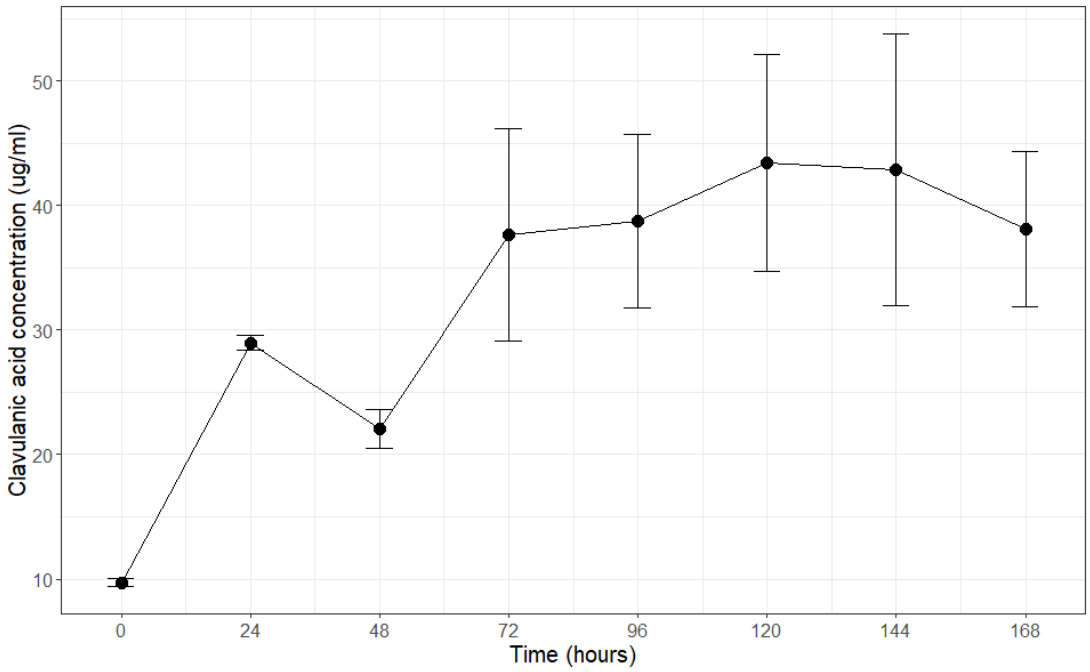


Figure 4.6. Clavulanic acid is produced throughout the course of 1 L multistage fermentations of *S. clavuligerus*. Clavulanic acid production was measured using an imidazole assay throughout the course of a 1 L S2A-CM5 fermentation. A drop in clavulanic acid concentration at 48 hours corresponds to the transfer from S2A to CM5 media. Due to the high viscosity of the media, normalization to dry weight was not feasible. Error bars correspond to SDs (n = 3).

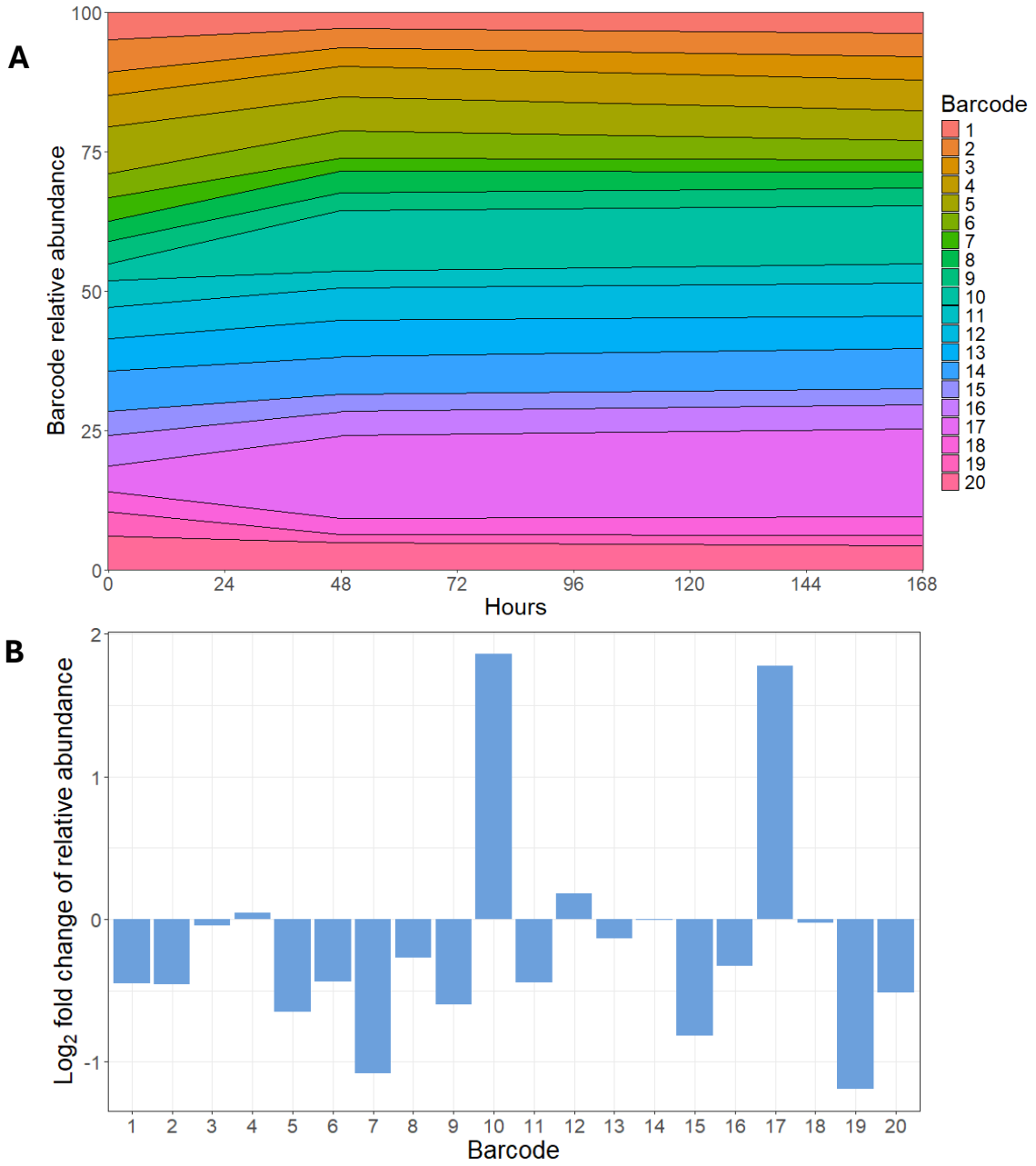


Figure 4.7. Scaling up multistage fermentations of *S. clavuligerus* does not increase strain variability. **(A)** A 50 mL fermentation in S2A media was grown for 48 hours and used as a 10% inoculum for 1 L CM5 media, where it was grown for an additional 120 hours, mimicking industrial multistage fermentations. **(B)** These results demonstrate that scaling up S2A-CM5 fermentations does not have the same negative impact on variability observed in TSB fermentations.

4.5 Batch-to-batch variability in 1 L multistage industrial fermentations

After the growth of industrial *S. clavuligerus* strains and use of Bar-seq had been validated within industrial media at a 1 L scale over the same timescale used within an industrial setting, it was decided to further investigate batch-to-batch variability. Therefore, the 1 L multistage fermentation previously completed (Figure 4.4) was repeated across six replicates, this would allow us to investigate batch-to-batch levels of variation at a scale closer to industrial levels.

As before, the 20 barcoded strains of *Streptomyces clavuligerus* SC6 pSET152 JCb1-20 were cultured for 48 hours and mixed in equal proportions based on OD_{600} measurements. This mixture was used to inoculate 50 mL of S2A seed media which was then grown for 48 hours. This seed media was then used as a 5% inoculum for 1 L CM5 complete production media which was grown for a further 120 hours (7 days growth in total) in a 5 L Erlenmeyer flask. Genomic DNA was extracted and barcode relative abundance determined at 0, 48 and 168 hours. The six replicates were biologically independent of each other (Figures 4.8 and 4.9).

It is again clear that all fermentations appear to be stable. Significantly, and as has been seen before, no single strain comes to dominate in any of the fermentations, nor does any strain reach an undetectable level of abundance. Across all six replicates the average absolute $\log_2$ fold change was 0.28, this compares to 0.57 of the previously completely 1L multistage fermentation. This further demonstrates the relative stability of barcoded strains within a 1 L fermentations mimicking industrial conditions. In no replicate does any strain come close to being dominant (>50% relative abundance) and the largest absolute $\log_2$ fold change across all replicates was 1.23 (JCb18 in A6). This

value compares to the average change of 1.98 in the 1 L (120 hours) TSB fermentations previously shown, further highlighting the stability of multistage fermentations.

Additionally, across the six replicates discussed here, while different strains increase or decrease in relative abundance, a very small level of batch-to-batch variability is seen. For the fermentation A1 to A6 the average absolute $\log_2$ fold change is 0.29, 0.30, 0.29, 0.29, 0.26 and 0.26 respectively. However, the small range of these values is likely down to the stable nature of these fermentations and therefore little batch-to-batch variation being able to be observed.

The 'bumps' in abundance seen in fermentation A2 and A4 are likely due to technical artefacts during barcode PCR amplification or sequencing, such as uneven amplification efficiency or low initial DNA input, likely related to isolating genomic DNA from the viscous/particulate industrial media. A return to expected abundance levels is observed in subsequent timepoints, confirming that this is an anomaly specific to datapoints rather than a biological effect.

Additionally, as several fermentation replicates have been performed, the number of times individual barcodes increased or decreased in abundance was previously not investigated. To test this analysis was carried out to determine the randomness of individual barcodes predominating in a fermentation and to further validate Bar-seq. The total times a barcode increased in abundance in all fermentation to this point was 112 and a total of 143 barcodes decrease in abundance, suggesting that fermentations are unbiased. Furthermore, to further assess any bias in the Bar-seq experiments to this point, all the data for changes in $\log_2$ fold change for all fermentation was plotted on a histogram (Figure 4.10). While the data appeared to be normally distributed around 0, a Shapiro-Wilk test on the $\log_2$ fold change data returned a p value of 1.47×10^{-12}

suggesting the data is not normally distributed. This is likely due to the long tails to the positive and negative values. Most of the data seemed to be centred around 0 suggested randomness for barcodes to either increase or decrease in abundance, validating the technique up to this point. Moreover, the potential emergence of mutations during fermentations mean that the measures are not truly random as each strain may not be identical.

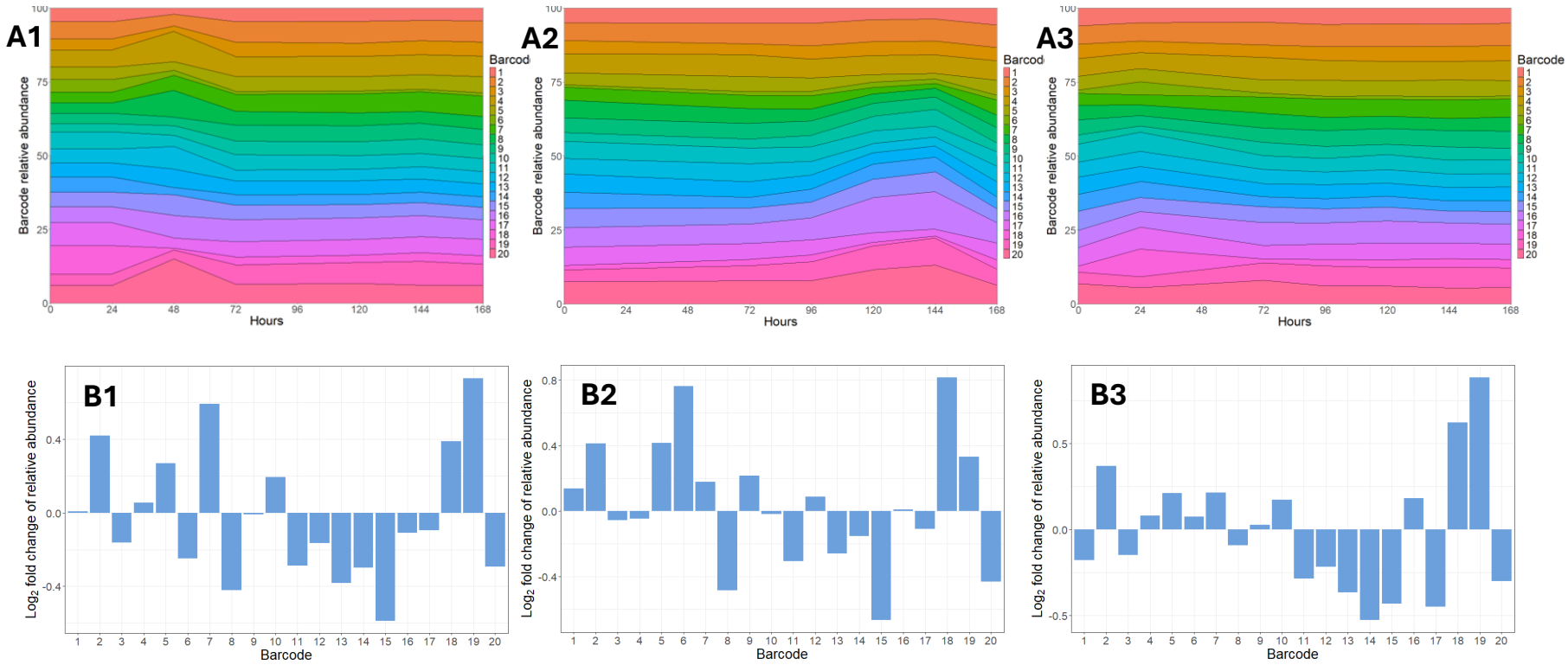


Figure 4.8. Minimal variation within 1 L multistage fermentations of *S. clavuligerus* is observed in fermentations 1 to 3. (A1, A2, A3) 50 mL fermentations in S2A media were grown for 48 hours and used as a 10% inoculum for 1 L CM5 media, where they were grown for an additional 120 hours, mimicking industrial multistage fermentations. **(B1, B2, B3)** Log_2 fold change in barcode relative abundance at hour 168 compared to hour 0 shows minimal variation between strains.

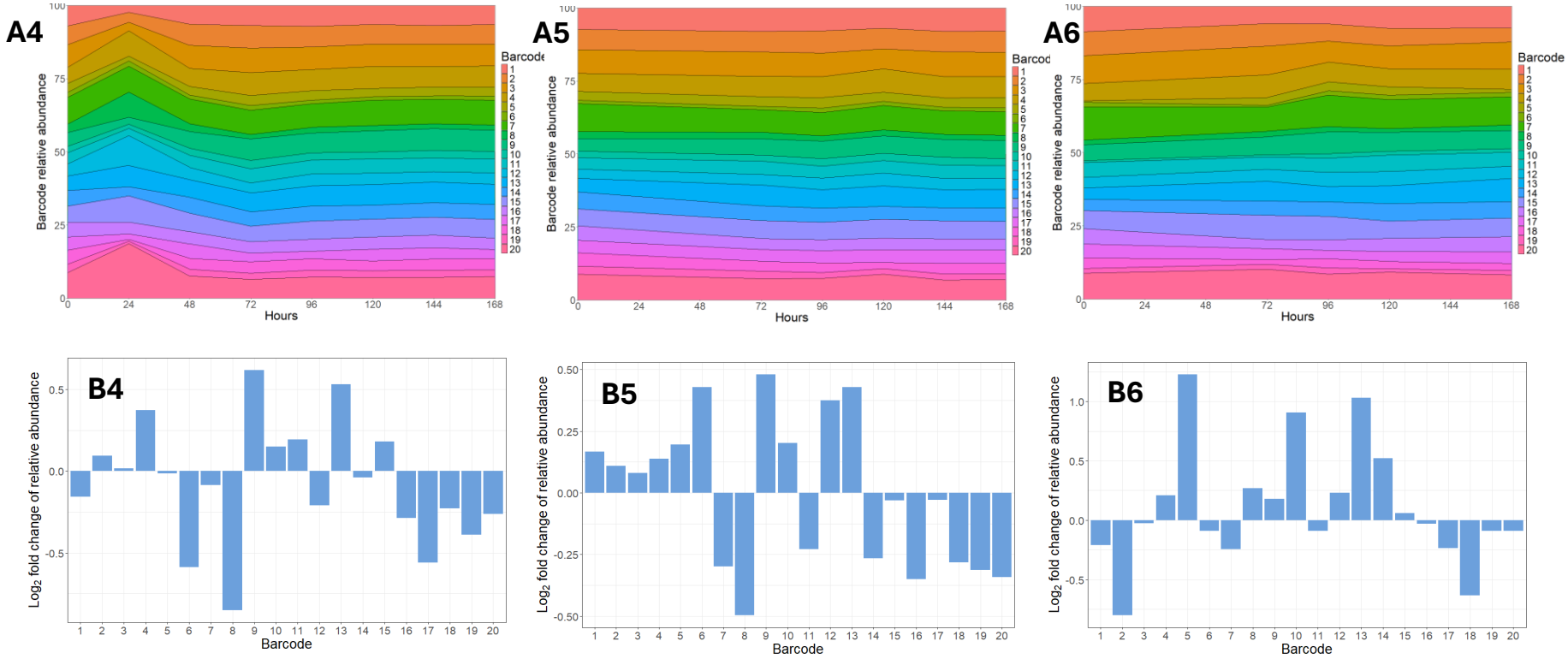


Figure 4.9. Minimal variation within 1 L multistage fermentations of *S. clavuligerus* is observed in fermentations 4 to 6. (A4, A5, A6) 50 mL fermentations in S2A media were grown for 48 hours and used as a 10% inoculum for 1 L CM5 media, where they were grown for an additional 120 hours, mimicking industrial multistage fermentations. **(B4, B5, B6)** Log_2 fold change in barcode relative abundance at hour 168 compared to hour 0 shows minimal variation between strains.

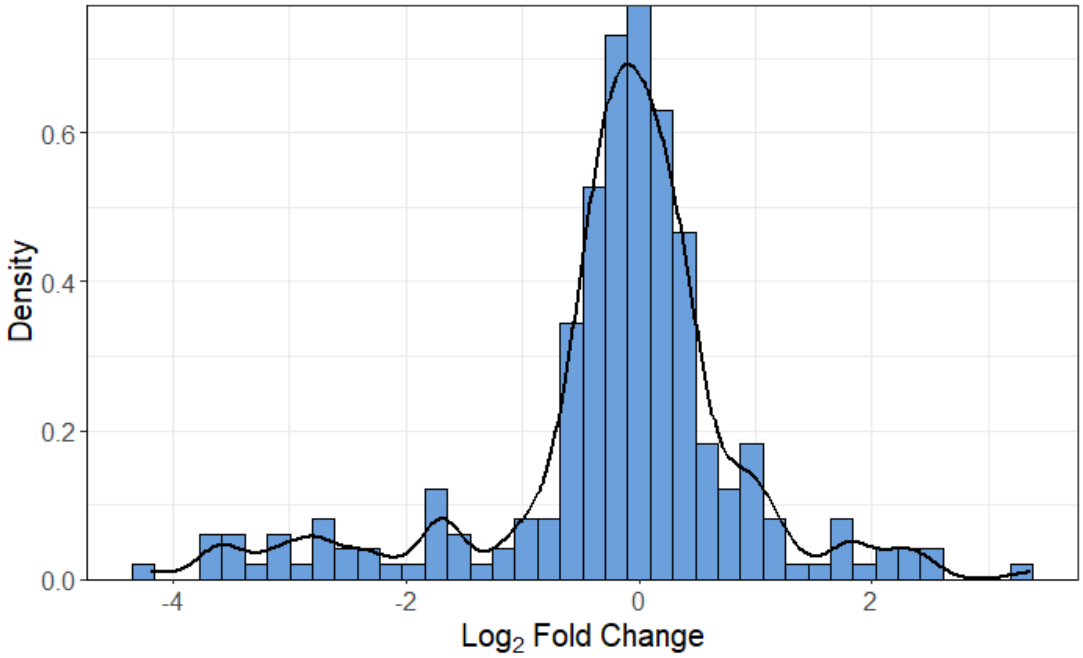


Figure 4.10. Distribution of barcode log_2 fold changes across all fermentations included in this study. The log_2 fold change in barcode relative abundance between the start and end of each fermentation was plotted for all barcoded strains to this point of the study. The resulting distribution is approximately symmetric and centred around zero, consistent with random strain behaviour. A Shapiro–Wilk test confirmed that the data are not normally distributed ($p = 1.47 \times 10^{-12}$), although the overall shape remains bell-like. This supports the use of Bar-seq to monitor strain competition without introducing systematic bias towards any barcode.

4.6 Statistical comparison across media, scale, and stage

Given there is an appropriate number of replicates across the conditions so far measured we can now start to analyse the levels of instability across fermentation volumes and medias. For each type of fermentation, the $\log_2$ fold change values were squared to allow all data to be plotted against each other as positive values. These squared values allows directly visualisation of instability in barcoded strains relative abundance compared to other fermentation types.

Firstly, all the squared $\log_2$ fold changes between 1L TSB and CM5 fermentation were plotted (Figure 4.11). While TSB fermentations were carried out over 120 hours and multistage fermentations were carried out over 168 hours, based on the growth data where fermentations were in stationary phase at this point, it was deemed valid to compare these data. Based on the resulting boxplot (Figure 4.11) the TSB fermentations gives rise to much greater instability than multistage fermentation, suggesting that the industrial media used within this study was much better at maintaining stability in fermentations. The data were not normally distributed; therefore, a Wilcoxon signed-rank test was used to compare the means of the two groups, resulting in a p value of 2.2×10^{-16}. This demonstrates that TSB gives rise to significantly more unstable fermentations that multistage industrial media fermentations. This is also highlighted by the squared $\log_2$ fold change value for TSB was 40.5 x higher than that of the industrial media.

Secondly while only one 50 mL multistage fermentation had been completed compared to 6 1 L multistage fermentations, the squared values of the $\log_2$ fold changes were plotted against each other to investigate whether scaleup within industrial media affected the levels of instability (Figure 4.12). This is because scale-up may cause

significant problems within the fermentation process (Viana Marques et al., 2018). Having data to elucidate how the scale-up process affects instability will offer insight into population dynamics within a fermentation. It was found that scaling the multistage fermentation from 50 mL to 1 L appeared to increase stability. Larger-scale fermentations show a reduction in barcode instability by an average of 4.00 x, suggesting improved strain stability at higher volumes in industrial media over 168 hours. The data are not normally distributed; therefore, a Wilcoxon signed-rank test was used to compare the means of the two groups, resulting in a p value of 1.97×10^{-2}. However, because the 50 mL multistage condition was represented by a single fermentation, this comparison should be treated as exploratory and would require additional 50 mL replicates to confirm any scale-associated difference.

Finally, as we have observed potentially greater instability within the S2A seed media compared to the CM5 production media within 1 L multistage fermentations (Figure 4.4), the squared $\log_2$ fold changes of all barcodes when comparing 0-48 h (seed media) and 48-168 h (complete media) were plotted (Figure 4.13). The mean $\log_2$ fold change values for the S2A medium and CM5 medium stages of fermentation were 0.627 and 0.706 respectively. This suggests that there is little difference between the two groups. Indeed, as the data were not normally distributed a Wilcoxon signed-rank test was used to compare the means of the two groups, resulting in a p value of 0.7998, confirming this lack of difference at this point.

However, it is important to consider the timescales at which each of these two stages are carried out within the fermentation. The S2A seed medium is used, as discussed, to promote growth and is fermented for 48 hours (2 days). This is then used as an inoculum for CM5 media which is used to maximise production of, in this case CA, over a further

120 hours (5 days). The similarity in instability levels between these two stages is notable. Despite the industrial strains being grown in CM5 for 2.5 times longer, they do not exhibit 2.5 times more instability. As a result, the squared $\log_2$ fold changes across both media types were normalised to the number of days to which *S. clavuligerus* is grown in that media for. When normalised, the average squared $\log_2$ fold change in S2A has an average value of 0.314 whereas CM5 has an average value of 0.1412. This means the level of squared $\log_2$ fold change abundance, when normalised, is 2.22x higher per day in S2A compared to CM5, suggesting that seed culture is twice as unstable than production medium. A Wilcoxon test returned a p value of 6.14×10^{-5} confirming the difference between the groups.

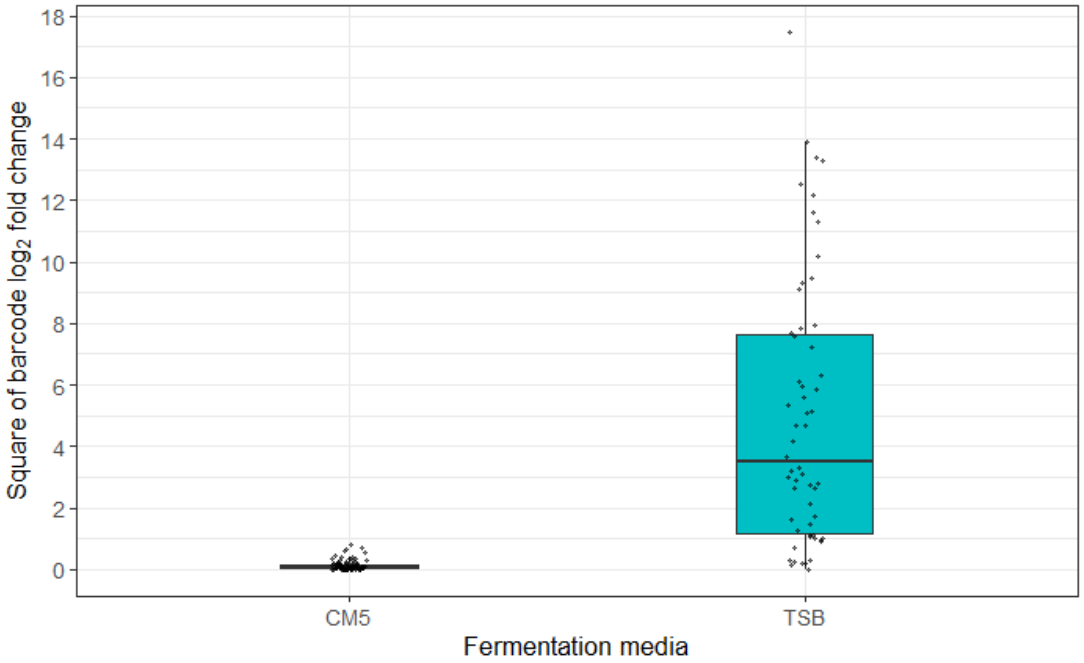


Figure 4.11. TSB causes greater variation in barcode relative abundance of *S. clavuligerus* fermentations compared to industrial media. The square of all barcode $\log_2$ fold changes for 168-hour, 1 L fermentations in TSB and S2A-CM5 media were compared. TSB fermentations show a large increase in variability compared to S2A-CM5 fermentations, indicating that S2A-CM5 media helps maintain strain stability. Data were not normally distributed; therefore, a Wilcoxon signed-rank test was used to compare the means of the two groups, resulting in $p = 2.2 \times 10^{-16}$.

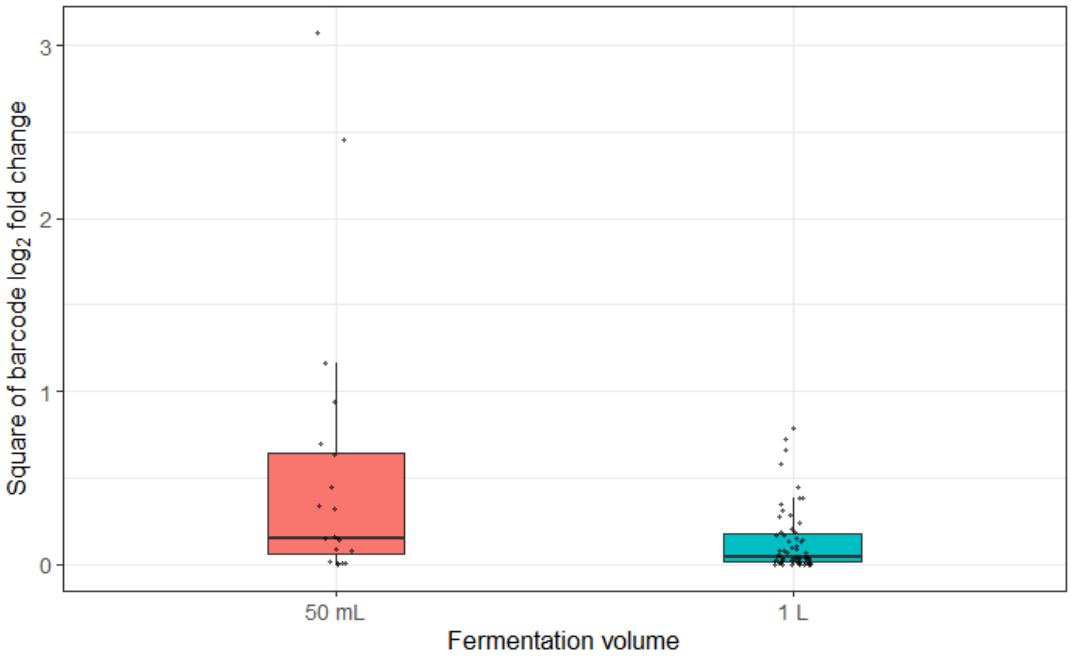


Figure 4.12. Scale-up of multistage fermentations of *S. clavuligerus* appears to increase strain stability. The square of all barcode $\log_2$ fold changes for 168-hour, 50 mL, and 1 L fermentations in S2A-CM5 media was plotted. Larger-scale fermentations show a reduction in barcode instability, suggesting improved strain stability at higher volumes. Data were not normally distributed; therefore, a Wilcoxon signed-rank test was used to compare the means of the two groups, resulting in $p = 0.01968$. This comparison includes one 50 mL multistage fermentation versus six 1 L multistage fermentations; conclusions should be interpreted cautiously.

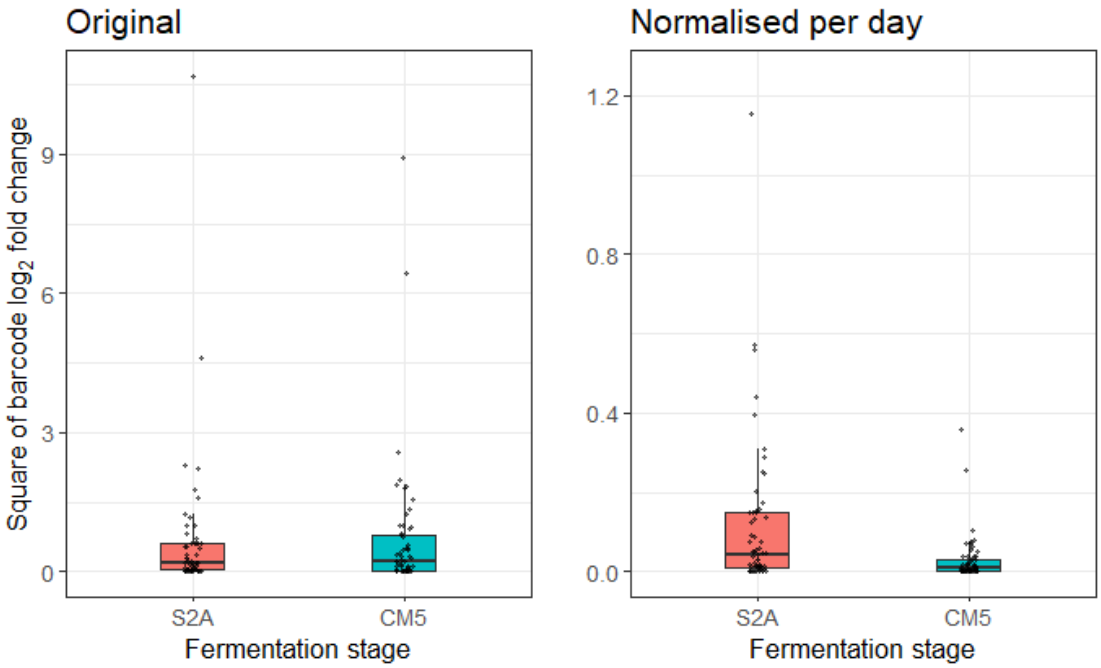


Figure 4.13. Strain variation is higher in S2A fermentations after normalising for time. The square of barcode $\log_2$ fold changes is shown for 48-hour, 50 mL S2A fermentations and 120-hour, 1 L CM5 fermentations of *S. clavuligerus*. Original shows the raw data. Normalised per day shows values normalised by fermentation duration ($\log_2$ fold change per day, squared). Data were non-normally distributed, so a Wilcoxon rank-sum test was used. While no significant difference was observed in the raw data ($p = 0.7998$), S2A showed significantly greater variation when normalised for fermentation time ($p = 6.14 \times 10^{-5}$), suggesting higher per-day instability. For clarity, the y-axis of the normalised panel is capped at 1.5 (excluding one data point). All data points, including outliers, were retained for statistical analysis.

4.7 Mimicking industrial fermentations in stirred-tank bioreactors

After these statistical analyses were completed on fermentations at a 1 L scale in shaking culture, further attempts were made to closer mimic industrial conditions to analyse how population dynamics were affected. To achieve this, the GSK Pilot Fermentation Facility at the GlaxoSmithKline (GSK) Worthing site that is used in the development of authentic industrial *S. clavuligerus* strains was used. The 20 barcoded strains of *S. clavuligerus* SC6 pSET152 JCb1-20 were cultured for 48 hours and mixed in equal proportions based on OD_{600} measurements. After, 40 mL of this mixture was used to inoculate 4 L of S2A media within a stirred tank bioreactor. This was grown for 48 hours after which 800 mL of this culture was used to inoculate four 10 L stirred tank bioreactors containing CM5 to a final volume of 7.5 L. The CM5 culture was grown for a further 6 days (24 hours longer than previously completed fermentation at 1 L scale). Across the two fermentation stages pH was controlled at 6.8 throughout the process and the cultures were agitated at an increasing rate between 300-700prm to account for the increase in viscosity. Airflow was controlled at 8 L / min and the temperature was maintained at 26 °C. These conditions closely mimic industrial testing conditions. Genomic DNA was extracted every 24 hours. The relative abundance of barcoded strains was plotted on a muller plot across four replicates (Figures 4.14 and 4.15).

It is immediately clear that across all four fermentations there is minimal change in barcoded strain relative abundance, which is of significance as although previous evidence has suggested scaling up from 50 mL to 1 L fermentation within industrial multistage fermentation does not adversely affect population dynamics, population dynamics have not before been explored within a single species at this level before.

Across the four replicate fermentations the average absolute $\log_2$ fold change in barcode relative abundance was 0.76 between hour 0 and hour 192. This compares to an average value of 0.28 for 1 L multistage fermentation carried out over 168 hours. This value of 0.76 is still significantly below the average value of 1.98 for 1 L TSB fermentation run over 120 hours, however. The largest change in $\log_2$ fold change was 2.74 (JCb18 in Fermentation A4), this compares to 1.23 and 4.18 for the 1 L multistage and TSB fermentation respectively.

The barcoded strain JCb16 appears to increase in abundance at a consistent rate (a $\log_2$ fold change of around 1). This anomaly is explained by the use of the same seed culture for all four fermentations. Therefore, the strain JCb16 had increased in abundance within the first 48 hours of growth within the S2A media. This aligns with evidence seen previously that the S2A stage of the fermentation gives rise to greater relative changes in barcode relative abundances than CM5. To that end, the differences were explored and plotted within a boxplot to visualise the difference between S2A and CM5 stages of this fermentation. As before the squared $\log_2$ fold changes were normalised to the number of days each culture was grown for (Figure 4.16).

Across these stirred tank conditions, the mean squared $\log_2$ fold change within the S2A stage of the fermentation was 1.57. This compares to the CM5 stage of the fermentation it was 0.56, 0.26, 0.48 and 0.28 respectively making an average of 0.40. This suggests that, with normalisation, the seed (S2A) stage of the fermentation is around four-fold more unstable than that of the production (CM5) stage of the fermentation. This aligns with what we have observed within a 1 L multistage fermentation were, once normalised the seed stage of the fermentation was around twice as unstable as the production stage of the fermentation.

As the pH of a culture is known to play a key role within *Streptomyces* stability, especially in fermentation (Baskaran and Muthukumarasamy, 2017). It is notable that the pH within the stirred tank bioreactor was controlled. To compare this to a 50 mL multistage fermentation the pH of the culture was measured at the start and end of the fermentation. A 50 mL fermentation ranges from pH ranged from 6.8 to 7.3 in S2A-CM5 and from 7.0 to 8.8 in TSB (Figure 4.17), between hours 0 and 168. However, in the stirred tank it was kept constant at 6.8 pH. Importantly however, across all of four fermentations at this scale one strain did not come to dominate the fermentation. Indeed, across the four fermentations the strain with the highest relative abundance at hour 192 is JCb17 representing 15.07% of the abundance in fermentation A3.

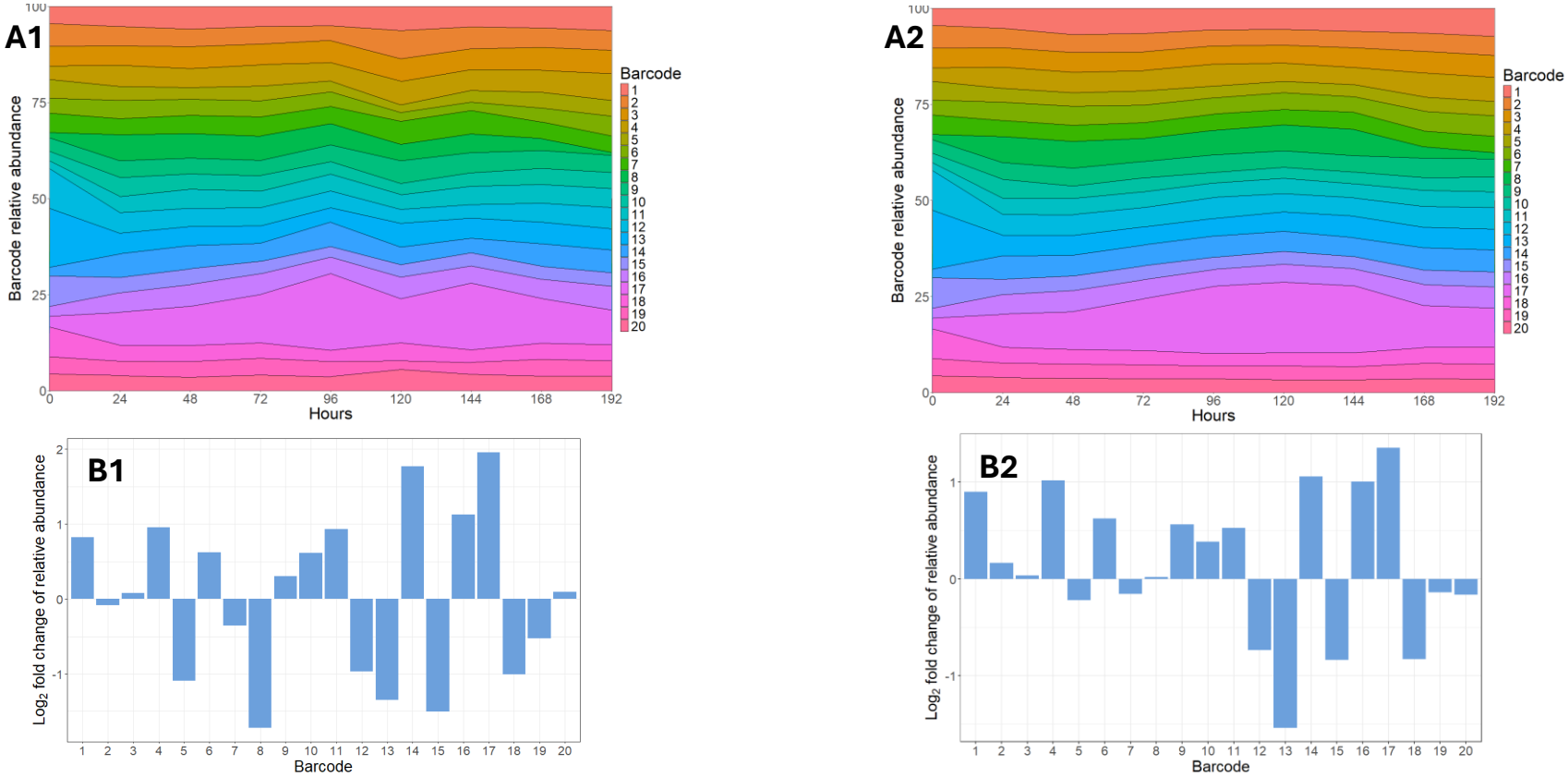


Figure 4.14. Scaling fermentation up to industrial conditions does not significantly increase strain variability on fermentations 1 and 2. (A1, A2) A 7.5 L fermentation of *S. clavuligerus* in S2A media was grown for 48 hours, and 800 mL was used as an inoculum for 7.5 L of CM5 media, where it was grown for an additional 120 hours, mimicking industrial multistage fermentation. Fermentations were paddle-driven and agitated at a maximum speed of 700 rpm, with pH monitored and maintained at 6.8. **(B1, B2)** Log_2 fold change in barcode relative abundance at hour 168 compared to hour 0 shows minimal variation between strains.

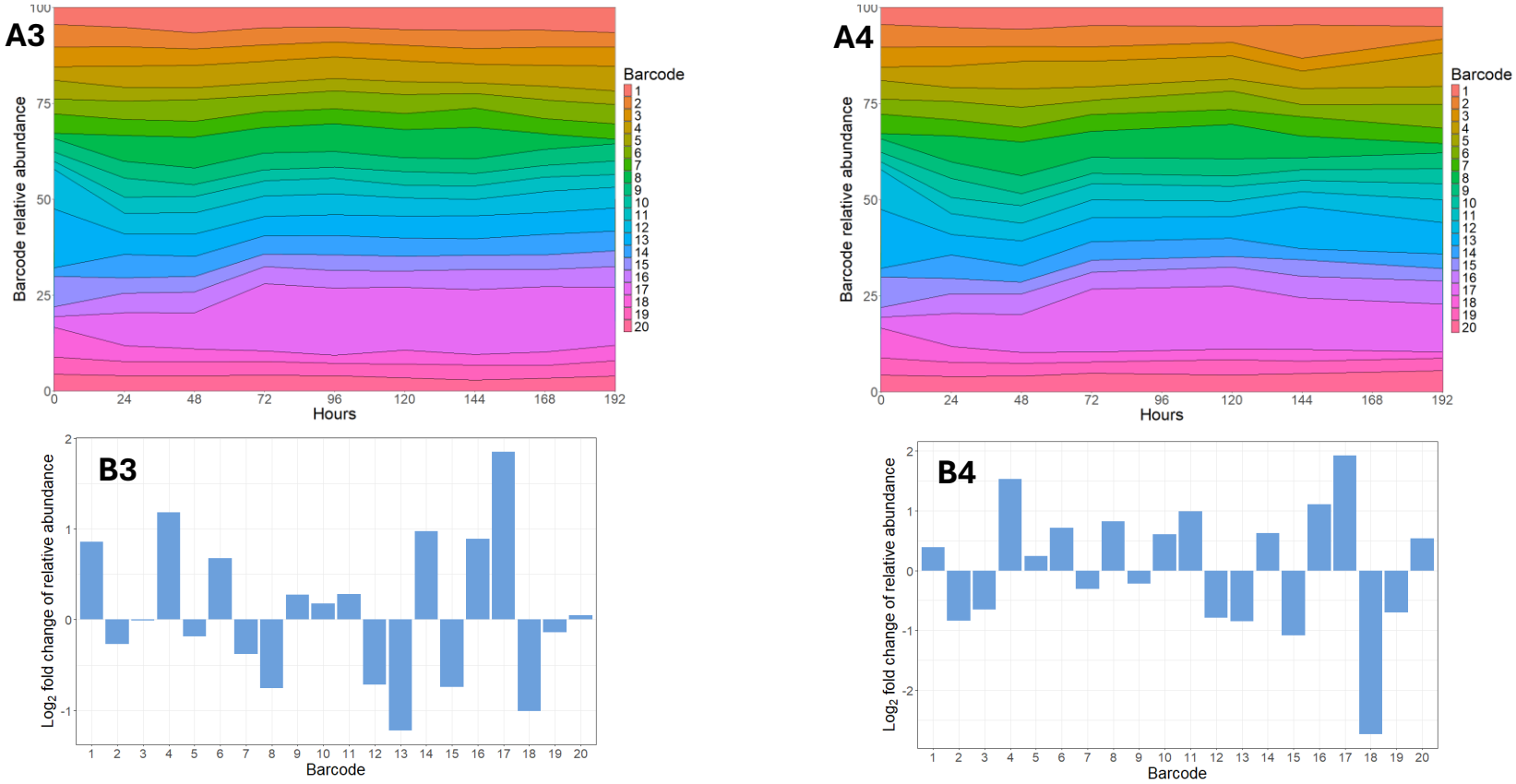


Figure 4.15. Scaling fermentation up to industrial conditions does not significantly increase strain variability on fermentations 3 and 4. (A3, A4) A 7.5 L fermentation of *S. clavuligerus* in S2A media was grown for 48 hours, and 800 mL was used as an inoculum for 7.5 L of CM5 media, where it was grown for an additional 120 hours, mimicking industrial multistage fermentation. Fermentations were paddle-driven and agitated at a maximum speed of 700 rpm, with pH monitored and maintained at 6.8. **(B3, B4)** Log_2 fold change in barcode relative abundance at hour 168 compared to hour 0 shows minimal variation between strains.

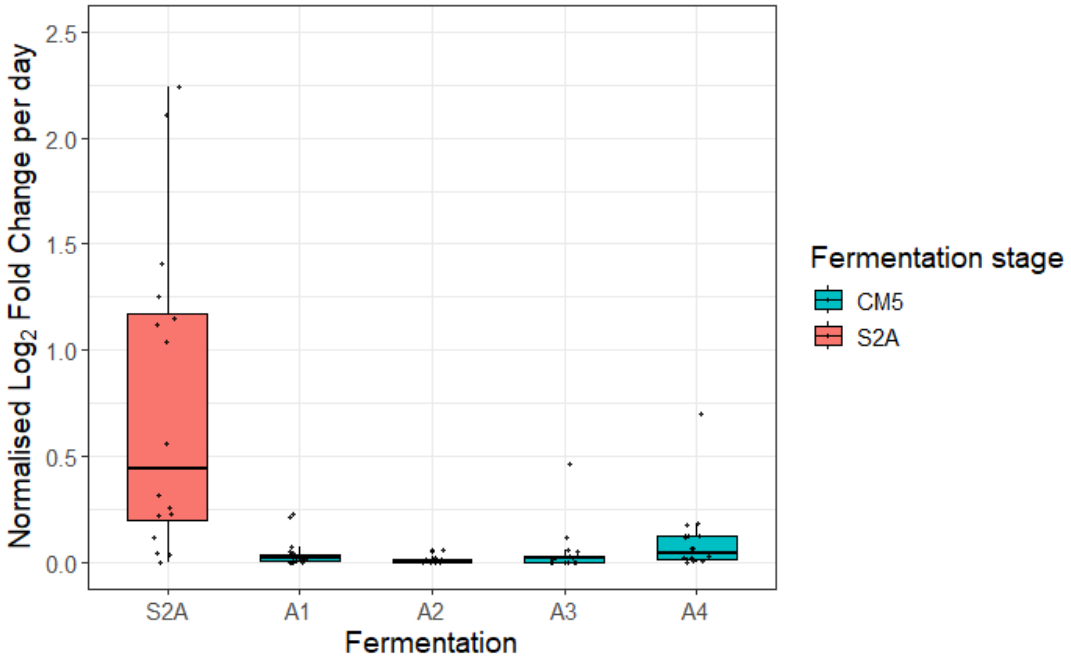


Figure 4.16. Normalised squared log_2 fold change per day across S2A and CM5 stages for four 7.5 L multistage fermentations. Each point represents the squared log_2 fold change in barcode abundance normalised to days of growth (2 days for S2A, 5 days for CM5). Normalisation was applied uniformly across all fermentations. The mean normalised log_2 fold change was 0.41 for S2A and 0.29 for CM5. For clarity, the y-axis of the normalised panel is capped at 2.5 (excluding seven data points). All data points, including outliers, were retained for statistical analysis. When considering the four replicates as one a Welch's t-test shows there is a significant different between S2A and CM5 stages of the industrial scale fermentation with a p value of 2×10^{-4}.

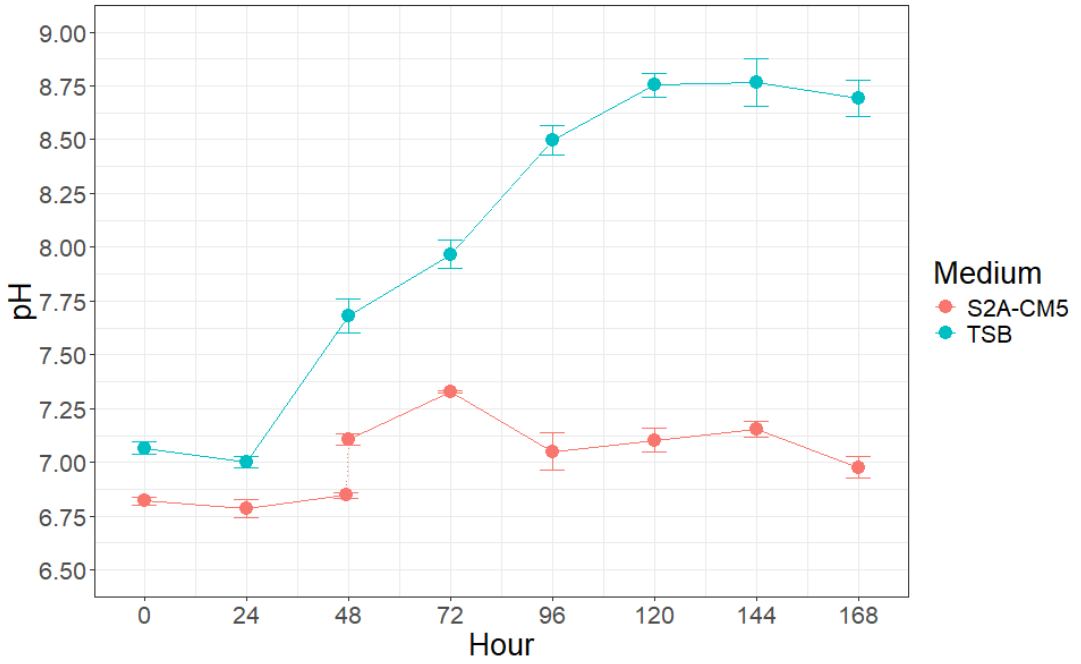


Figure 4.17. pH changes during 50 mL fermentations of *S. clavuligerus* in TSB and S2A-CM5 media. pH was measured at regular intervals throughout 168-hour fermentations in TSB or S2A-CM5. In TSB, pH increased from 7.0 to 8.8 over the course of the fermentation, whereas in S2A-CM5 pH remained stable between 6.8 and 7.3. Error bars correspond to SDs (n = 3). The dotted line between the two S2A-CM5 data at hour 48 represent the transition between the two media.

4.8 Bar-seq for competition assays

Following the determination that Bar-seq is suitable for the study of population dynamics in industrial fermentations, it was explored if Bar-seq could be used to study inter-strain competition assay. A key requirement within a competition assay is to be able to distinguish strains with a simple phenotypic test, for example to count colonies on different discriminatory media to determine differential fitness (Finkel and Kolter, 1999). One way to do this is to use antibiotic markers or auxotrophic strains (Lenski et al., 1998) but this can have potential fitness costs associated with carriage of the marker gene and selection (Harmand et al., 2018). However, if two or more strains of the same species are being competed against each other this phenotypic differentiation may not be possible. In this case, the only way to elucidate the quantities between two strains is tagging the genomes of those strains with some sort of unique marker and using this to quantify the abundance of each strain within a population. Therefore, the ability of Bar-seq to track the relative abundance of two strains of the same species was explored within the context of a competition assay.

The industrial strains *S. clavuligerus* SC2 and *S. clavuligerus* SC6 were used within a competition assay. As discussed previously *S. clavuligerus* SC2 is a precursor industrial strain to *S. clavuligerus* SC6 following multiple rounds of random mutagenesis (Benjamin Huckle, Personal communication). The iterative selection for *S. clavuligerus* strains during the CA development process for increased CA production has also resulted in adaptation to industrial media (Munnoch et al., in prep). Therefore, *S. clavuligerus* SC2 is less adapted to industrial media than *S. clavuligerus* SC6, and also has a greater specific growth rate than SC6, which has been selected through more rounds of mutagenesis for its ability to grow and produce CA is industrial media.

To explore this, two versions of each industrial strain were created. *S. clavuligerus* SC2 had barcodes JCb1 and JCb2 introduced via conjugation. The previously made strains of *S. clavuligerus* with JCb1 and JCb2 were also used in this experiment. To ensure that the barcode was not inferring with a specific aspect of strain fitness reciprocal experiments were performed. S. *clavuligerus* SC2 JCb1 was competed against *S. clavuligerus* SC6 JCb2. In a separate competition *S. clavuligerus* SC2 JCb2 was competed with *S. clavuligerus* SC6 JCb1. As with the previous experiments involving Bar-seq each strain was grown for 48 hours in 50 mL TSB and mixed in equal proportions based on OD_{600} measurements. The mixture of the two strains was further grown for 48 hours in 1 L TSB, with genomic DNA being extracted at the start and end of the fermentations to enable barcode relative abundance to be determined.

Both combinations of barcodes show similar results demonstrating that the barcodes themselves do not infer a fitness advantage on either strain. In both fermentation *S. clavuligerus* SC2 becomes dominant throughout the fermentation ending up at a final relative abundance of 98.8 and 94.9% for fermentation A1 and A2 respectively (Figure 4.18).

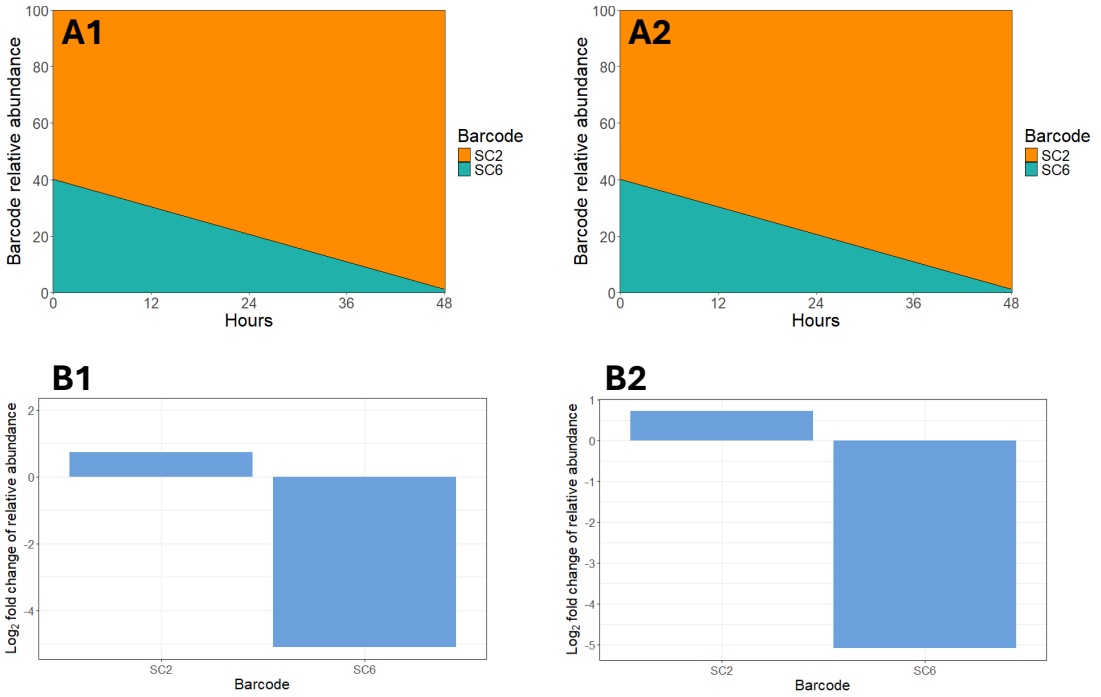


Figure 4.18. Bar-seq can be used to track competition between *S. clavuligerus* SC2 and SC6. **(A1)** Barcode JCb1 was used to track SC2, and **(A2)** barcode JCb2 was used to track SC6. Strains were mixed and fermented in 1 L TSB for 48 hours. **(B1, B2)** Log_2 fold changes in barcode relative abundance show SC2 versus SC6 and SC6 versus SC2, respectively. The data reveal that SC2 consistently outcompetes SC6 under these conditions.

To build on these findings and further explore both inter-strain and intra-strain competition, five barcoded variants of each industrial strain (*S. clavuligerus* SC2 and SC6) were constructed. Barcodes JCb1-10 were used with JCb1-5 being used for one strain and then JCb6-10 being used for the opposing strain in that competition. Each barcode was introduced to *S. clavuligerus* via conjugation into an otherwise isogenic background, meaning that within each strain (*S. clavuligerus* SC2 or SC6), the only difference between the five barcoded strains was the barcode itself, as before. As with previous experiments, strains were grown for 48 hours in 50 mL TSB, then mixed in equal proportions based on OD_{600} and co-cultured in 1 L TSB for 48 hours. To control for barcode-specific effects, a second fermentation was set up in which the barcodes were swapped between strain backgrounds.

Across both biological replicates, SC2 again became dominant, reaching final cumulative relative abundances of 96.7% and 89.2% respectively (Figures 4.19 and 4.20). These results are consistent with the previous two-barcode competition assay (Figure 4.18), where SC2 reached 98.8% and 94.9% when using a single barcode for each strain. This confirms that SC2 is consistently more fit than SC6 in TSB, regardless of which barcode is used, and further validates the use of multiple barcodes per strain for tracking.

Interestingly, this experimental design also revealed evidence of intra-strain competition. Within the *S. clavuligerus* SC2 group, some barcoded variants increased in relative abundance over the course of the fermentation, while others decreased. Despite *S. clavuligerus* SC2 becoming dominant, not all five *S. clavuligerus* SC2 barcodes followed the same trajectory.

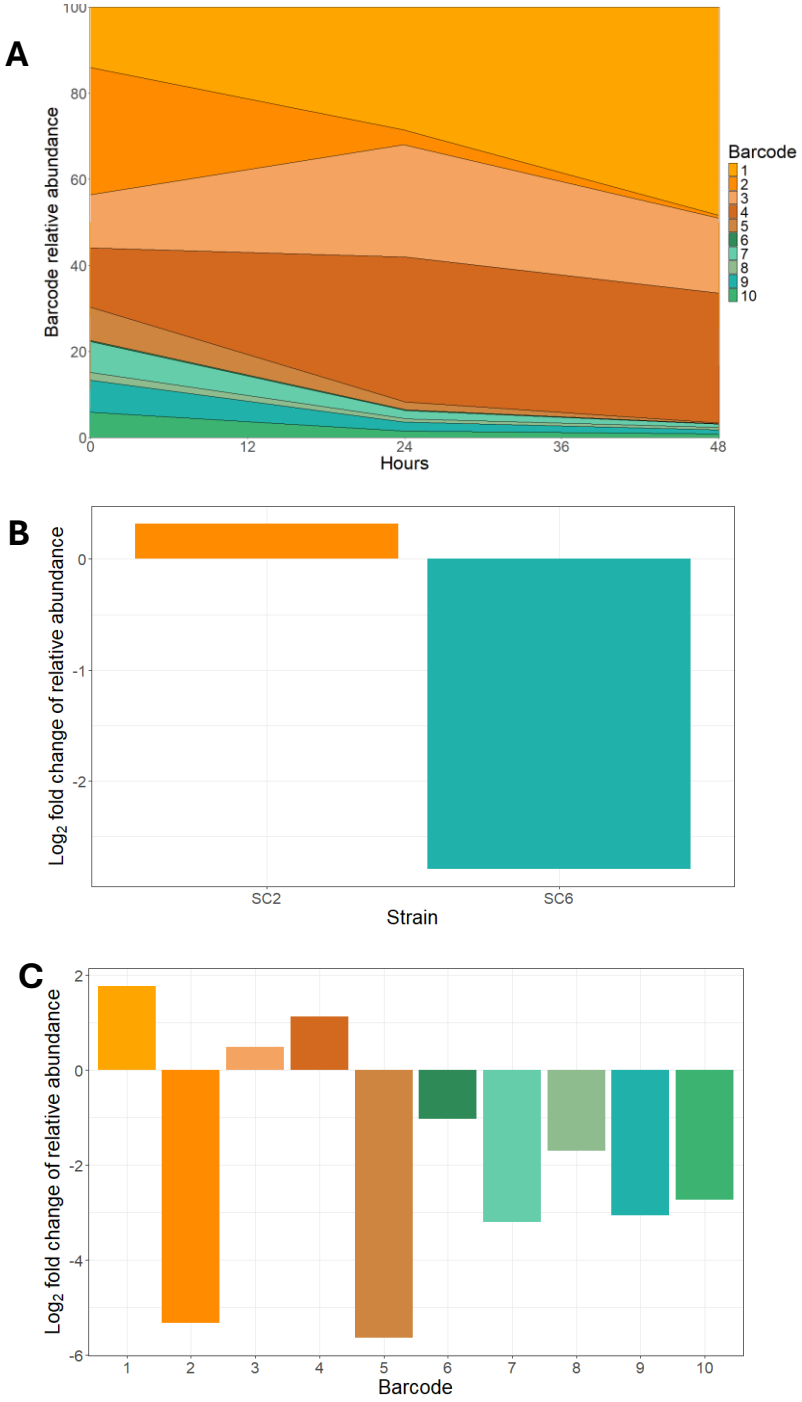


Figure 4.19. When multiple barcodes are used per *S. clavuligerus* strain, intra-strain competition can be observed. (A) Barcodes JCb1-5 were conjugated into otherwise isogenic strains of SC2, and barcodes JCb6-10 were conjugated into otherwise isogenic strains of SC6. All 10 strains were mixed in equal abundance and fermented in 1 L TSB for 48 hours. **(B)** The $\log_2$ fold change of the cumulative relative abundance for both strains was plotted, showing SC2 significantly outcompetes SC6, as is consistent with previous experiments. **(C)** Intrastrain competition is observed within SC2. While all SC6 barcodes decrease in relative abundance, unexpectedly, two of the five SC2 barcoded strains also decrease in abundance. This suggests that the overall increase in SC2 abundance is attributable to three of the five barcoded strains and demonstrates variation even within strains.

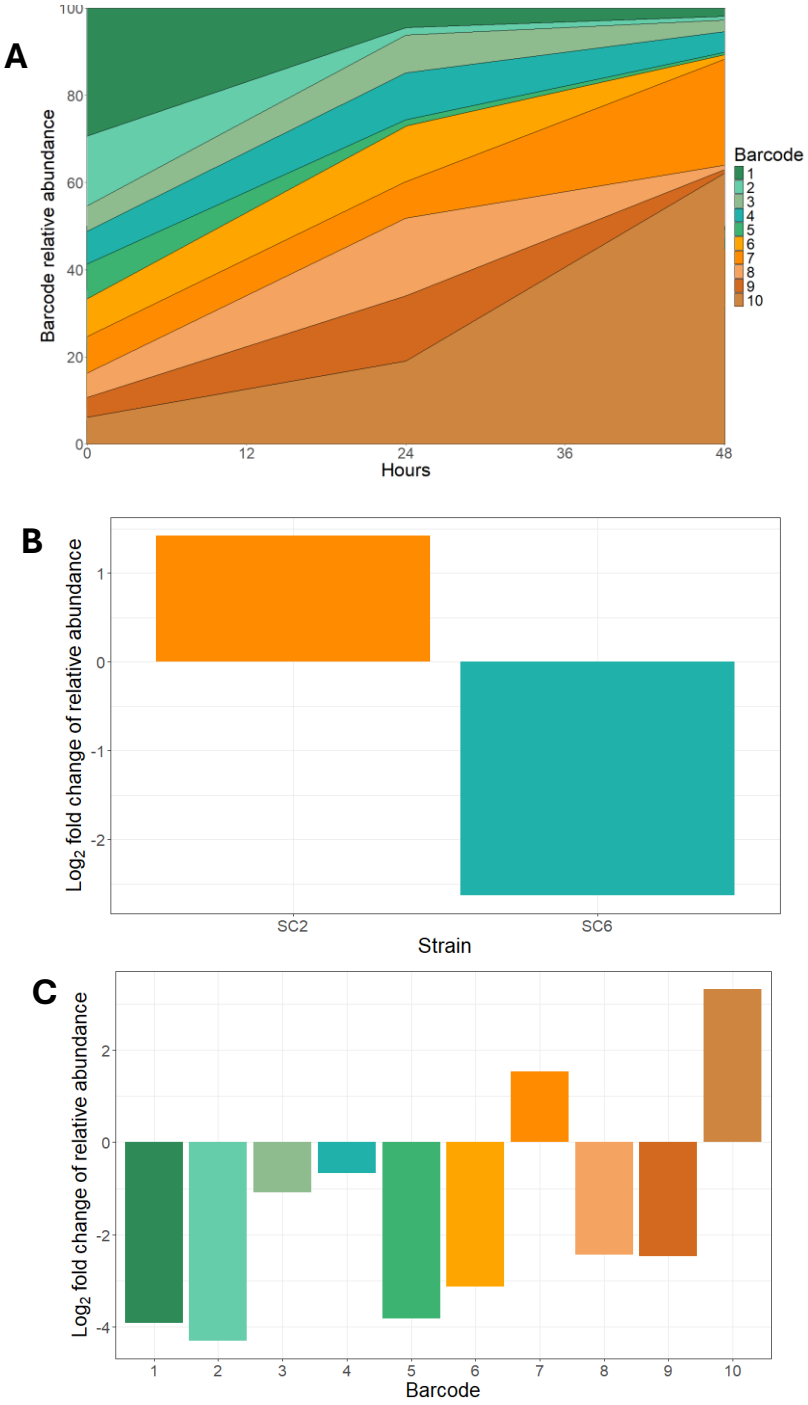


Figure 4.20. When multiple barcodes are used per strain, intrastrain competition can be observed. (A) Barcodes JCb1-5 were conjugated into otherwise isogenic strains of SC6, and barcodes JCb6-10 were conjugated into otherwise isogenic strains of SC2. All 10 strains were mixed in equal abundance and fermented in 1 L TSB for 48 hours. **(B)** The $\log_2$ fold change of the cumulative relative abundance for both strains was plotted, showing SC2 significantly outcompetes SC6, as is consistent with previous experiments. **(C)** Intrastrain competition is observed within SC2. While all SC6 barcodes decrease in relative abundance, unexpectedly, three of the five SC2 barcoded strains also decrease in abundance. This suggests that the overall increase in SC2 abundance is attributable to two of the five barcoded strains and demonstrates variation even within strains.

4.9 Summary

This chapter aimed to assess how well Bar-seq can track population dynamics under increasingly relevant industrial fermentation conditions.

Bar-seq was first tested in a 72-hour, 50 mL TSB fermentation to check whether extending the timescale introduced additional instability (Figure 4.1). Variation was slightly reduced compared to the 48-hour fermentations described in Chapter 3, suggesting that Bar-seq remains reliable over longer timeframes and that instability is not strictly time-dependent.

Scaling up to 1 L fermentations in TSB led to much greater variability in barcode abundance between the start and end points (Figure 4.2), with replicate-to-replicate variation also becoming more pronounced. Despite this, no strain reached dominance (>50%), and most barcodes remained detectable, showing that even with increased instability, the full population could still be resolved.

Introducing a multistage setup using S2A and CM5 media, a more industrially relevant approach, led to much more stable barcode dynamics. Both 50 mL and 1 L multistage fermentations showed limited changes in barcode abundance (Figures 4.3 and 4.7), and even across six biological replicates, batch-to-batch variation remained low (Figures 4.8 and 4.9). These findings show that industrial media formulations help stabilise strain composition, possibly due to different selective pressures or reduced growth rates in CM5. Additionally, the increased stability within industrial media correlated with media-dependent pH changes. TSB fermentations became increasingly alkaline, whereas CM5 maintained a stable pH range (Figure 4.17).

However, across both the 1 L multistage shake flasks and the 7.5 L stirred-tank bioreactor fermentations, the S2A seed stage consistently showed more instability than

the CM5 production stage when normalised to time (Figures 4.13 and 4.16). This was observed even when pH, agitation, and aeration were tightly controlled (Figures 4.14 and 4.15), suggesting that the nature of the seed medium, rather than physical conditions, is a key contributor to early population shifts. Once in CM5, populations remained stable over several days, and no single barcode came to dominate in any fermentation.

Taken together, these experiments show that Bar-seq is not only suitable for tracking complex mixed populations under laboratory conditions, but also robust enough for use across a range of fermentation scales and industrial media types. It consistently resolved population structure without introducing bias (Figure 4.10) and was sensitive enough to detect both inter- and intra-strain competition. This was further supported by competition assays, where SC2 consistently outcompeted SC6 regardless of the barcodes used (Figures 4.18–4.20), and intra-strain variation was observed when multiple barcodes were introduced per background.

Finally, a comparison of overall instability across conditions (Figure 4.21) highlights that the highest variation occurred in TSB at 1 L scale, and the lowest in CM5-based multistage fermentations. These findings emphasise that Bar-seq is well-suited for assessing strain stability in industrial fermentations and provides a useful tool for evaluating how strain mixtures behave during scale-up.

Figure 4.21. Summary of change in absolute relative abundance across different medias and scales.

Chapter 5 Long term evolution of barcoded strains

5.1 Introduction

During industrial fermentations, populations of *S. clavuligerus* are subjected to extended growth in nutrient-limited environments designed to maximise CA production (López-Agudelo et al., 2021; Schmidt, 2005). However, these environments also impose strong selective pressures that may drive genetic and phenotypic changes over time (Harms et al., 2017; Hoskisson and Hobbs, 2005). Spontaneous mutations are well-characterised mechanisms of genetic change in bacteria (Martinez and Baquero, 2000), and *Streptomyces* species are known to be particularly susceptible in this regard because of their large, G+C-rich genomes, high capacity for recombination and linear genomes (Lerminiaux and Cameron, 2018; Zhang et al., 2022). In natural environments, this evolutionary adaptability enables resilience and survival (Ensign, 1978; Williams et al., 1972). In contrast, in an industrial setting, it may lead to instability and reduced consistency of product formation (Rugbjerg et al., 2018).

While fermentations completed in the previous chapters revealed media and scale dependent changes in barcode abundance, they were not sufficient to assess the long-term evolutionary potential of barcoded *S. clavuligerus*. Additionally, the lack of population dynamics within industrial media over an industrially relevant timescale, suggested the anecdotal links between genotypic and phenotypic differences leading to non-production in industrial fermentations may not be as obvious as originally purported.

Previous work, most notably the Long-Term Evolution Experiment (LTEE) in *Escherichia coli* (Lenski et al., 1991), has shown that extended culture under a single condition can lead to adaptation, with the fixation of specific mutations and the emergence of dominant lineages. Whether similar dynamics occur in *S. clavuligerus* under industrial fermentation conditions remains unknown.

This chapter investigates the evolutionary potential of barcoded *S. clavuligerus* strains over extended serial passage in industrially relevant CM5 medium. Barcoded strains were grown over 70 days with regular passaging, and barcode abundance was tracked over time to identify changes in population structure. Strains that became dominant were selected for further characterisation to identify putative adaptive mutations and assess their impact on product yield. This approach provides insight into the evolutionary forces acting during prolonged culture and the extent to which these changes affect industrially important phenotypes such as antibiotic production.

All experimental evolution work within this chapter (e.g. serial passaging, sampling, phenotyping etc.), sequencing, analysis, and figure preparation in this chapter were performed by James Croxford. Genome sequencing analyses in this chapter used the annotated *S. clavuligerus* SC6 reference genome generated by John Munnoch (DNA isolation, sequencing, genome assembly, and annotation), as described in Section 2.3 of this Thesis, which was used as the reference for downstream analyses. Transcriptomic analyses additionally used this annotated reference genome generated by John Munnoch.

5.2 The evolution experiment

To investigate adaptation of *S. clavuligerus* under industrial conditions, an evolution experiment was performed using the 20 barcoded strains of *S. clavuligerus* SC6. While previous fermentations in 1 L of multistage industrial media showed significant shifts in population dynamics, a 7.5 L bioreactor did not yield similar changes. This suggested that either the environment was insufficiently selective or that adaptation occurs on longer timescales. To explore this further, an experiment was designed in which

adaptation was studied on longer timescales to test if strains may emerge within the population that may then go on to be dominant, as a result of adaptations. The design was conceptually modelled on the Long-Term Evolution Experiment with *E. coli* (Lenski et al., 1991).

Barcoded strains of *S. clavuligerus* SC6 pSET152 JCb1–20 were grown in TSB for 48 hours and then mixed in equal proportions (2.7×10^7 CFUs in total, equivalent to 5 mL of 5.4×10^6 CFUs/mL within the inoculating culture) based on OD_{600} readings. This inoculum was used to seed 50 mL of CM5 medium in triplicate, generating three independent evolution experiments (A, B and C). Cultures were grown for five days, then serially passaged by transferring 10% volume into fresh CM5 medium. This process was repeated every five days for 70 days (14 passages).

Genomic DNA was extracted at days 25, 40, 50, 60, and 70. Barcoded regions were amplified and sequenced as described previously to quantify the relative abundance of each strain. A Muller plot was generated to visualise the dynamics of barcoded strain populations over time (Figure 5.1). Additionally, $\log_2$fold changes in abundance were calculated for each strain at day 70 relative to the starting abundance.

The doubling time of *S. clavuligerus* during the evolution experiments was estimated using two independent methods, due to the variability in CFU counts observed in CM5. First, CFU counts were taken at the start and end of a five-day passage in each of the three experiments, in triplicate, yielding an average doubling time of 10.63 hours across replicates. Second, a Bradford assay was used to measure total protein concentration in cultures sampled every 24 hours over five days. Exponential phase was identified, and the corresponding growth rate indicated a doubling time of 11.58 hours. Averaging these

two estimates gave a final doubling time of 11.11 hours, suggesting that approximately 152 generations occurred over the 70-day course of the evolution experiment.

Despite identical starting inoculum, the three biological replicates diverged markedly in their population trajectories. In experiment A, six strains increased in abundance, while one and seven strains did so in experiments B and C, respectively. At day 70, experiment A contained 10 strains each at <1% relative abundance, while JCb13 and JCb2 represented 24.20% and 36.62%, respectively. Notably, JCb2 peaked at 55.90% at day 60 before declining in relative abundance.

Experiment B showed a markedly different pattern, with JCb10 reaching 90.96% relative abundance by day 70. Apart from JCb2 (2.04%), all other strains finished at below 1%. The expansion of JCb10 began after day 40; at that earlier timepoint, JCb12 was dominant (31.84%) but subsequently declined.

Experiment C exhibited a more stable profile, with multiple strains maintaining moderate abundance. At day 70, JCb4, JCb13, JCb14, JCb15, and JCb18 were present at 12.11%, 14.29%, 23.06%, 19.63%, and 4.36%, respectively. JCb14 remained consistently abundant throughout the experiment. Only five strains fell below 1% abundance, though fluctuations were still observed over time.

The differing dynamics between replicates suggest there is no bias between barcoded strains. Experiment B was selected for further characterisation in subsequent analyses, as it resulted in the clear dominance of a single barcoded strain, suggesting this dominant strain had successfully adapted to its environment.

CA production was also monitored across all experiments. CA assays were performed every five days, with three replicates per timepoint. As before, CM5 medium prevented dry weight normalisation, so results were reported as µg/mL only. At day five, CA titres

reached 203.94 µg/mL, 243.00 µg/mL, and 195.45 µg/mL respectively. Over time, all three showed a consistent decline. By day 70, CA production had dropped to 69 µg/mL, 88 µg/mL in experiments A and C respectively, and just 25 µg/mL in experiment B.

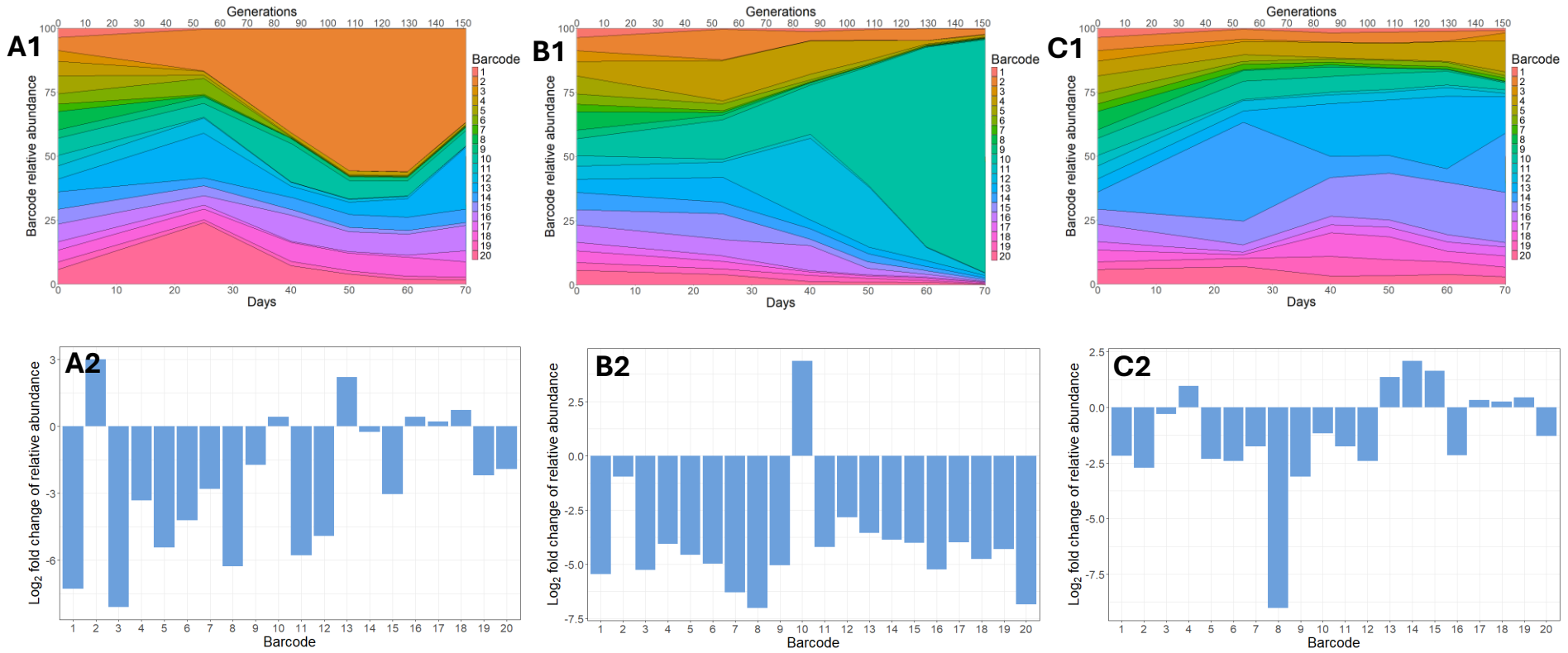


Figure 5.1. Long-term evolution of industrial *S. clavuligerus* shows inconsistent strain dominance. (A1, A2, A3) Barcode relative abundances were tracked in 50 mL CM5 fermentations over 70 days. Cultures were grown for five days, passaged into fresh CM5, and repeated for 70 days. **(B1, B2, B3)** Log_2 fold change in strain abundance varied across biological replicates, with six, one, or seven strains increasing in relative abundance in fermentations A, B, and C, respectively.

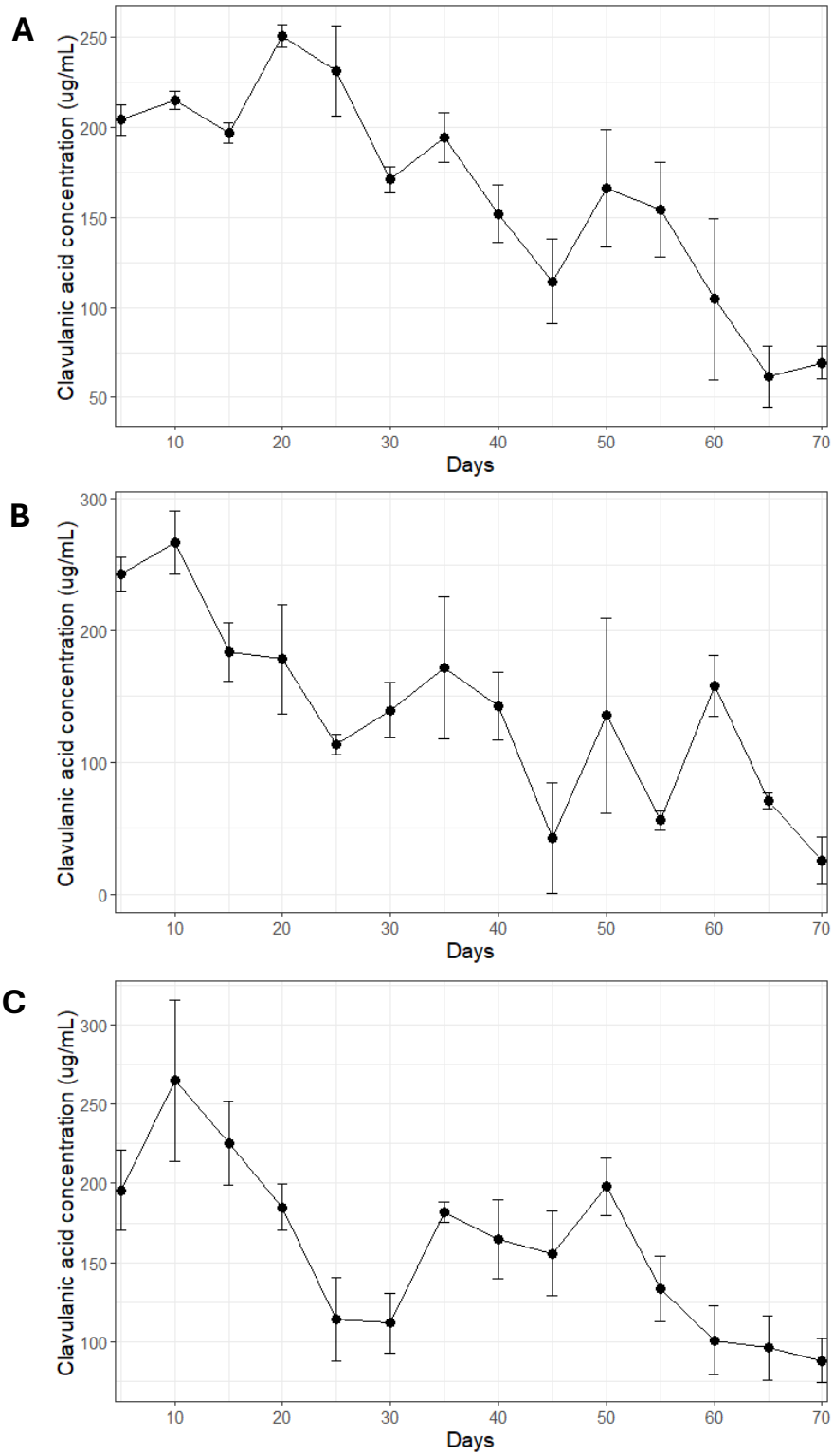


Figure 5.2. Clavulanic acid levels decrease during long-term strain evolution. Triplicate samples were taken every five days from the fermentations shown in Figure 5.1. Clavulanic acid (CA) was quantified in each sample; results could not be normalised to cell dry weight due to high medium viscosity and particle content. Error bars indicate standard deviation of triplicates. After 70 days, fermentations A and C retained CA levels of 69 µg/mL and 88 µg/mL, respectively, while fermentation B dropped to 25 µg/mL.

5.3 Analysis of evolution experiment B

As evolution experiment B was selected for further analysis, a method was required to distinguish colonies belonging to the dominant barcoded strain (JCb10) from other strains. To achieve this, PCR primers were designed with the forward primer (JCb10 check fwd) targeting the unique barcode region of JCb10, and the reverse primer (JCb10 check rev) located 1 kb upstream. The design accounted for the genomic integration of pSET152, such that the primers would face away from each other in the circular plasmid, but toward each other once integrated, allowing for specific amplification only from integrated pSET152 JCb10.

To ensure the designed primers amplified only DNA from JCb10, and no other barcode, a validation PCR was performed. Genomic DNA from all 20 barcoded strains were tested using the primers JCb10 check fwd and rev, only genomic DNA from JCb10 amplified and created a band at the expected size (Figure 5.3). As a result, it was confirmed that these primers were appropriate for the detection of JCb10.

Following validation, genomic DNA was extracted from 96 colonies isolated from evolution experiment B. PCR was performed using the JCb10 check fwd and rev primers. A subset of the resulting PCR products (colonies 25–36) were analysed via agarose gel electrophoresis (Figure 5.4). Amplification was observed in all samples except colony 27, which lacked the expected band. Barcode sequencing confirmed that this colony was JCb6, a minor strain within the population at day 70. For downstream genotypic and phenotypic characterisation, five confirmed JCb10 colonies and the single JCb6 isolate (colony 27) were selected for further analysis.

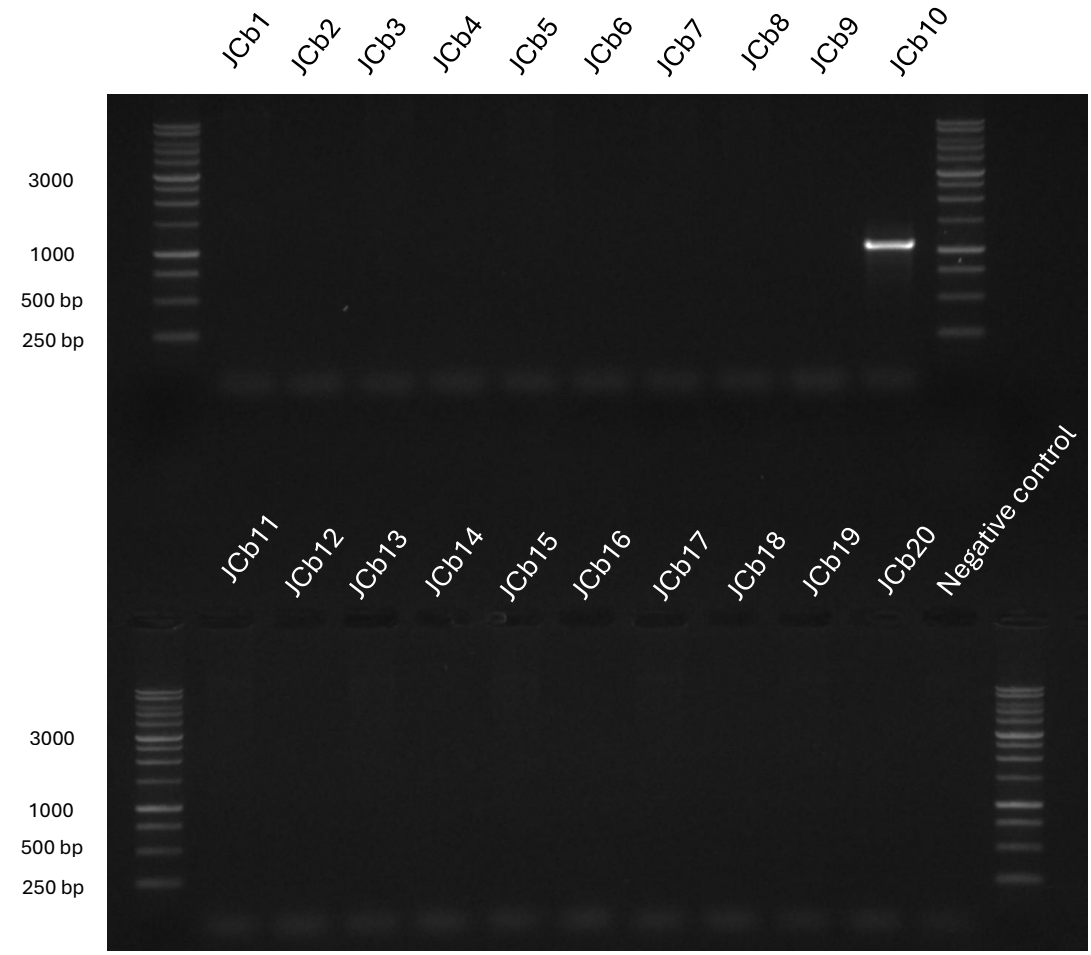


Figure 5.3. Validation that primers JCb10 check fwd and rev only amplify JCb10 genomic DNA. Genomic DNA was extracted from *S. clavuligerus* SC6 pSET152 JCb1-20 and PCR was performed using primers JCb10 check fwd and rev. Amplification of JCb10 DNA shows specificity of the primers for this strain. Negative control was nuclease free water.

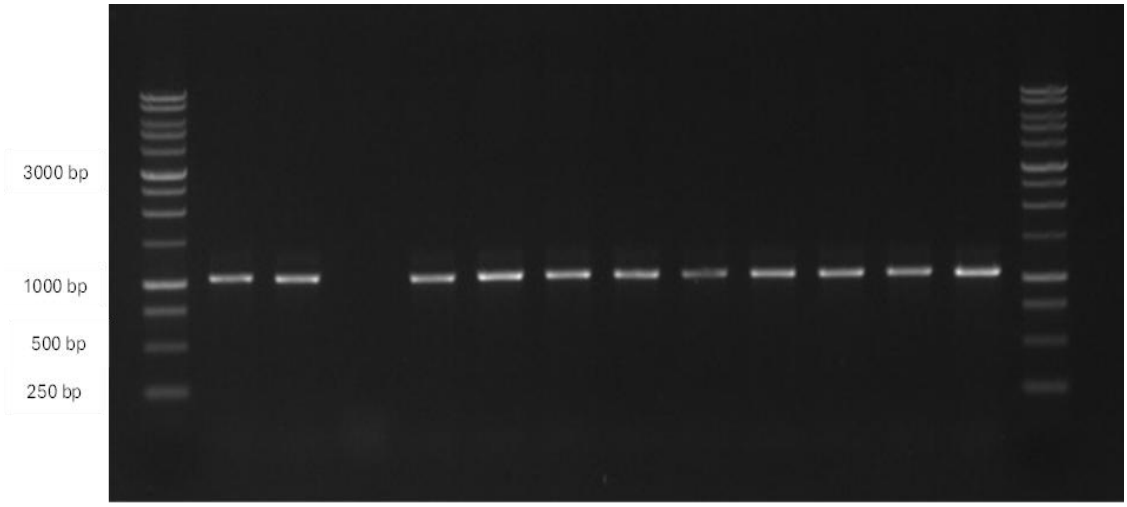


Figure 5.4 Identification of a rare strain within evolution experiment B using JCb10-specific PCR. Genomic DNA was extracted from 96 colonies isolated from evolution experiment B, and PCR was performed using primers JCb10 check fwd and rev. A subset of colonies (25–36) was run on an agarose gel, showing amplification in all except colony 27. The absence of amplification in colony 27 indicated it was not JCb10, subsequent barcode sequencing confirmed this colony as JCb6.

5.4 Genotyping of evolution experiment B

Following confirmation of barcodes from colonies isolated during evolution experiment B, whole-genome sequencing was performed to characterise their mutational profiles. This analysis aimed to explore the evolutionary capacity of *S. clavuligerus* within industrial media and to compare genotypic differences between dominant (JCb10) and non-dominant (JCb6) barcoded strains. The five randomly selected colonies identified as JCb10 were labelled A–E for downstream analysis. The non-dominant colony, identified as JCb6, was referred to as the rare isolate (JCb6).

All six evolved strains (A–E and JCb6), along with three reference strains (JCb6, JCb10, and JCb17 from parental strain stocks), were grown in TSB for 48 hours. Genomic DNA was extracted and sequenced using the Illumina NovaSeq X Plus platform at PE150 (Novogene UK). Mutations were identified using the Breseq pipeline (Deatherage and Barrick, 2014), with comparisons made against the *S. clavuligerus* SC6 reference genome (BioProject PRJNA1127551) . Mutations present in the reference barcoded strains were used as a background control for subsequent analyses due to the presence of SNP mutations within the chromosome and pSCL4 of the parental barcoded strains compared to the SC6 reference genome. Although all work was carried out with *S. clavuligerus* SC6, transcriptomic data were mapped to the SC2 annotated genome (BioProject PRJNA1127551) because this reference provides more complete annotation while still containing all genes present in SC6. As a result, all transcript identifiers are presented as SC2 locus tags throughout the following section.

A complete deletion of plasmid pSCL3 was observed in four of the five dominant isolates (B–E) and in the rare (JCb6) strain. In addition, three non-synonymous mutations were consistently found across all dominant isolates (A–E; Table 5.1). These affected genes include an ATP-binding protein (SCLAV SC2 10490, (TGCGGGG)2→1), responsible for the

use of ATP to power ABC transport proteins (Baichwal et al., 1993). Another non-synonymous mutation is to a bifunctional diguanylate cyclase/phosphodiesterase (SCLAV SC2 21505, Δ5 bp), an important gene in the regulation of the global regulator ci-di GMP, of which knockout strains have shown to have an increased life-cycle and a reduced secondary metabolism (Gallagher et al., 2024; Nesbitt et al., 2015). The final non-synonymous mutation was seen in a zinc metalloprotease HtpX (SCLAV SC2 17525, 58 bp→T), which is a heat shock protein that has implications in membrane protein folding and assemblies (Dalbey et al., 2012). All three mutations were located on the chromosome and were assumed to cause a complete loss of function. Four additional mutations were detected in individual dominant isolates:

- Strain B: a homopolymer deletion, (G)10→9, in a non-coding chromosomal region. The homopolymer starts 6bp downstream of the ABC transporter protein (SCLAV SC2 16920), potentially affecting transcription.
- Strain C: a homopolymer deletion, (G)10→9, within a CHAT domain-containing protein on pSCL4 (SCLAV SC2 00150). Resulting in a nonsense mutation of a protein with an unknown function.
- Strain D: a substitution (C→G, Asp→Glu) in a tetratricopeptide repeat protein on pSCL2 (SCLAV SC2 28310), which itself may play a role in protein-protein interactions (Cerveny et al., 2013).
- Strain E: a substitution, G→A, in a non-coding region on pSCL2. The substitution occurs within the 'right hand' telomere of pSCL2.

The rare JCb6 isolate also carried a deletion of the 455.3 kb plasmid pSCL3, along with a tandem repeat expansion, (GCTGCG)8→9, in a non-coding chromosomal region. This expansion occurs 46bp upstream of a 3-oxoacyl-ACP reductase but 111bp downstream

of a Zn dependent alcohol dehydrogenase. The final mutation is a synonymous substitution, C→T, in a serine/threonine protein kinase (SCLAV SC2 14025; Table 5.1). Compared to the dominant isolates, the JCb6 genome contained fewer mutations and no shared alterations with strains A–E beyond the loss of pSCL3.

Although the mutations observed across the six evolved isolates do not converge on a single metabolic pathway, they are broadly associated with regulation, developmental, and stress response functions. No mutations are directly implicated in primary metabolism.

Table 5.1. Mutations detected in genomes of parental, dominant, and rare strains after evolution. Five colonies of the evolution experiment B JCb10 (dominant strain), one colony of evolution experiment B JCb6 (rare strain), and three parental strains were grown in TSB for 48 hours. Genomic DNA was extracted and sequenced, and Breseq was used to identify mutations relative to the parental genome. Unique mutations were detected in most evolved strains.

Replicon	Position	Mutation	Gene	Present in
Dominant strain (JCb10)				
Chrm	2,376,315	(TGCGGGG)2→1	ATP binding protein	All
Chrm	4,925,454	Δ5 bp	Bifunctional diguanylate cyclase/ phosphodiesterase	All
Chrm	4,005,502	58 bp→T	Zinc mettaloprotease	All
pSCl3	1	Δ444,894 bp	-	B, C, D, E
Chrm	3,862,021	(G)10→9	Non-coding	B
pSCl4	876,194	(G)10→9	CHAT domain containing protein	C
pSCL2	82,985	C→G	Tetratricopeptide repeat protein	D
pSCL2	149,351	G→A	Non-coding	E
Rare strain (JCb6)				
Chrm	1,638,383	(GCTGCG)8→9	Non-coding	-
Chrm	3,209,793	C→T	Serine/threonine protein kinase (synonymous)	-
pSCL3	1	Δ444,894 bp	-	-

Following the identification of a pSCL3 deletion in multiple evolved isolates, the presence and relative copy number of all four plasmids were assessed across sequenced strains. Plasmid content was investigated due to the known variability of plasmid copy number in *S. clavuligerus* and their role in biosynthetic gene cluster (BGC) carriage (Gomez-Escribano et al., 2021; Medema et al., 2010).

Illumina reads from parental strains (JCb6, JCb10, JCb17), dominant isolates (A–E), and the rare JCb6 isolate were mapped to the SC6 reference chromosome and plasmids using Bowtie2 (Langmead and Salzberg, 2012). Average coverage, including zero-coverage bases, were calculated for each replicon using Bedtools (Quinlan and Hall, 2010), and plasmid coverage was normalised to chromosomal coverage. Normalised plasmid copy numbers were plotted using the average of the three parental strains as a reference (Figure 5.5).

Consistent with the breseq results, strains B–E and the rare isolate showed near-complete loss of pSCL3. Although normalised coverage in these strains was not exactly zero (0.022 x), inspection with JBrowse (Diesh et al., 2023) revealed that mapped reads within these strains originated from pSCL3 orthologues present on the chromosome, confirming true loss of the plasmid. Strain A retained pSCL3 with a normalised coverage of 0.88 x, also consistent with breseq output. JBrowse analysis confirmed low but even coverage across the entire pSCL3 replicon in strain A, indicating retention at reduced copy number compared to the parental average (1.56 x)

Coverage of pSCL4 remained relatively stable across all strains. The parental strains had an average normalised coverage of 1.25x, compared to 0.98x in dominant strains and 1.20x in the rare strain.

The greatest variability in coverage between strains was observed for pSCL1. Normalised coverage averaged 6.63 x in parental strains, 3.90 x in dominant strains, and 7.38 x in the rare strain. Among dominant isolates, individual values ranged from 2.18 x (strain E) to 4.78 x (strain A).

pSCL2 displayed a consistent reduction in copy number across all evolved strains. Parental strains had an average normalised coverage of 9.63 x, compared to 3.27 x in dominant strains and 5.47 x in the rare strain.

To confirm that coverage drops were not due to structural variation, JBrowse was used to visualise coverage across pSCL1, pSCL2, and pSCL4. No large deletions or new junctions were detected in any strain, supporting the breseq result that no structural plasmid changes occurred beyond pSCL3 loss.

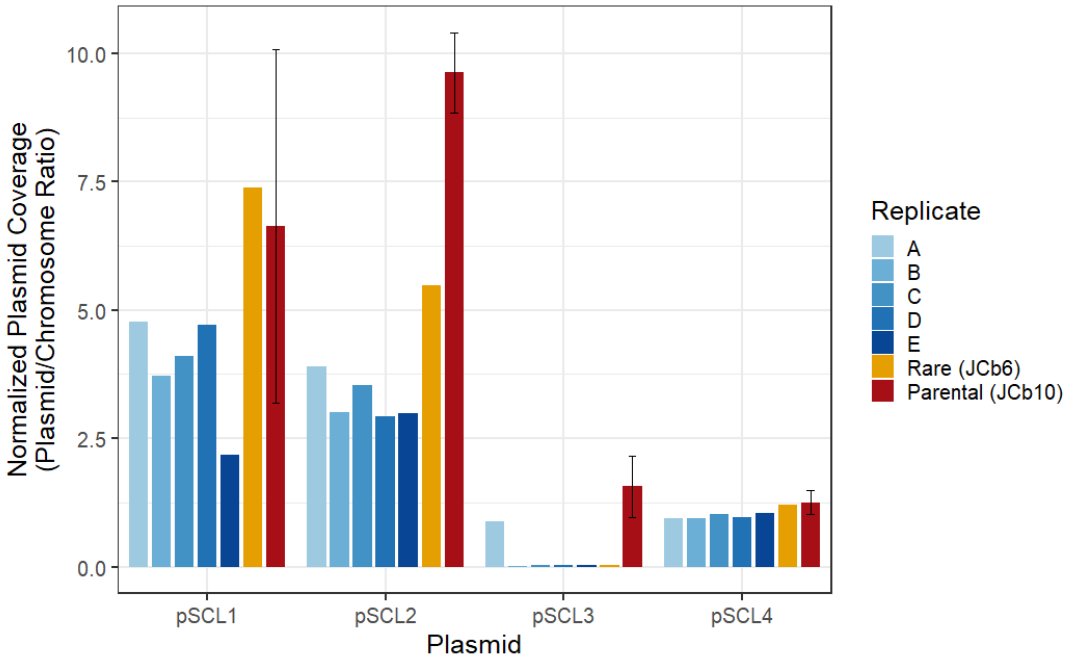


Figure 5.5. Plasmid copy number in parental, dominant, and rare strains after evolution. Five colonies of the evolution experiment B JCb10 (dominant strain), one colony of experiment B JCb6 (rare strain), and three parental strains were grown in TSB for 48 hours. Genomic DNA was extracted and sequenced, and plasmid-to-chromosome coverage ratios were calculated for pSCL1–4. Error bars on parental values represent standard deviation across the three replicates. In the dominant strain, all but one colony lost pSCL3 entirely. Levels of pSCL4 remained comparable to the parental strain, pSCL2 levels differed among groups, and pSCL1 levels were variable.

In addition to analysing plasmid content and structural variation, the potential circularisation of the chromosome was investigated. Previous work demonstrated that Cas9-mediated loss of pSCL4 in *S. clavuligerus* DSM738 was associated with chromosomal circularisation, potentially due to loss of the terminal proteins Tap-Tpg4 (Mohit, 2023). While loss of pSCL3 alone was not sufficient to cause circularisation in the previous study, it was proposed that cumulative loss of multiple *tap-tpg* loci could disrupt end-patching and lead to circularisation (Mohit, 2023). Given that the current study was performed in *S. clavuligerus* SC6, an industrial derivative, it was important to assess whether plasmid loss and reduced copy number might also cause chromosomal circularisation in this background.

Although no novel chromosomal junctions were identified by Breseq, three independent approaches were used to assess potential circularisation. First, a circularised version of the SC6 genome was generated in SnapGene, and a 50 bp sequence spanning the artificial chromosomal junction was extracted. This sequence was then queried against raw FASTQ reads from all sequenced strains using SeqIO.parse in BioPython (using the code available at https://github.com/jamescroxford/Bar-seq). No matches to the bridging sequence were detected.

Second, a 1.3 kb circular junction sequence was extracted from the SnapGene circularised genome and saved as a FASTA file (https://github.com/jamescroxford/Bar-seq). Reads from all sequenced strains were mapped to this region using Bowtie2, and the resulting BAM files were visualised in JBrowse. No reads mapped to the junction-spanning region, without mismatched mapping. Reads from both dominant strain E and the strain still possessing pSCL3 (A) were compared against each other. Read mapping did not differ between the two strains, again supporting the absence of circularisation.

Finally, a PCR-based assay was performed using primers designed to bind 500 bp upstream and downstream of the linear chromosomal ends. If the chromosome were circularised, these primers would amplify a 1 kb product, as previously described (Mohit, 2023). No amplification was observed in any strain, further confirming that the chromosome remained linear throughout the evolution experiment. No positive control was possible for this PCR reaction due to the lack of a strain containing a circularised chromosome.

5.5 Transcriptomic analysis of evolution experiment B

To investigate transcriptional changes associated with strain adaptation, RNA-seq was performed on the parental (JCb10), a representative dominant isolate (E), and the rare strain (JCb6). All strains were grown in TSB for 48 hours using the Cell Growth Quantifier to determine log and stationary phases. Following growth phase determination, each strain was regrown in triplicate and RNA was extracted at both timepoints. Libraries were created using an rRNA depletion approach and sequenced using the Illumina NovaSeq X Plus platform at PE150 (Novogene, UK).

Differential gene expression analysis was conducted using DESeq2, with $\log_2$ fold change (log2FC) values based on triplicate means. Nine pairwise comparisons were performed, covering intra-strain growth phase transitions, dominant vs. parental, dominant vs. rare, and rare vs. parental comparisons. Principal component analysis (PCA) on normalised gene expression confirmed clear clustering by condition, with biological replicates grouping closely and no clear outliers observed. Volcano plots were generated for each comparison (Figures 5.6 to 5.9). All annotated genes were used for analysis, and the 30 most significant differentially expressed genes (DEGs) (i.e. lowest

adjusted p-values) were manually annotated via BLAST where required. Functional patterns were identified across core metabolism, sulphur assimilation, membrane-associated uptake, stress responses, and regulatory proteins.

After analysing the transcripts, it was found that parental JCb10 maintained higher expression of genes associated with primary metabolism compared with the dominant strain E. Key glycolytic and pyruvate pathway enzymes were upregulated in JCb10 during both exponential and stationary phases. For example, genes involved in the commitment of pyruvate utilisation showed differing expression levels. Pyruvate phosphate dikinase (SCLAV_SC2_08115) showed log2FC = +2.67 (padj = 4.8×10^{-75}), and the pyruvate dehydrogenase E2 component (SCLAV_SC2_14435) had a log2FC = +1.84 (padj = 1.3×10^{-27}; Minges et al., 2017; Song and Jordan, 2012). Several NADH-quinone oxidoreductase subunits (SCLAV_SC2_17335/17340/17360/17365/17370), important for the oxidation of NADH to NAD+ (Spero et al., 2015), were more strongly expressed in JCb10 (log2FC = +1.7 to +2.6, padj ≤ 10^{-20}). ATP synthase subunits (SCLAV_SC2_20745–20775) were also more strongly expressed (log2FC ≈ +2.0, padj ≤ 10^{-10}). These results indicate that JCb10 sustains stronger flux through primary metabolism, while E shows downregulation of genes involved with primary metabolism.

Sulphur-associated pathways were also downregulated in the two evolved strains. In stationary phase, JCb10 expressed sulfite reductase, an enzyme that catalyse the reduction of sulfite to sulfide, enabling the assimilation of sulfate into amino acids and other sulfur-containing molecules (Fischer et al., 2012; SCLAV_SC2_23930) at much higher levels (log2FC = +5.01, padj = 1.4×10^{-241} vs. E). Thiosulfate sulfurtransferase (SCLAV_SC2_15495) and adenylyl-sulfate kinase (SCLAV_SC2_23915) showed similar differences (log2FC = +4.2, –4.8, padj ≤ 10^{-226}). These two genes are also associated with

the assimilation and trafficking of sulphate (Lee et al., 2023; Luo et al., 2025) . Together these results point to strongly reduced levels of sulfur assimilation and thiol redox metabolism in dominant strain E relative to the parental strain.

Additionally, membrane transporter expression was different between strains. Parental JCb10 had increased levels of the important antioxidant glutathione and nitrate import gene transcripts. For example, the glutathione importer ATPase GsiA (SCLAV_SC2_19655) was strongly upregulated (log2FC = +6.05, padj < 10^{-23}), demonstrating an increased import of glutathione (Wang et al., 2017). The nitrate permease NrtD (Llarena et al., 2006; SCLAV_SC2_34030) also has increased levels of transcripts (log2FC = +4.2, padj ≤ 10^{-12}). However, strain E upregulated several amino acid transporters. The cystine permease TcyB (SCLAV_SC2_09535) and ABC transporter responsible for the import of some amino acids YxeN (SCLAV_SC2_09540) were more highly expressed in E (log2FC = −4.3 to −6.2, padj ≤ 10^{-60}). This points towards a general change in membrane transporters within the two evolved strains.

Stress-associated genes were generally downregulated in the evolved strains also. An amino-acid efflux pump, known to remove excess or toxic amino acids (Gaurav et al., 2023; SCLAV_SC2_20380) showed log2FC = +5.23 (padj = $1.4x10^{-178}$ in stationary phase). Additionally, a daunorubicin/doxorubicin resistance ABC transporter (SCLAV_SC2_13555) was also more strongly expressed in JCb10 (log2FC = +2.99, padj = $1.3x10^{-78}$). The daunorubicin and doxorubicin associated efflux pump has broad substrate specific for potentially harmful compounds in *Streptomyces* (Li et al., 2014). These patterns indicate reduced levels of stress associated genes and efflux pumps in evolved strains, especially E.

A change in global regulators was also seen. In exponential phase, strain E expressed WhiB, a transcription factor important for the progression of *Streptomyces* sporulation and cell division (Lilic et al., 2023) at higher levels (SCLAV_SC2_19225, log2FC = −3.93, padj = 2.1×10^{-42}). The regulator SigF, also important for spore development (Homerová et al., 2000) was also more highly expressed in E (SCLAV_SC2_03350, log2FC = −3.64, padj = 3.8×10^{-21}) relative to JCb10. However, parental JCb10 had higher levels of housekeeping and stationary regulators such as SigA, the principal sigma factor of RNA polymerase (Gomez et al., 1998; SCLAV_SC2_22915, log2FC = +1.69, padj = 8.7×10^{-5}) and the sigma factor SigJ (SCLAV_SC2_11425, log2FC = +1.34, padj = 9.6×10^{-11}). These shifts suggest strain E has an altered global regulatory system, while JCb10 has a broader transcriptional profile as would be expected.

The CA gene cluster showed the starkest differences in transcription levels. In stationary phase, JCb10 expressed clavaminate synthase 2 (SCLAV_SC2_20420) and proclavaminate amidinohydrolase (SCLAV_SC2_20425) at extremely high levels compared to strain E (log2FC = +7.5 and +6.9, padj ~ 0), along with other genes responsible for the biosynthesis of CA. The same was seen in exponential phase, though at reduced levels (log2FC +1.6, padj = 0.026 of clavaminate synthase 2 for example). In comparisons with the rare strain (JCb6), expression was consistently higher in JCb10 and JCb6 relative to E (log2FC = +6.5–7.0, padj ≤ 10^{-67} of clavaminate synthase 2 for example). Growth-phase transitions highlighted further differences in expression levels. JCb10 displayed strong induction of CA genes in stationary phase (e.g. SCLAV_SC2_20420 log2FC = −5.96, padj = 2.0×10^{-22} i.e. stationary > exponential), whereas strains E and JCb6 showed little to no growth-phase differences.

The expression of genes with previously identified mutations using Breseq (section 5.3) within dominant strain E and rare strain JCb6 was also assessed during stationary phase growth, compared to the parental strain. The four mutated loci were analysed: ATP-binding protein (E), bifunctional diguanylate cyclase/phosphodiesterase (E), zinc metalloprotease phosphatase (E), and serine/threonine kinase (JCb6).

- ATP-binding protein (SCLAV SC2 10490, (TGCGGGG)2→1): an increased expression within the evolved strain despite a presumed loss of function within the gene (log2FC 2.03, padj 1.38 x 10^{-52}). Neighbouring genes also show an increased transcription within the evolved strains (SCLAV_SC2_1047 log2FC 1.84, SCLAV_SC2_10480 log2FC 4.55, SCLAV_SC2_10485 log2FC 5.16, SCLAV_SC2_10500 log2FC 1.21). This data suggests a likely operon induction as compensation for a non-functional transporter.
- Diguanylate cyclase/phosphodiesterase (SCLAV SC2 21505, Δ5 bp): No differential gene expression detected
- Zinc metalloprotease (SCLAV SC2 17525, 58 bp→T): an increased expression within the evolved strain despite a presumed loss of function within the gene (log2FC 1.48, padj 9.43 x 10^{-46}). Neighbouring genes also show increased transcription within the evolved strains (SCLAV_SC2_17525 log2FC 1.48, SCLAV_SC2_17530 log2FC 1.74), indicating co-induction of the htpX–yccF locus. This pattern suggests activation of an envelope stress regulon, likely as compensation for the non-functional HtpX protease caused by the 58 bp deletion.
- Serine/threonine kinase (SCLAV SC2 14025, C→T, synonymous): No differential gene expression detected

In the comparison between the dominant strain (E) and the rare strain (JCb6), a clear difference was observed between growth phases. During exponential growth, 1,773 DEGs were identified, with a balanced distribution of up- and downregulated genes (49.7% and 50.3%, respectively). In contrast, during stationary phase, the total number of DEGs fell to 1,437, but the proportion of upregulated genes increased to 56.7% (Table 5.2). This indicates that while fewer genes were differentially expressed overall in stationary phase, a larger share of these were activated rather than repressed.

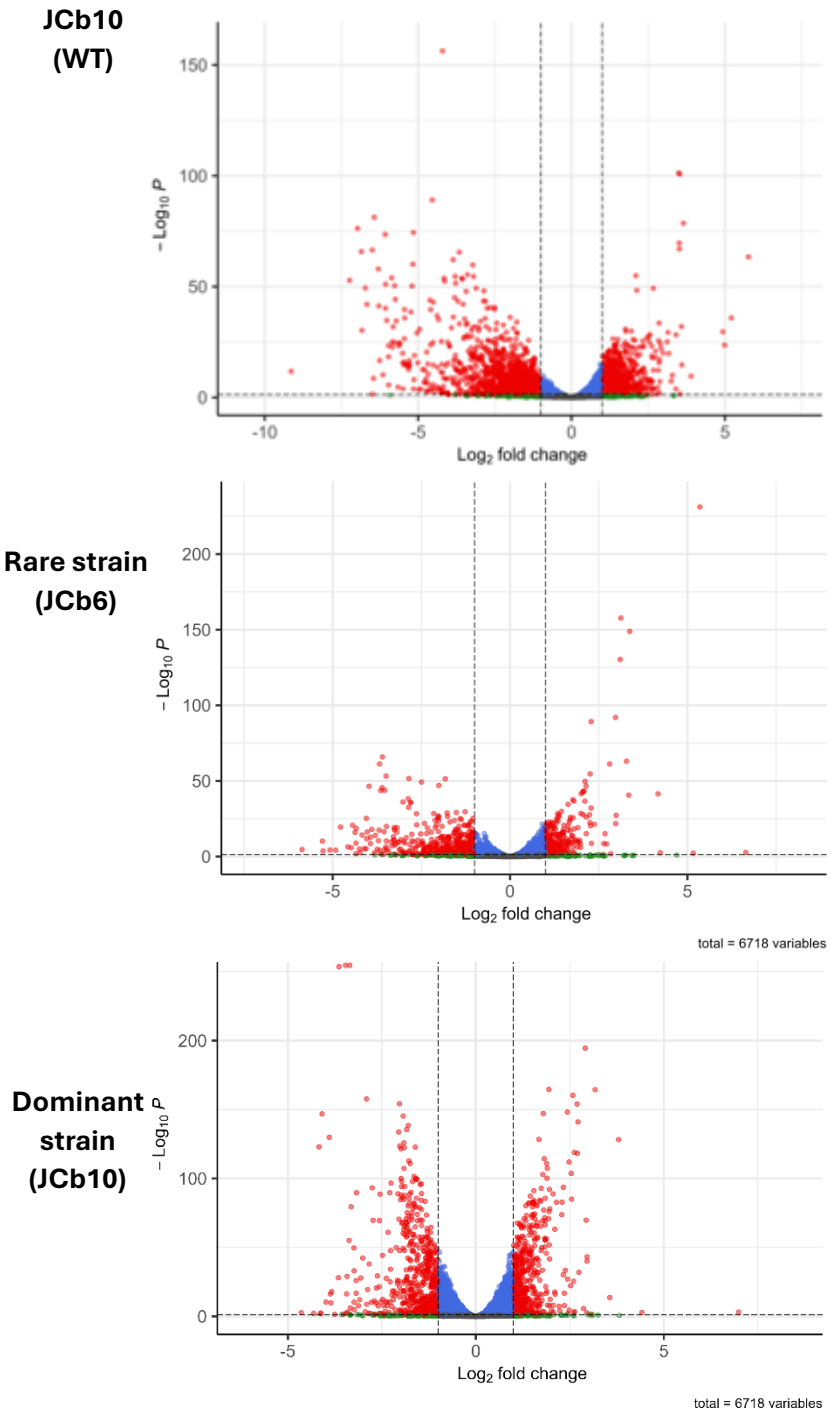


Figure 5.6. Differential transcript expression between stationary and exponential phases in parental, dominant, and rare strains. *S. clavuligerus* SC6 (parental), evolution experiment B colony E (dominant), and evolution experiment B JCb6 (rare) were grown in TSB for 48 hours. Growth phases were determined using a cell growth quantifier. RNA was extracted at both exponential and stationary phases and sequenced. Volcano plots show $\log_2$ fold change versus p value for each transcript. Points are coloured as follows: grey = not significant, green = significant $\log_2$ fold change only, blue = significant p value only, red = significant for both $\log_2$ fold change and p value.

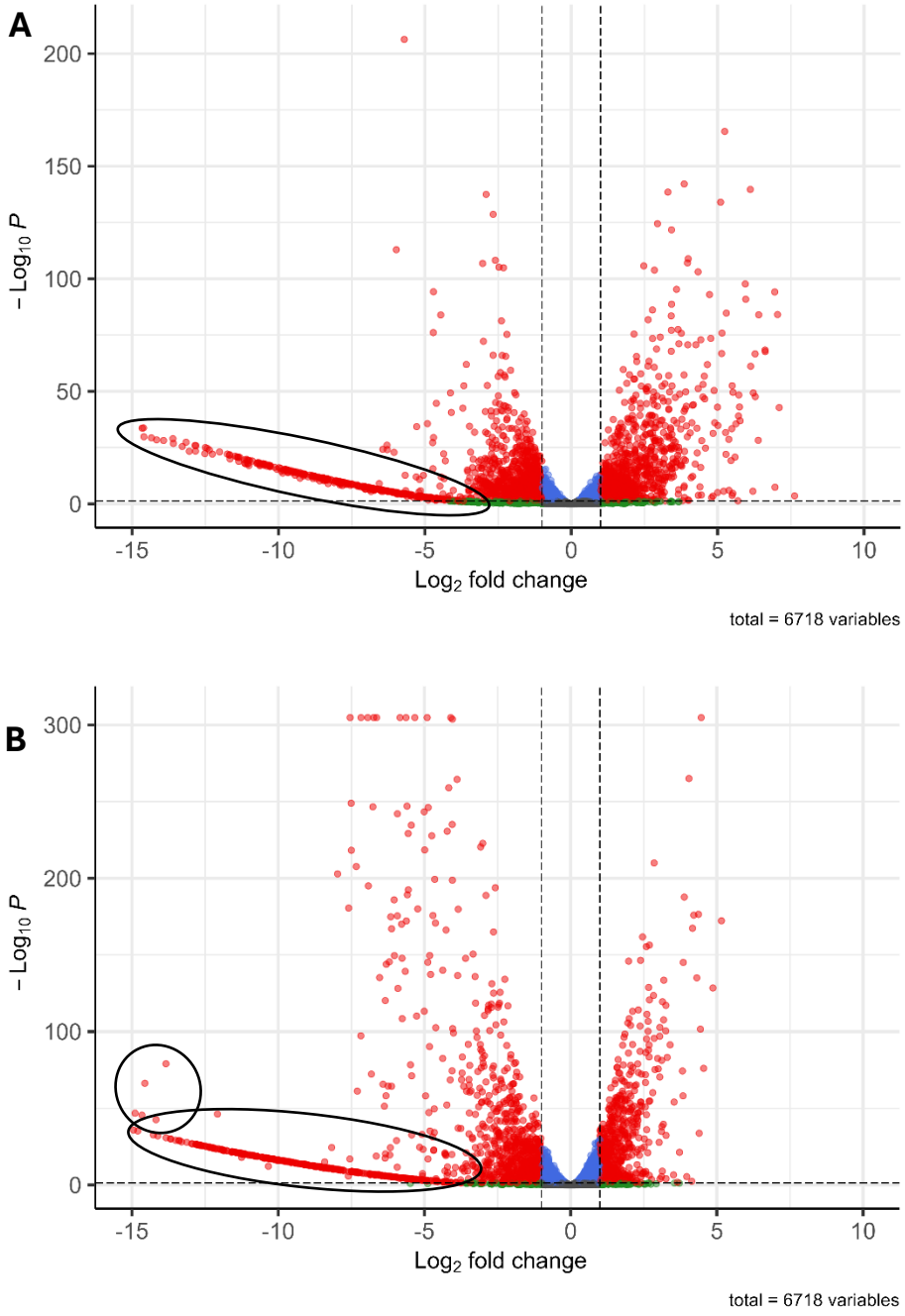


Figure 5.7. Differential transcript expression between parental and dominant strains in stationary and exponential phases. *S. clavuligerus* SC6 (parental) and evolution experiment B colony E (dominant) were grown in TSB for 48 hours. Growth phases were determined using a cell growth quantifier. RNA was extracted at both exponential and stationary phases and sequenced. Volcano plots compare transcript expression between parental and dominant strains during **(A)** exponential and **(B)** stationary phases. Transcripts highlighted within the circle correspond to genes located on pSCL3, which is lost in the dominant strain. Points are coloured as follows: grey = not significant, green = significant $\log_2$ fold change only, blue = significant P value only, red = significant for both $\log_2$ fold change and P value.

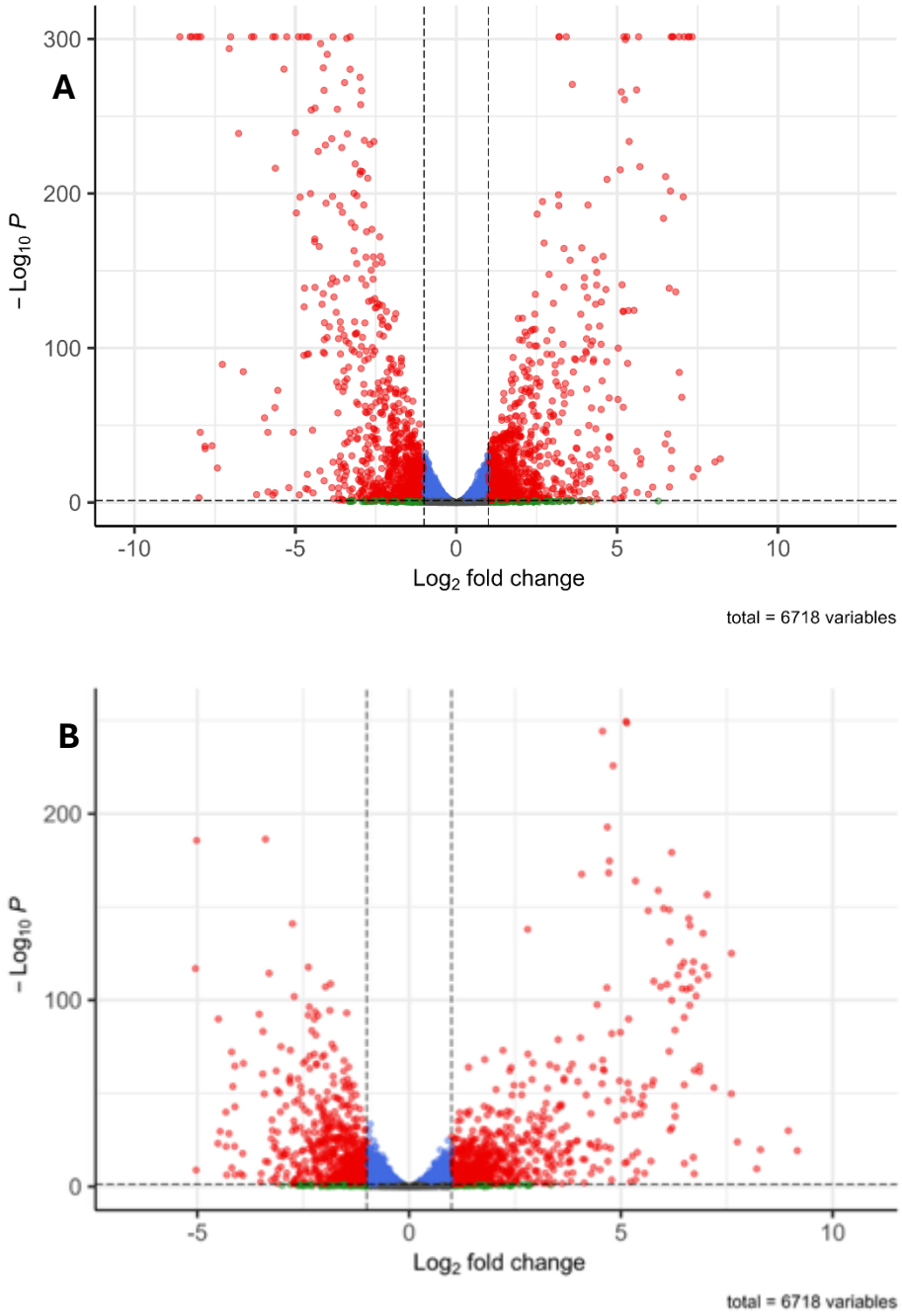


Figure 5.8. Differential transcript expression between dominant and rare strains in stationary and exponential phases. Evolution experiment B colony E (dominant) and evolution experiment B JCb6 (rare) were grown in TSB for 48 hours. Growth phases were determined using a cell growth quantifier. RNA was extracted at both exponential and stationary phases and sequenced. Volcano plots compare transcript expression between dominant and rare strains during **(A)** exponential and **(B)** stationary phases. Points are coloured as follows: grey = not significant, green = significant $\log_2$ fold change only, blue = significant P value only, red = significant for both $\log_2$ fold change and P value.

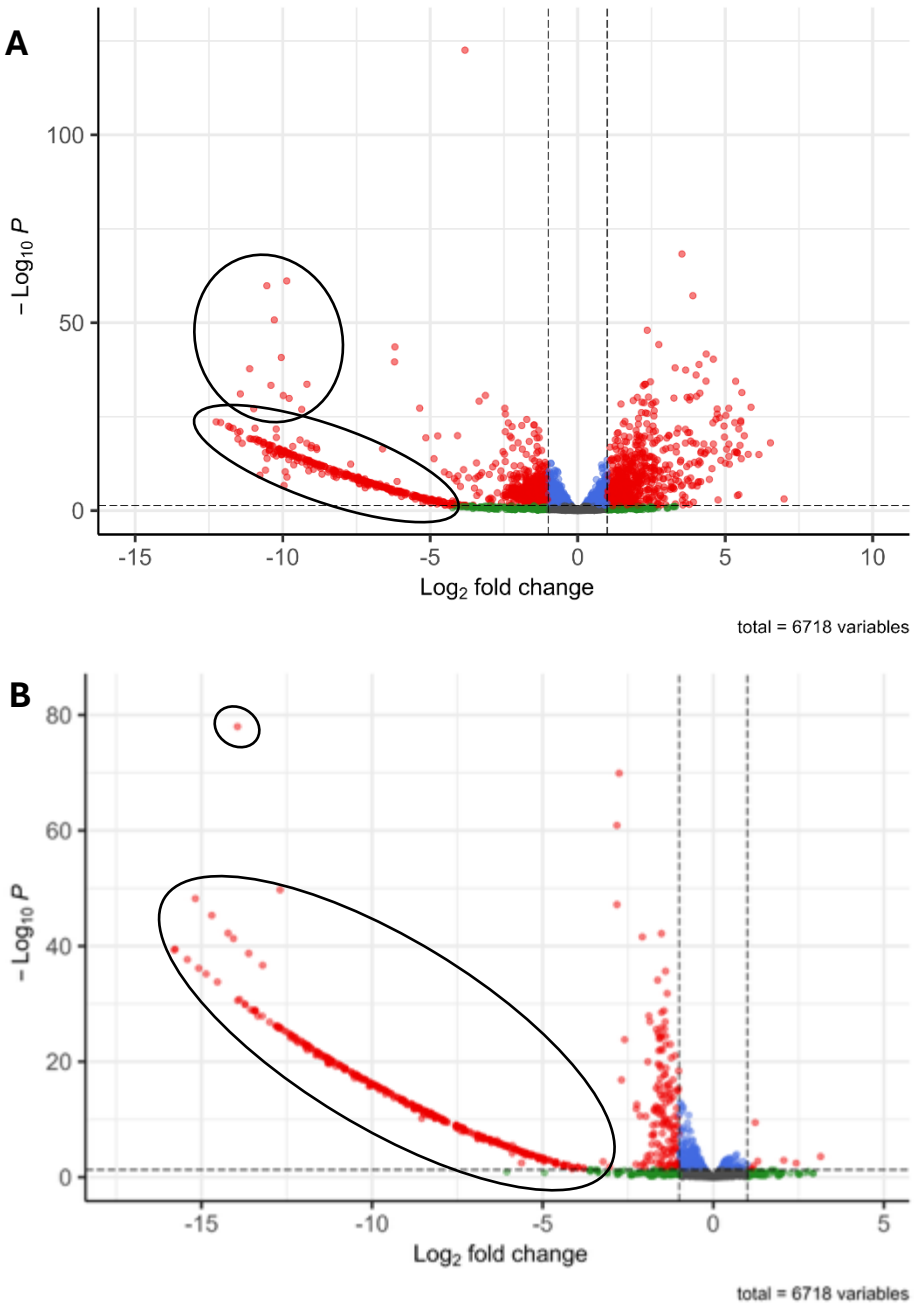


Figure 5.9. Differential transcript expression between parental and rare strains in stationary and exponential phases. *S. clavuligerus* SC6 (parental) and evolution experiment B JCb6 (rare) were grown in TSB for 48 hours. Growth phases were determined using a cell growth quantifier. RNA was extracted at both exponential and stationary phases and sequenced. Volcano plots compare transcript expression between parental and rare strains during **(A)** exponential and **(B)** stationary phases. Transcripts highlighted within the circle correspond to genes located on pSCL3, which is lost in the dominant strain. Points are coloured as follows: grey = not significant, green = significant $\log_2$ fold change only, blue = significant P value only, red = significant for both $\log_2$ fold change and P value.

Table 5.2. Percentage of differentially expressed genes (DEGs) up- and downregulated across all pairwise comparisons and growth phases. DEGs were identified using DESeq2 with an adjusted p-value threshold of ≤0.05. Each comparison shows the proportion of DEGs that were upregulated and downregulated. Comparisons are grouped by strain and growth phase.

Comparison	Growth phase	Total DEGs	% Upregulated	% Downregulated
Parental strain (JCb10)	Stationary vs exponential	2180	40.8	59.2
Dominant strain (E)	Stationary vs exponential	1053	45.4	54.6
Rare (JCb6)	Stationary vs exponential	790	34.9	65.1
Dominant vs. Parental	Exponential	2333	57.0	43.0
	Stationary	1901	69.5	30.5
Dominant vs. Rare	Exponential	1773	49.7	50.3
	Stationary	1437	56.7	43.3
Parental vs. Rare	Exponential	1539	47.6	52.4
	Stationary	502	1.0	99.0

5.6 Phenotypic analysis of evolution experiment B

Given that both genomic and transcriptomic differences were observed between evolved and parental strains, as well as between dominant and rare strains after experimental evolution, phenotypic analyses were carried out to determine whether these differences translated into measurable phenotypic changes, allowing the link between genotype and phenotype.

As transcriptomic data indicated changes in central carbon metabolism, carbon catabolite repression was investigated. This was based on previous findings that loss of carbon catabolite repression in some *Streptomyces* strains enabled the use of the otherwise toxic 2-deoxyglucose as a carbon source (Hodgson, 1982). To test this, *S. clavuligerus* DSM738, SC2, SC6, and all evolved strains were plated on minimal medium agar containing 2-deoxyglucose as the sole carbon source. After 10 days, no growth was observed on any plate, indicating that glucose repression and sensing mechanisms remained intact in the evolved strains due to the fact all tested strains grew on agars containing a standard carbon source (e.g. L3M9 agar).

In addition to investigating the carbon catabolite repression within evolved strains, the CA yield within evolved strains was also investigated. This was because while no mutations were identified within the genomes of evolved strains, transcriptomic data suggests that CA genes are downregulated after adaptation. Assessing the levels of CA yield would therefore aid in linking the genotype to the phenotype of these strains. The five dominant strains (A to E), rare (JCb6) strain and parental strain JCb10 were all grown for 48 hours in 50 mL TSB. After 48 hours the clavulanic acid titre was calculated and normalised to dry weight, the data were plotted (Figure 5.10).

The mean CA titre for the parental strain was 14.01 µg/mg cell dry weight. This compares to a mean titre of 9.60 and 13.08 µg/mg cell dry weight for the five dominant and rare strains respectively, representing a reduction in CA production in all evolved strains. The CA yield of dominant strain A was investigated to explore if the retention of pSCL3, compared to other dominant strains (B to E), affected the production of CA. The CA yield of strain A was 12.61 µg/mg cell dry weight, whereas the average CA yield of strains B to E was 8.85 µg/mg cell dry weight.

Statistical tests were then performed between groups. An ANOVA detected significant differences between groups ($p = 3.67 \times 10^{-5}$). Among dominant strain isolates, CA titre also varied significantly. For example, a Mann Whitney test returned $p = 0.0026$ between colonies C and A. When grouped, dominant colonies showed significantly lower CA titres than both parental and rare strains ($p = 0.0039$ and $p = 0.0213$ respectively), while no difference were observed between parental and rare strains ($p = 0.8141$).

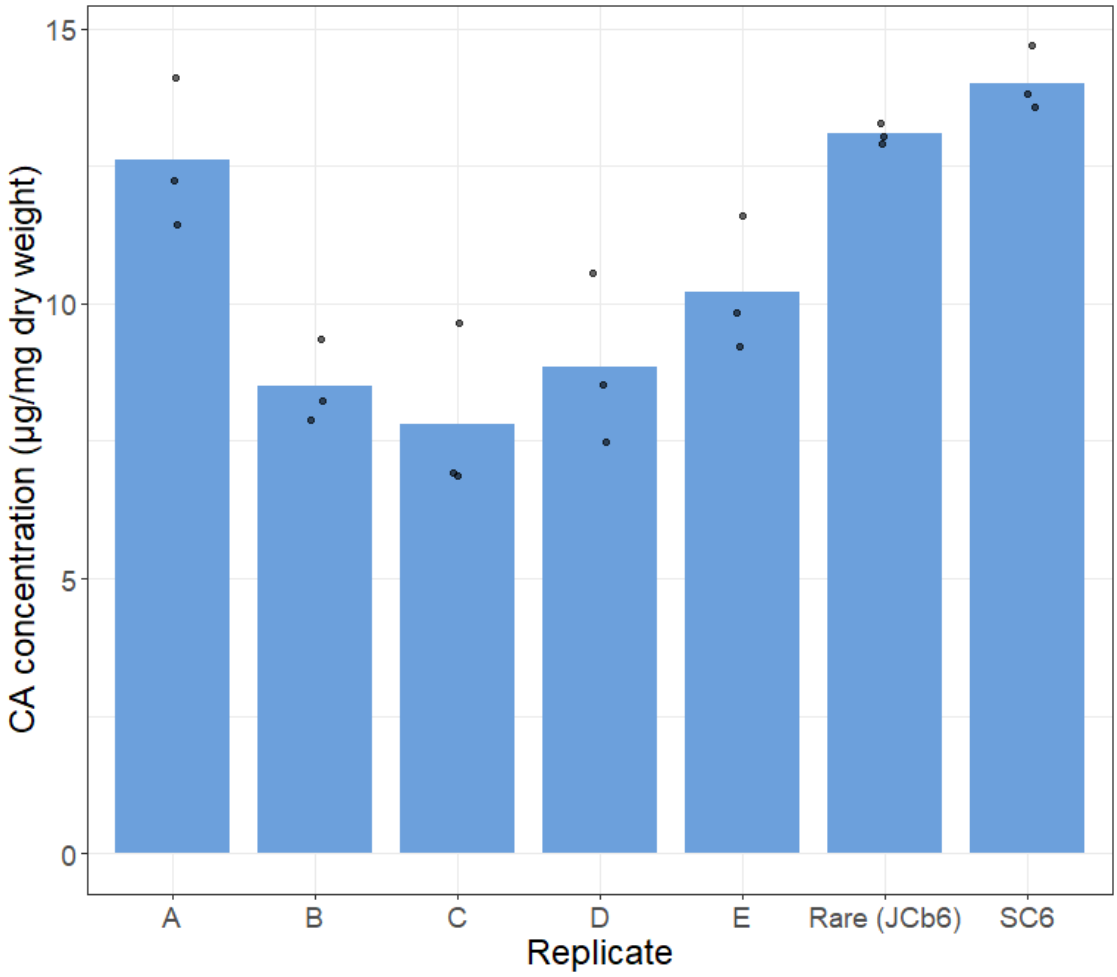


Figure 5.10. Clavulanic acid titre in parental, dominant, and rare strains after evolution. Five colonies of the evolution experiment B JCb10 (dominant strain), one colony of evolution experiment B JCb6 (rare strain), and a parental strain were grown in TSB for 48 hours, and clavulanic acid (CA) titre was measured. ANOVA detected significant differences between groups ($p = 3.67 \times 10^{-5}$). Among dominant strain isolates, CA titre also varied significantly (Mann–Whitney p = 0.0026 between colonies C and A). When grouped, dominant colonies showed significantly lower CA titres than both parental and rare strains (p = 0.0039 and p = 0.0213, respectively), while no difference was observed between parental and rare (p = 0.8141). Average CA titres were 14.00, 13.08, and 9.60 µg/mL for parental, rare, and grouped dominant strains, respectively.

In addition to investigating carbon catabolite repression and CA yield within evolved strains, the central carbon metabolism of evolved strains was further investigated. This is because, like CA production, while no genomic mutations had occurred within central carbon metabolism genes, transcriptomic data suggest alterations in the transcription levels within several key genes. Additionally, an increased level of central carbon metabolism would suggest a diversion of flux away from secondary metabolite production towards primary metabolism and growth. G6PDH represents a branch point in metabolism between glycolysis and the pentose phosphate pathway (Stincone et al., 2015) and in *Streptomyces* there are generally two parologous proteins found in the genome (Schniete et al., 2018). It has been demonstrated that these are differentially expressed (Schniete et al., 2020) and their activity can affect antibiotic production (Butler et al., 2002). To investigate this, the five dominant strains (A to E), rare (JCb6) strain and parental JCb10 were all grown for 48 hours in 50 mL TSB. After 48 hours a Glucose 6 Phosphate Dehydrogenase Activity Assay Kit (Fluorometric, Abcam) was used to assess the activity of glucose 6 phosphate dehydrogenase (G6PDH). The levels of G6PDH activity were normalised to the total protein concentration (using a Bradford assay) within each sample as recommended by the manufacturer.

A high level of variability between all samples was observed (Figure 5.11). For example, a range of 17.84 nmol/min/mL/mg total protein was observed between dominant strains D (35.97 nmol/min/mL/mg total protein) and A (18.13 nmol/min/mL/mg total protein). It is also notable that dominant colony A, previously noted for its retention of pSCL3 and high production of CA, also had the lowest level of G6PDH activity within the dominant strains. While dominant strain A had a G6PDH activity of 18.13 nmol/min/mL/mg total protein, the average activity of dominant strains B to E was 26.94 nmol/min/mL/mg total protein.

The average G6PDH activity of the parental strain, dominant (A to E) and rare (JCb6) strains was 13.24, 25.18 and 16.33 nmol/min/mL/mg total protein respectively, representing an increase in G6PDH activity in evolved strains.

Statistical analysis was performed between groups. An ANOVA indicated significant differences between groups ($p = 5.1 \times 10^{-13}$). Among dominant strain isolates, G6PDH activity also varied significantly. For example, a Mann Whitney test returned a p value < 0.0001 between colonies D and A. When grouped, dominant colonies differed significantly from both rare and parental (p = 0.036 and 0.0003 respectively). However, little statistical difference was observed between rare and parental strains (P = 0.685). These results are consistent with G6PDH activity found within the literature of ~20 nmol/min/mL/mg total protein (Obanye et al., 1996).

Additionally, transcript levels of the *zwf* gene, which encodes glucose-6-phosphate dehydrogenase (G6PDH) and is a key regulator of PPP flux and NADPH generation (Sandoval et al., 2011; Spaans et al., 2015), were examined. The *zwf* sequence from *Streptomyces avermitilis* (WP_010987713.1) was used as a representative query and aligned by BLAST (Altschul et al., 1990) against the *S. clavuligerus* SC2 annotated genome. This identified SCLAV_SC2_05510, annotated as glucose-6-phosphate 1-dehydrogenase 2, as the orthologous *zwf* gene. Transcriptomic analysis in stationary phase showed no significant differential expression of this gene between the parental strain and the dominant isolate (E), nor between the parental and the rare isolate (JCb6).

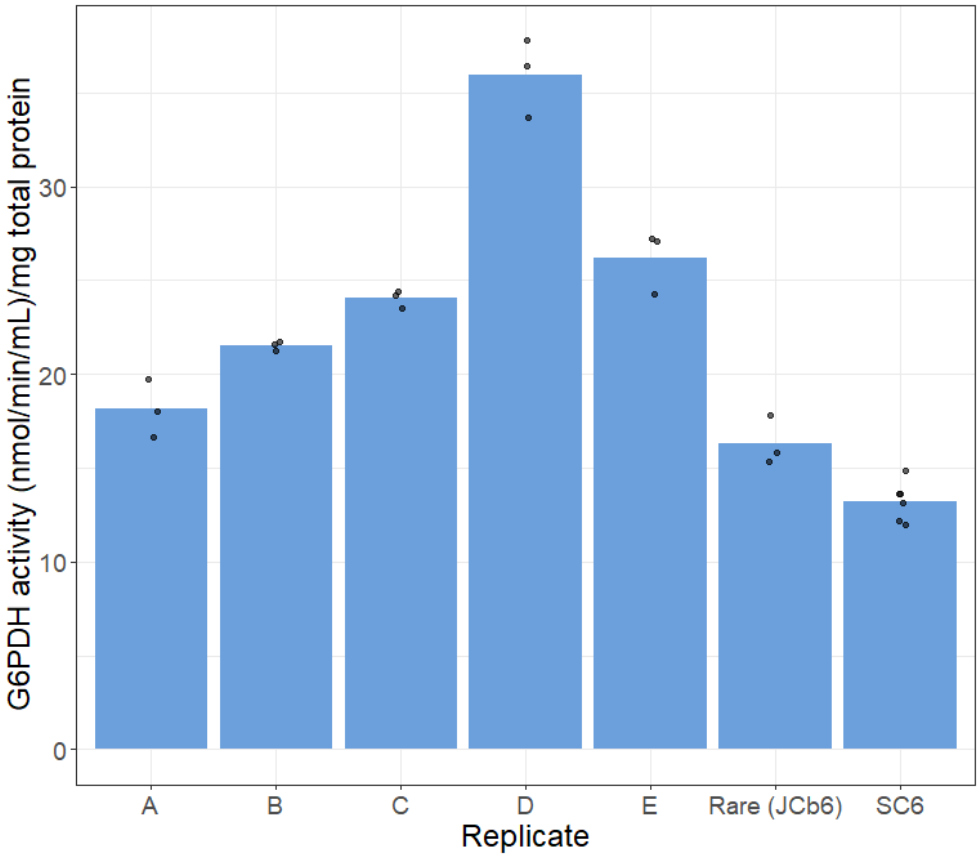


Figure 5.11. Glucose-6-phosphate dehydrogenase activity in parental, dominant, and rare strains after evolution. Five colonies of the evolution experiment B JCb10 (dominant strain), one colony of evolution experiment B JCb6 (rare strain), and the parental strain were grown in TSB for 48 hours, and G6PDH activity was measured and normalised to total protein concentration (Bradford assay). ANOVA indicated significant differences between groups ($p = 5.1 \times 10^{-13}$). Among dominant strain isolates, G6PDH activity also varied significantly (Mann–Whitney $P < 0.0001$ between colonies D and A). When grouped, dominant colonies differed significantly from both rare and parental strains ($p = 0.036$ and 0.0003, respectively), while no difference was observed between rare and parental strains ($p = 0.685$). Average G6PDH activities were 25.18, 16.33, and 13.24 for dominant, rare and parental strains, respectively.

In addition to investigating carbon catabolite repression, CA yield, and G6PDH activity, the specific growth rate (μ) of all strains was measured. The five dominant strains (A–E), the rare (JCb6) strain, and the parental JCb10 were grown in 50 mL TSB for 48 hours, with backscatter measured continuously using the aquilabiolabs Cell Growth Quantifier (Figure 5.12). This allowed for calculation of specific growth rate (μ) based on the exponential phase of each growth curve (Table 5.2).

The specific growth rates of all evolved strains were significantly lower than that of the parental. The average μ for parental strain, dominant, and rare strains were 0.0079 h^{-1}, 0.0031 h^{-1}, and 0.0046 h^{-1} respectively. ANOVA indicated significant differences between groups ($p = 5.35 \times 10^{-6}$). Within dominant strains, inter-isolate variation was low, the largest pairwise Mann–Whitney p value was 0.141 (colonies B vs. D). When grouped, dominant strains differed significantly from both parental ($p < 0.001$) and rare strains ($p = 0.0007$).

Given that other phenotyping results suggest strain A is an outlier, its μ was compared to the average of strains B–E. While A did not have the lowest μ overall, its value (0.0026 h^{-1}) was lower than the mean of the other dominant strains (0.0032 h^{-1}).

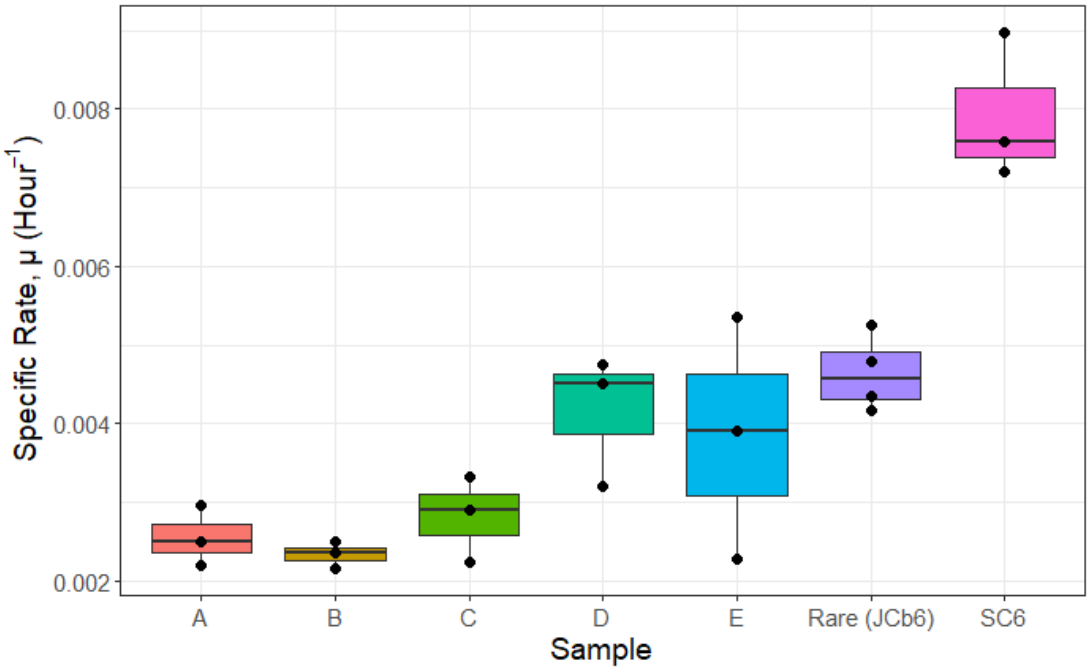


Figure 5.12. Specific growth rates of parental, dominant, and rare strains after evolution. Five colonies of the evolution experiment B JCb10 (dominant strain), one colony of evolution experiment B JCb6 (rare strain), and a parental strain were grown in TSB for 48 hours. Backscatter was measured using a cell growth quantifier to calculate specific growth rates. ANOVA indicated significant differences between groups ($p = 5.35 \times 10^{-6}$). Average specific growth rates were 0.0031, 0.0046, and 0.0079 h^{-1} for dominant, rare, and parental strains, respectively. Variation among dominant strain isolates was minor (largest Mann–Whitney $p = 0.141$ between colonies B and D). When grouped, dominant strains differed significantly from both rare and parental strains ($p < 0.001$ and $p = 0.0007$, respectively).

Table 5.2. Specific growth rates of parental, dominant, and rare strains after evolution. Specific growth rates (μ) of individual and grouped evolved strains are summarised. The loss of pSCL3 in strain A did not improve μ compared to other dominant strains (B–E).

Strain(s)	μ (h^{-1})
B - E	0.0032
A	0.0025
All dominant (A - E)	0.0031
Rare (JCb6)	0.0046
SC6	0.0079

In addition to investigating carbon catabolite repression, CA yield, G6PDH activity, and specific growth rate within evolved strains, pellet morphology in liquid culture was examined to assess whether adaptation had altered pellet size. Although qualitative observations during the evolution experiment in CM5 medium indicated an absence of pellet formation at all timepoints, this was investigated further. Pellet formation is a known trait of *Streptomyces* strains, often associated with reduced antibiotic production (Nieminen et al., 2013b; van Dissel et al., 2015).

The five dominant strains (A–E) and parental JCb10 were grown in 50 mL TSB for 48 hours before imaging by phase-contrast microscopy. One representative image (containing multiple cell aggregates) was selected for each strain (Figure 5.13). Evolved strains showed visibly larger pellets compared to the parental strain, which warranted quantitative analysis.

Hyphae could be seen extending from the edges of all pellets, consistent with *Streptomyces* morphology in liquid culture (Paul and Thomas, 1998). No obvious outliers were noted among samples. Some pellets were larger or showed unassociated hyphae, but no strain appeared morphologically distinct from others. Quantitative assessment of pellet diameter was therefore performed.

ImageJ (Schneider et al., 2012) was used to measure pellet diameters (Figure 5.14). Parental JCb10 showed the smallest average pellet diameter (128 µm), while dominant strains exhibited an average 8.6-fold increase. ANOVA confirmed significant differences between groups ($P = 1.09 \times 10^{-48}$), and variation among dominant strains was also significant (e.g. Mann–Whitney P = 0.0002 between colonies A and D). Average pellet diameters for colonies A–E were 1005, 980, 1154, 1186, and 1127 µm, respectively.

Dominant strain A, consistent with previous phenotypic assays, had a smaller average pellet size (1005.15 µm) than the average of strains B–E (1126.59 µm).

It is important to note that all pellet morphology and phenotypic measurements were performed after growth in TSB, not CM5, the medium used during the evolution experiment. This was due to the inability to isolate and regrow individual pellets once removed from culture and placed onto slides. Nonetheless, the consistent differences observed between groups, and the identification of outliers such as strain A, suggest that evolved phenotypic traits persist across growth conditions.

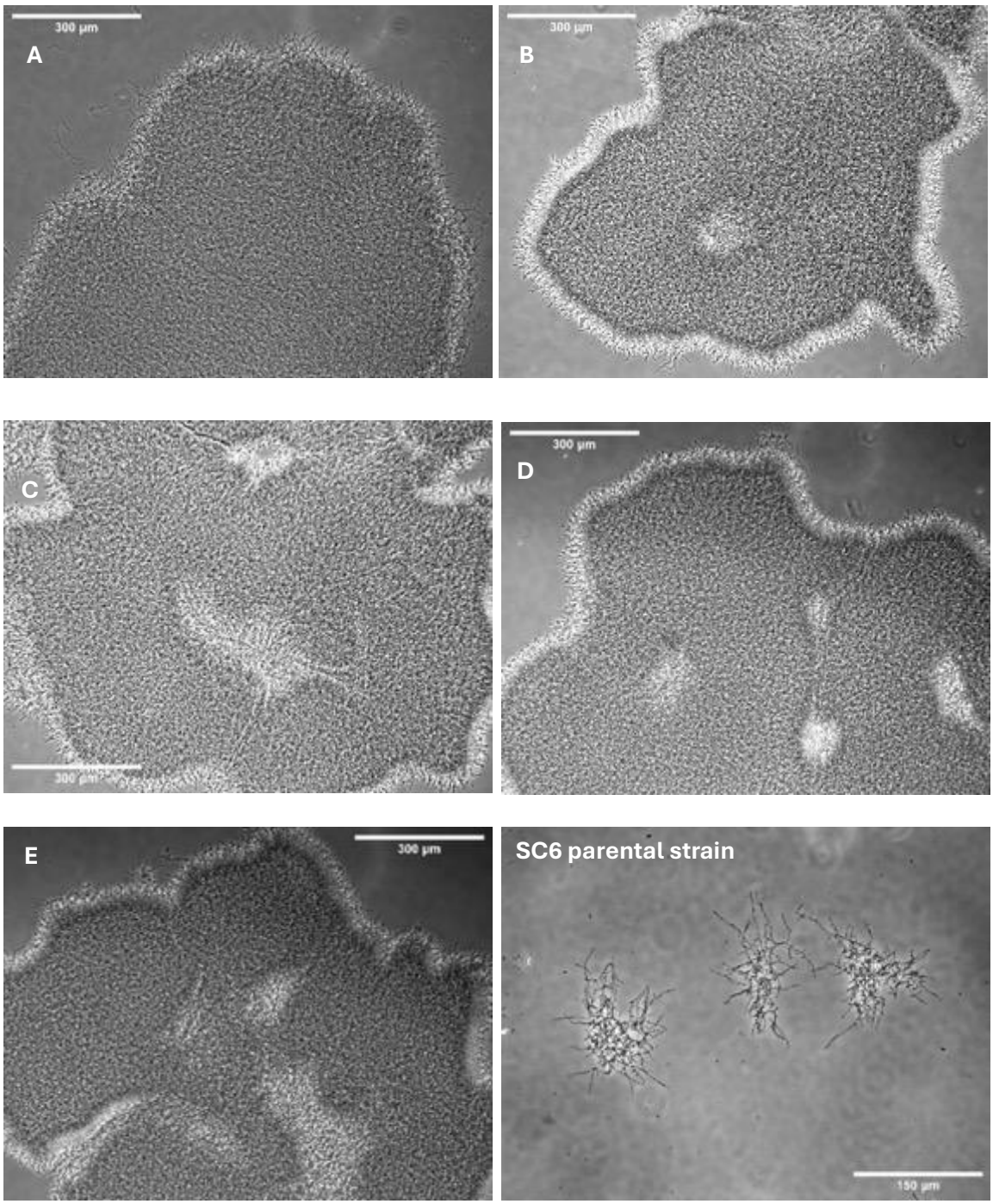


Figure 5.13. Pellet morphology of parental and dominant strains after experimental evolution. Five colonies of the evolution experiment B JCb10 (dominant strain) and a parental strain were grown in TSB for 48 hours. Pellets were imaged by phase contrast microscopy, and one representative image is shown for each strain.

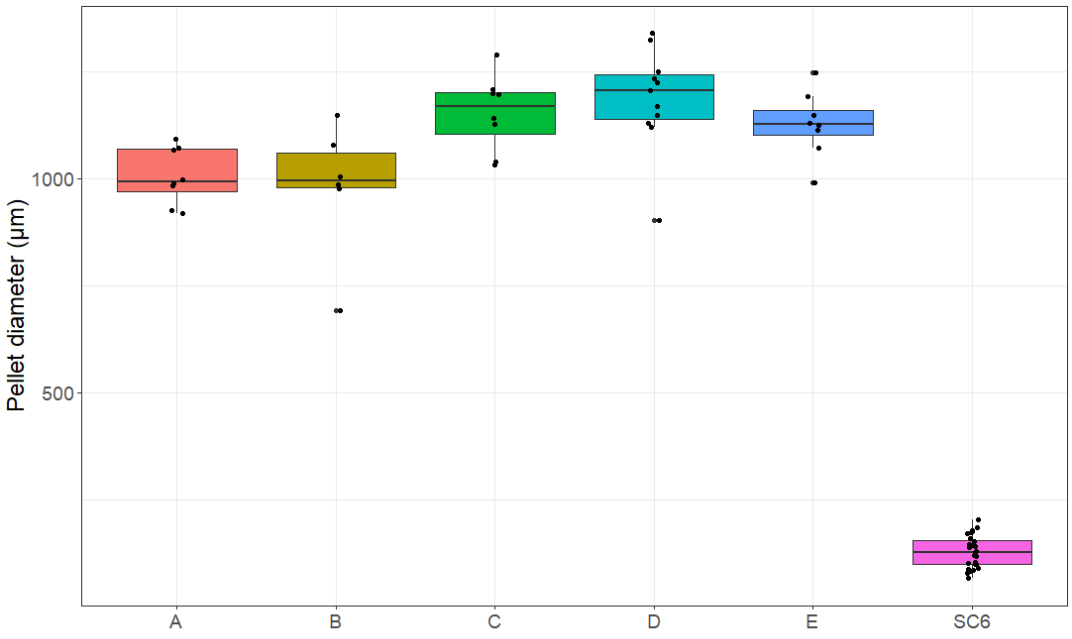


Figure 5.14. Pellet size quantification of parental and dominant strains after evolution. Pellets from parental *S. clavuligerus* and five evolution experiment B JCb10 (dominant strain) colonies were imaged by phase contrast microscopy, and diameters were measured using ImageJ. The parental strain displayed the smallest average diameter (128 µm), while the dominant strains showed an average 8.6-fold increase. ANOVA confirmed significant differences between groups ($p = 1.09 \times 10^{-48}$). Significant variation was also observed among dominant strain isolates (Mann–Whitney P = 0.0002 between colonies A and D). Average pellet diameters for colonies A–E were 1005, 980, 1154, 1186, and 1127 µm, respectively.

5.7 Competition of evolved strains

After observing phenotypic differences between evolved and parental strains, including CA yield, G6PDH activity, μ, and pellet morphology, a key remaining question was whether the evolved strains were in fact adapted to CM5 medium. This question was especially relevant given that parental JCb10 remained dominant in evolution experiment B but only had modest changes in phenotype.

To directly assess fitness, a competition assay using Bar-seq was performed between the dominant experimental evolution isolate E (selected due to its representative genomic mutation profile) and the parental strain *S. clavuligerus* SC6 pSET152 JCb20 (chosen arbitrarily). Both strains were grown separately in 50 mL TSB for 48 hours. They were then mixed in ratios of 25:75, 50:50, and 75:25 according to their OD_{600}. These mixtures were used to inoculate fresh 50 mL cultures of both TSB and CM5, each in triplicate, using the same inoculum across all 18 flasks. Cultures were grown for an additional seven days. Genomic DNA was extracted at timepoints 0 and 168 hours, and the relative abundance of each strain was determined (Figure 5.15).

Sequencing of the initial mixtures revealed that the parental strain (JCb20) represented 41.78%, 59.40%, and 77.18% in the 25:75, 50:50, and 75:25 input ratios, respectively. While these did not precisely match the intended ratios, they were deemed sufficient to assess competitive dynamics over time.

The results clearly showed that the parental strain outcompeted the evolved strain in TSB across all replicates, with parental strain occupying >95.7% of the final population after 168 hours. In contrast, the evolved strain outcompeted the parental train in CM5 medium, with the parental representing <17.7% of the final population across replicates. These findings confirm that adaptation to CM5 medium conferred a selective fitness

advantage specifically under CM5 growth conditions. A parallel assay with the rare isolate JCb6 showed the same qualitative pattern, with the parental strain favoured in TSB and the evolved strain favoured in CM5, although the advantage of JCb6 was weaker than that of JCb10.

The use of three different mixing ratios allowed determination of the presence of frequency-dependent selection. Previous studies have shown that fitness advantages can depend on starting abundance, where certain strains perform better when initially rare (Ribeck and Lenski, 2015). However, in this case, the relative fitness outcomes were consistent across all ratios, indicating that the competition was not frequency dependent.

Having established that the dominant isolate consistently displaced the parental strain in CM5, a further question was whether this advantage was unique to dominant lineages. To address this, an equivalent assay was performed using the rare experimental evolution strain (JCb6) against the parental strain (JCb20). Strains were cultured and mixed in the same ratios as the previous competition assay. As before, these mixtures were used to inoculate fresh 50 mL cultures of both TSB and CM5, each in triplicate. Cultures were incubated for 168 hours, after which genomic DNA was extracted at 0 and 168 hours and the relative abundance of each strain was quantified (Figure 5.16).

Sequencing of the initial inoculum showed that the parental represented 40.27%, 51.05%, and 70.86% of the populations in the 25:75, 50:50, and 75:25 ratios, respectively. While not precisely matching the intended proportions, these values were sufficient to evaluate competitive outcomes.

In TSB, the parental strain increased in relative abundance across all starting ratios, reaching final frequencies of approximately 62%, 77%, and 85% after 120 hours for the

25:75, 50:50, and 75:25 input ratios, respectively. In CM5, the inverse trend was observed, with the parental declining to ~21%, 48%, and 65% at the same ratios. These outcomes parallel those seen with the dominant isolate, with the parental strain consistently outcompeting the evolved strain in TSB and the evolved strain gaining an advantage in CM5. As with the previous assay, the results were consistent across starting frequencies, indicating that competition between JCb6 and JCb20 was not frequency dependent.

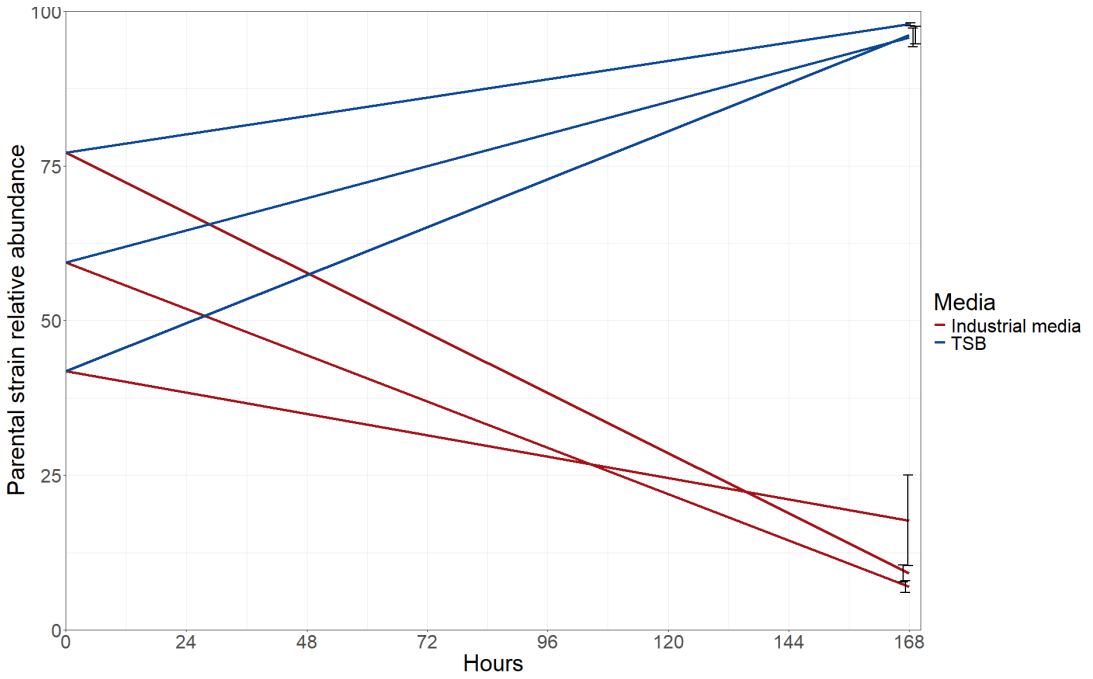


Figure 5.15. Competitive fitness of evolved and parental strains in TSB and CM5. A representative dominant evolved strain (*S. clavuligerus* SC6 pSET152 JCb10 E) was mixed with parental (*S. clavuligerus* SC6 pSET152 JCb20) at OD_{600} adjusted ratios of 25:75, 50:50, and 75:25 after 48 hours of growth in TSB. These mixtures were used to inoculate TSB and CM5 in triplicate and grown for a further seven days. Relative barcode abundances were determined at 168 hours. Error bars represent standard deviation between triplicates. In TSB, the parental strain dominated (>95.7% at 168 hours), while in CM5 the evolved strain dominated (<17.7% parental). Competitive outcome was independent of initial frequency.

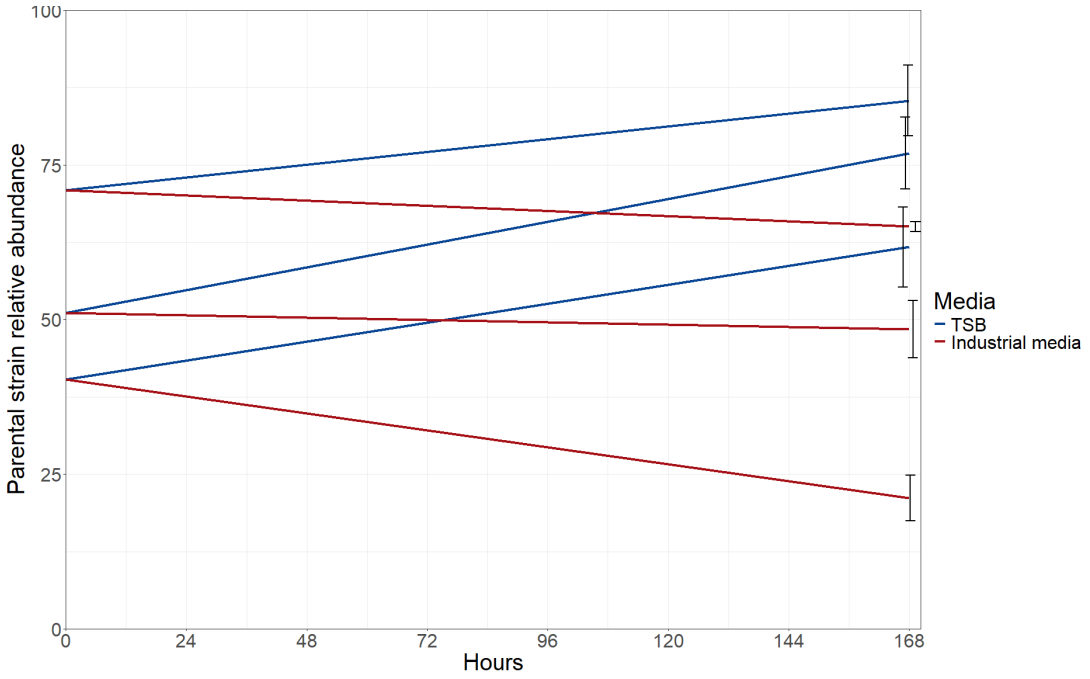


Figure 5.16. Competition between the rare evolved isolate JCb6 and the parental strain JCb20. Strains were mixed at OD_{600} adjusted ratios of 25:75, 50:50, and 75:25 after 48 hours of growth in TSB. These mixtures were used to inoculate TSB and CM5 in triplicate and grown for a further seven days. Relative barcode abundances were determined at 168 hours. Error bars represent standard deviation between triplicates. In TSB, the parental strain increased to ~62–85% across input ratios, whereas in CM5 it declined to ~21–65%. Competitive outcome was independent of initial frequency.

5.8 Summary

This chapter set out to explore the evolutionary potential of *Streptomyces clavuligerus* strains under industrially relevant fermentation media. Using the barcoded strains previously shown to be stable in short-term experiments, populations were serially passaged in CM5 media over 70 days (152 generations). This design aimed to identify whether sustained growth and nutrient limitation would drive reproducible evolutionary changes, favour dominance of strains, and ultimately impact industrially relevant phenotypes such as CA production.

The results revealed substantial divergence between the three biological replicates. Despite being inoculated from the same starting mixture, the three experiments followed markedly different trajectories. In fermentation A, strain JCb2 rose in abundance before being outcompeted by JCb13. In contrast, fermentation B saw a clear and sustained selective sweep, with JCb10 reaching over 90% relative abundance by day 70. Fermentation C maintained a more diverse population throughout, with several strains persisting at moderate abundance. These results indicate that evolutionary outcomes under these conditions are random, with limited parallelism even across identically treated populations.

The link between population dynamics and phenotype was also investigated. CA production declined across all three evolution experiments over time. This decline was most severe in fermentation B, where strain JCb10 became dominant. This suggests that the mutations which conferred competitive advantage to JCb10 may have come at a cost to CA biosynthesis. While selection under these conditions appears to favour faster growing or more robust variants, these strains are not necessarily the most productive from an industrial perspective.

Whole genome sequencing of isolates from fermentation B identified several parallel mutations in dominant strains. These included mutations in genes associated with regulation such as diguanylate cyclase and zinc metalloprotease. These adaptations may have contributed to the observed dominance of JCb10. Additionally, the loss of plasmid pSCL3 was seen within both dominant and rare strains within evolution experiment B. This suggests that pSCL3 deletion alone is not sufficient to drive adaptation. Importantly, no mutations were found directly in genes involved in primary metabolism.

However, transcriptomic analysis revealed a more nuanced fine tuning of *S. clavuligerus* strains compared to genotypic mutations alone. Across dominant isolates, extensive changes in gene expression were observed, including many genes linked to primary metabolism, stress response, and development. This discrepancy between a small genomic mutational profile compared to a widespread change in transcriptome, suggests that regulatory tinkering, compared to broad genetic changes is the primary driver of adaptation in this context. This is consistent with a broader trend observed in microbial evolution, where adaptation under selective pressure can occur via changes in global regulatory states or epigenetic mechanisms, rather than direct mutations in core metabolic genes.

Overall, the use of evolution experiments shows that under conditions mimicking industrial fermentation, *S. clavuligerus* populations can and do evolve but not predictably. Strain dominance is not guaranteed, nor is it always beneficial for production yield, as is expected. The data highlight the importance of monitoring not just genetic changes, but also regulatory and transcriptional responses over time. Adaptive evolution under industrial conditions may negatively impact secondary metabolite

production, and even small-scale serial passaging can result in the fixation of strains with lower yields. These findings also underscore the potential importance of epigenetic regulation and transcriptional plasticity as key contributors to fitness and adaptability in *S. clavuligerus*.

Direct competition assays further confirmed that adaptation to CM5 medium conferred a selective advantage. Both a dominant evolved isolate (JCb10 E) and a rare isolate (JCb6) displaced the parental strain in CM5, whereas the parental strain remained superior in TSB. Competitive outcomes were consistent across input ratios, indicating that fitness differences were not frequency dependent. While both isolates showed adaptation to CM5, the advantage of JCb6 was consistently smaller than that of JCb10.

Chapter 6 Genotypic and phenotypic characterisation of post-fermentation strains

6.1 Introduction

Streptomyces exhibit complex morphological differentiation as part of their life cycle, with traits such as colony size, pigmentation, and sporulation regulated by developmental pathways involving the key global regulatory *bld* and *whi* genes that often have pleiotropic effects (Chater, 2001; McCormick and Flärdh, 2012). These traits are tightly linked to secondary metabolism and environmental sensing and are influenced by both genetic and environmental factors (Holmes et al., 2013; Lawlor et al., 1987). In industrial contexts, strain morphology can affect fermentation performance, for example through pellet formation or altered oxygen transfer (Nieminen et al., 2013b; van Dissel et al., 2015). While this has been explored within the previous chapter over a longer timescale than is industrially relevant, some genotypic variation has been linked to its phenotypic outcomes in that work.

Although barcode sequencing allows temporal tracking of strain abundance, it does not itself capture phenotypic traits that may arise as a consequence of evolutionary selection in the fermentation conditions. To determine this, further analysis is required to assess whether any of the strains recovered from different fermentation conditions, particularly those that emerged as dominant or rare lineages, exhibit distinct characteristics. This is especially relevant given that phenotypic based screening remains a widely used tool in strain selection and optimisation in industry. Additionally further focus must be paid to industrially relevant timescales and volumes for a better understanding of the industrial production of CA.

In this chapter, strains isolated from various fermentation experiments were replica plated onto a range of nutrient agars to determine the phenotype of the strains, such as ability to grow and sporulate on a range of media. This included isolates from the 1 L and 7.5 L multistage fermentations, as well as from the evolution experiment presented in

Chapter 5. Phenotypic traits were compared across strains and growth conditions, and assessed relative to the isogenic parental strain, *S. clavuligerus* SC6 stock to determine the extent to which fermentation conditions influence the genotype and the phenotype of the strains as they emerge under the conditions they were grown. This analysis also aimed to evaluate whether morphology could serve as a reliable proxy for strain identity or performance, and to better understand the baseline variability within fermented *Streptomyces* populations.

All experimental work, quantitative analyses (including image-based measurements), sequencing analysis, and figure preparation in this chapter were performed by James Croxford. For stirred-tank bioreactor fermentations (>1 L) conducted at GSK, reactor inoculation, operation, and sample collection were performed by Wei Li Thong and Alistair Middlemiss (GSK), while James Croxford prepared the inoculum and performed all downstream sample processing, sequencing, and data analysis.

6.2 Analysis of colony morphology after fermentation

Following completion of fermentations at various timescales and volumes, phenotypic analyses were carried out to determine whether fermentation conditions influenced the resulting strain morphology. While the previous chapter focused on linking genotype to phenotype, further examination was needed to assess how fermentation scale might affect both in a wider range of conditions. The first step was to evaluate colony morphology.

Firstly, to further explore the baseline variability within clonal populations, an isogenic stock of *S. clavuligerus* SC6 was aliquoted into 48 wells of a 96-well plate and replica plated onto the agars that had previously supported growth. All plates were incubated

for 10 days and imaged, with an L3M9 plate shown as an example (Figure 6.1). Even among this isogenic population, variation in colony size and sporulation was observed. Some colonies showed sporulation only in the centre of the colony, while others did not. Similarly, colony outlines varied from smooth and circular to irregular and fanned. This background level of variability demonstrates that minor morphological variation between colonies is normal, even in clonal populations, making it difficult to confidently identify true outliers within the evolved or fermentation derived groups based on morphology alone.

After, 94 colonies were randomly selected from three fermentations: a 1 L multistage fermentation (Figure 4.8), a 7.5 L multistage fermentation (Figure 4.14), and the evolution experiment B from the previous chapter (Figure 5.1B). Each colony was used to inoculate 50 µL of 20% glycerol in a 96-well plate using the following layout (left to right, top to bottom):

- Wells 1 and 96: *S. clavuligerus* SC6 (parental control)
- Wells 2–16: Evolution experiment B (final timepoint)
- Wells 17–55 and 86–95: 7.5 L multistage fermentation (final timepoint, Figure 4.15 A3)
- Wells 56–84: 1 L multistage fermentation (final timepoint, Figure 4.2 A3)

After inoculation, a microplate replicator (Boekel Scientific) was used to transfer cultures onto L3M9 agar. Plates were incubated for 10 days before imaging (Figure 6.2).

Clear differences in colony morphology emerged between strains derived from different fermentation conditions. Colonies from evolution experiment B formed visibly larger colonies. Those from the 1 L multistage fermentation showed improved sporulation

compared to colonies from the 7.5 L fermentation. Overall, colony size, pigmentation, and sporulation patterns varied between groups. However, within each group, most colonies appeared morphologically consistent, no obvious outliers were observed at this stage. The presence or absence of such outliers may later help explain whether a particular strain became dominant or rare within a fermentation.

It is worth noting that colonies 2 and 3 from the evolution experiment B group failed to grow, along with others observed later. This was attributed to poor inoculation or storage in the 96-well plate.

Interestingly, even the two SC6 parental strain control colonies (wells 1 and 96) differed in both size and sporulation, raising further questions about within-strain phenotypic variability. These differences, as well as potential group outliers, are addressed in more detail in the following sections.

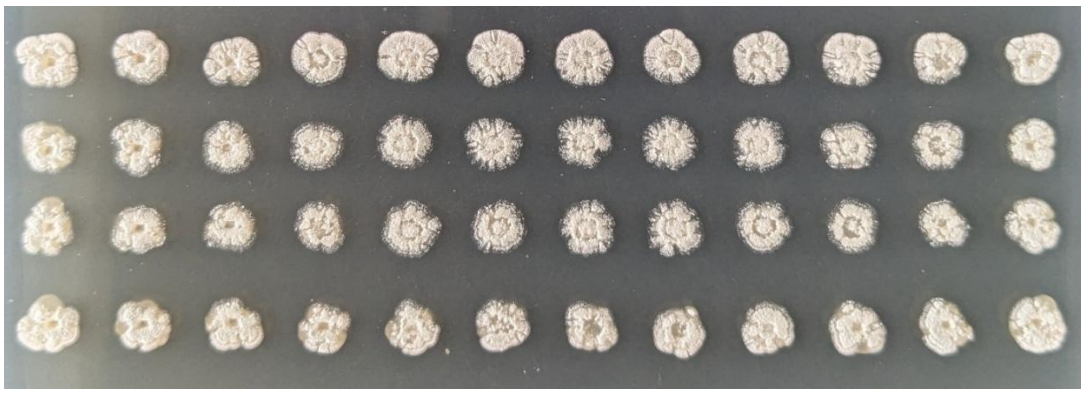


Figure 6.1. Colony morphology varies within an isogenic stock of *S. clavuligerus* SC6. An isogenic stock of *S. clavuligerus* SC6 was aliquoted into 48 colonies and plated onto L3M9 agar using a replica plater. Plates were incubated for 10 days. Variation in colony size and sporulation is observed within the isogenic population.

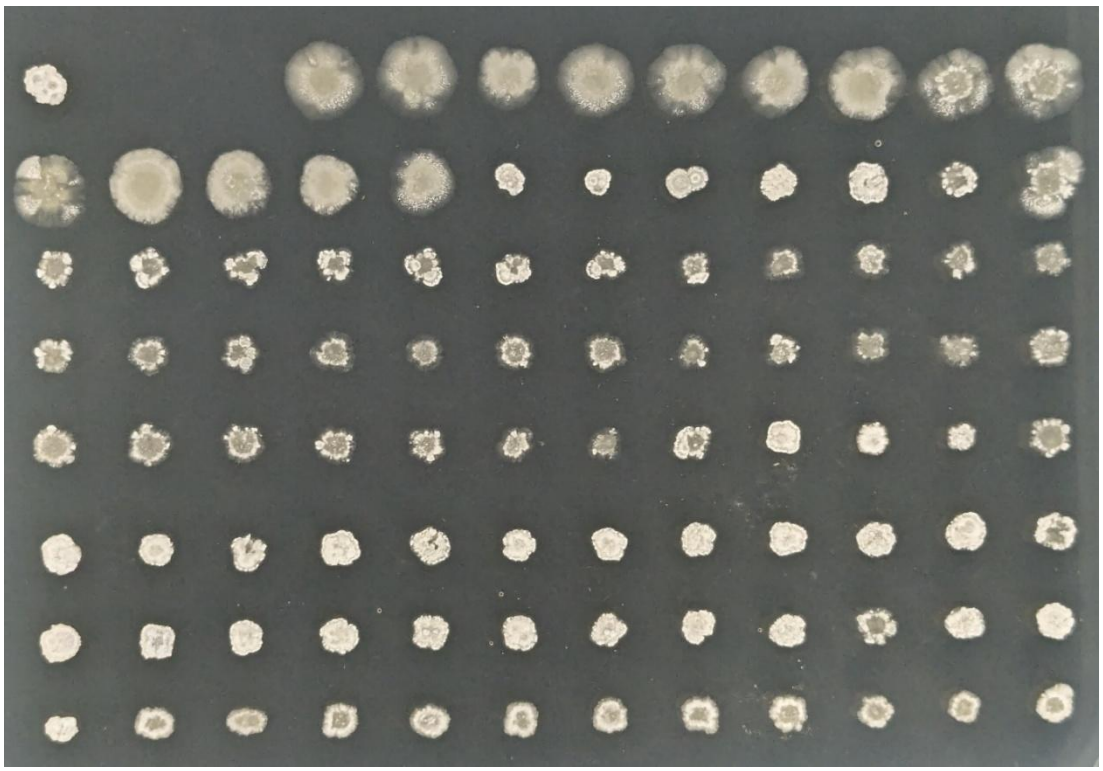


Figure 6.2. Colony morphology varies between fermentation conditions. Colonies from fermentations described in chapters 4 and 5 were replica-plated onto L3M9 agar and grown for 10 days. Colonies are arranged left to right, top to bottom: (1) SC6; (2–17) evolution experiment B; (18–56) 7.5 L multistage fermentation; (57–85) 1 L multistage fermentation; (86–95) 7.5 L multistage fermentation; (96) SC6. Distinct differences in colony size and morphology are visible between fermentation types.

To investigate whether outliers existed within groups based on colony morphology, and to test whether the large differences observed between groups could be reproduced, the same 96 colonies were replica plated onto a wider range of agars. The aim was to determine whether differences in nutrient source would accentuate morphological variability within and between fermentation groups. As *Streptomyces* are known to grow on a wide variety of nutrient sources (Hodgson, 2000) and 19 agars containing diverse carbon, nitrogen and other substrates were selected.

Colonies were grown on all 19 agars for 10 days. Growth was observed on 2xYT, instant mash agar, GYM (Figure 6.3), ISP4, L3M9, and LB (Figure 6.4), as well as milk agar, R2A, starch agar (Figure 6.5), and TSB agar (Figure 6.6). No growth occurred on cornmeal agar, ISP3, ISP7, minimal agar with fructose or glucose, MYM, Sabouraud, MS, or YEME.

Across the media that supported growth, a range of morphologies were observed. Sporulation was most apparent on L3M9, consistent with this being the medium that the strains are propagated on as standard. The largest colonies consistently formed on GYM. In contrast, LB and starch agar produced very small colonies with little visible variation, likely a result of poor growth that limited phenotypic diversity. While a range of morphologies were observed across the usable agars, no single colony stood out as a consistent outlier within its group across all conditions.

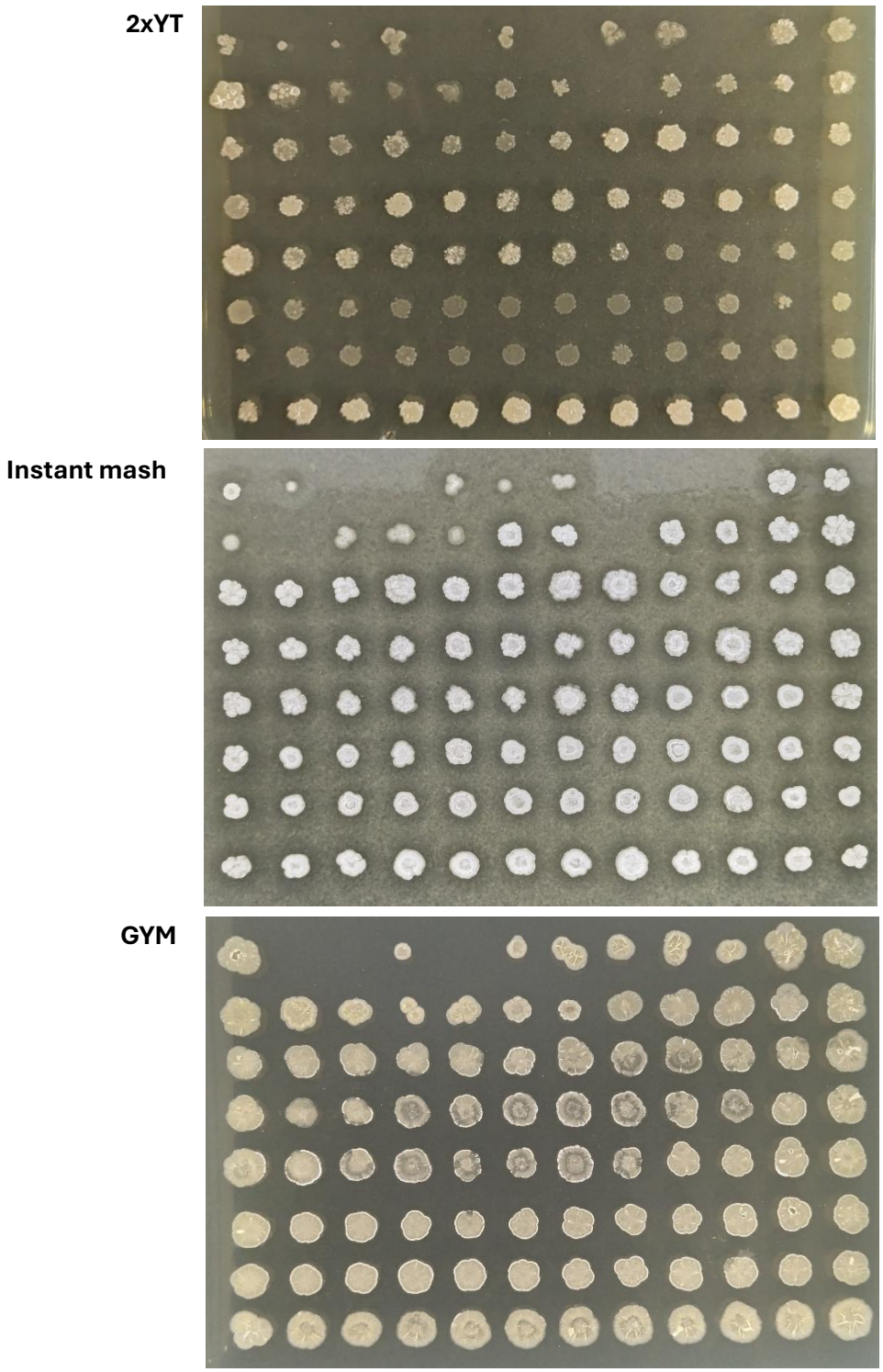


Figure 6.3. Colony morphology of fermented strains on 2xYT, instant mash, and GYM agars. Colonies selected by replica plating from different fermentations were transferred to 2xYT, instant mash, and GYM agars and grown for 10 days. Colonies are arranged in the same layout as previously described. Variation in colony size and sporulation is observed between strains and agars.

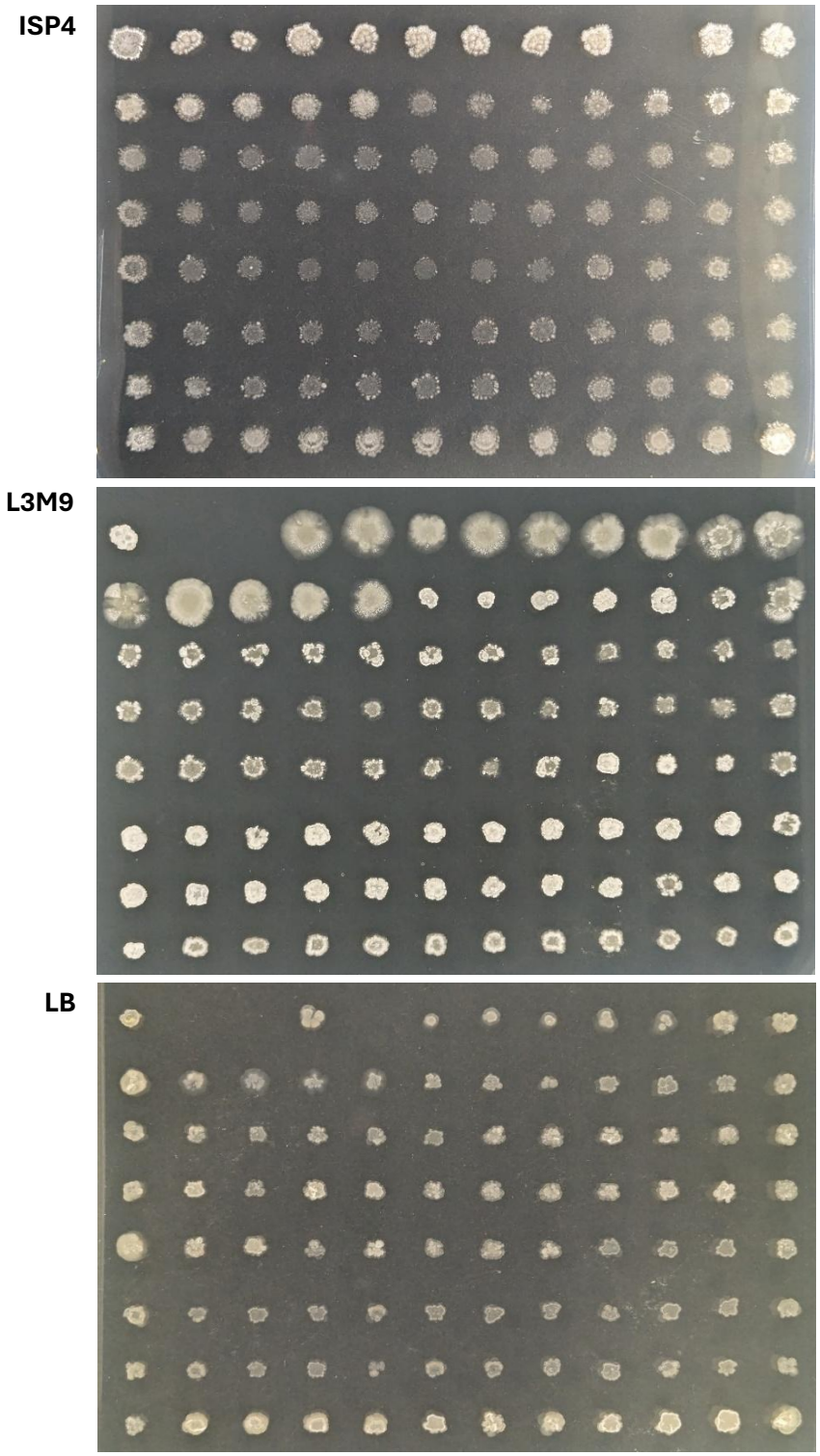


Figure 6.4. Colony morphology of fermented strains on ISP4, L3M9, and LB agars. Colonies selected by replica plating from different fermentations were transferred to ISP4, L3M9, and LB agars and grown for 10 days. Colonies are arranged in the same layout as previously described. Variation in colony size and sporulation is observed between strains and agars.

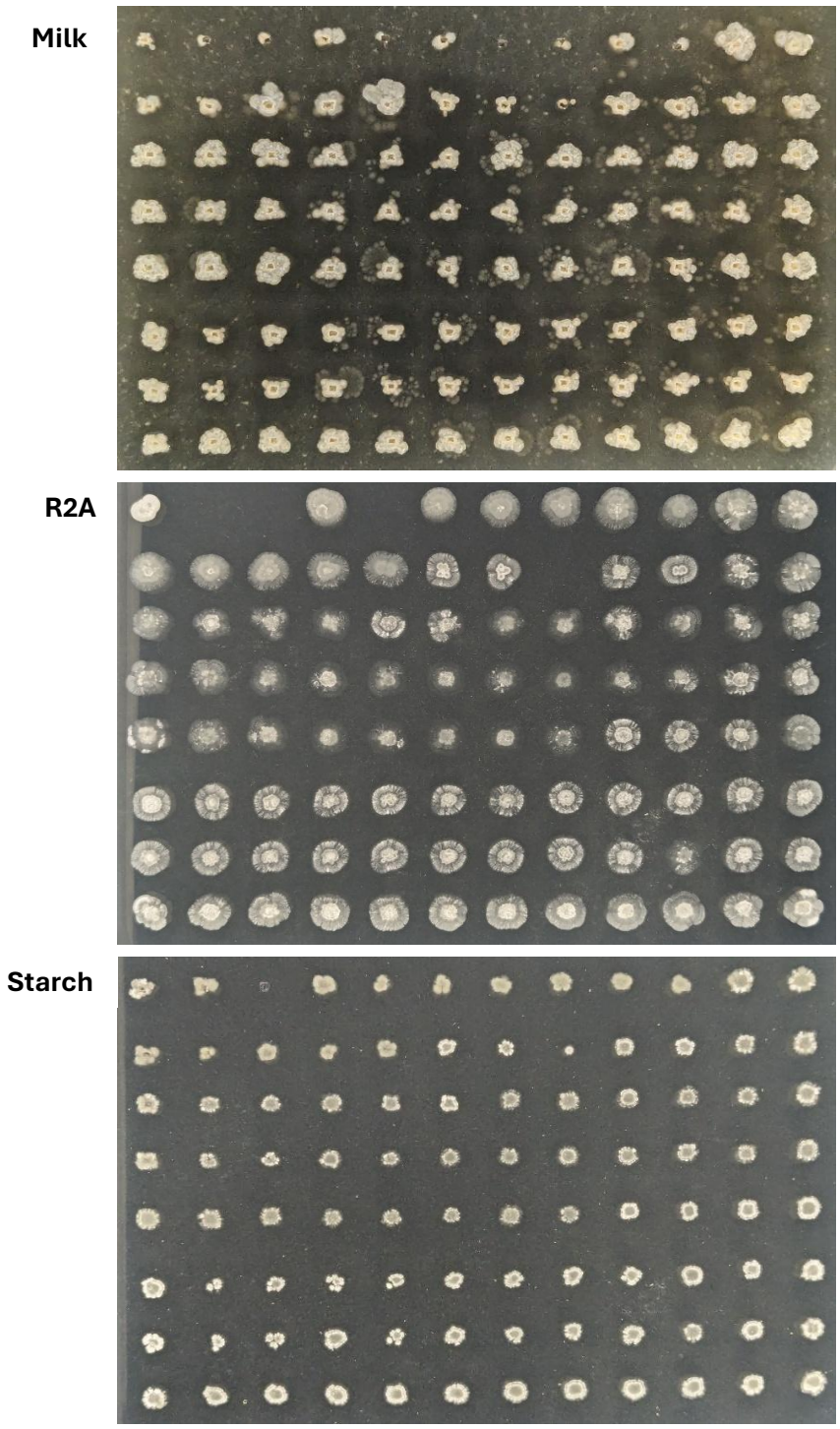


Fig 6.5. Colony morphology of fermented strains on milk, R2A, and starch agars. Colonies selected by replica plating from different fermentations were transferred to milk, R2A, and starch agars and grown for 10 days. Colonies are arranged in the same layout as previously described. Variation in colony size and sporulation is observed between strains and agars.

TSB

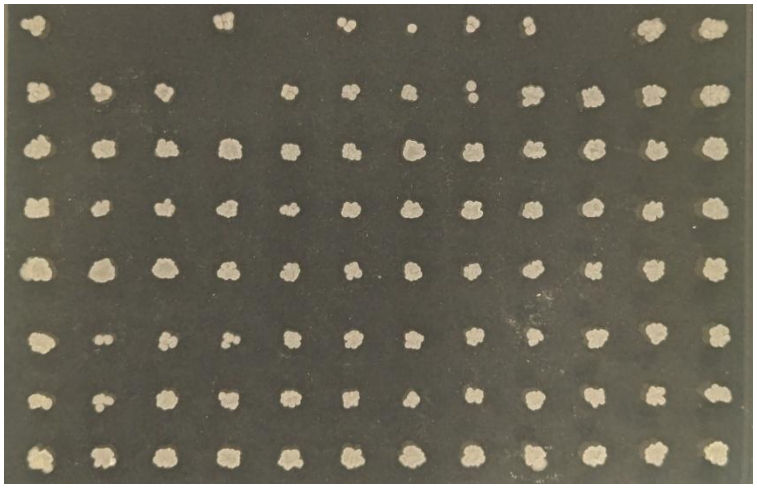


Figure 6.6. Colony morphology of fermented strains on TSB agar. Colonies selected by replica plating from different fermentations were transferred to TSB agar and grown for 10 days. Colonies are arranged in the same layout as previously described. Variation in colony size and sporulation is observed between strains.

To further analyse the colony morphology of the replica-plated fermentations (Figures 6.2 to 6.6), colony diameters were measured for all colonies with observable growth using ImageJ (Figures 6.7 and 6.8). This aimed to identify potential outliers within groups and assess whether colony morphology could distinguish specific strains or fermentation types.

Outlier analysis was performed on colony size measurements by treating each group on each agar type independently. Outliers were defined as values exceeding 1.5 times the interquartile range (IQR) from the group median. Colonies from the *S. clavuligerus* SC6 stock were analysed in parallel as a parental strain control to establish a baseline of natural variability on each medium and contextualise any observed deviations in the experimental groups.

On 2xYT agar, two SC6 colonies (16.7 and 14.9 mm) were flagged as outliers, reflecting a relatively high level of baseline variability. Similar outliers were observed in experimental groups, including colony diameters of 16.3 and 20.4 mm in the 7.5 L multistage group and 11.4 mm in the 1 L multistage group. These values fell within the range of variability already seen in SC6, suggesting that the experimental group deviations are not likely to represent meaningful biological differences.

On GYM agar, one SC6 colony (41.3 mm) was marked as an outlier. Additional outliers were observed in evolution experiment B and 7.5 L multistage fermentation groups. Again, the high variability already present in SC6 limits the interpretation of these deviations as functionally significant.

Across the remaining agars that supported growth, instant mash agar, ISP4, L3M9, LB, milk agar, starch agar, and TSB agar, 18 experimental group outliers were identified. However, most of these agars also produced outliers in the SC6 control, and in nearly all

cases, the size of the deviations in experimental groups was comparable to or smaller than those seen in SC6. For instance, on instant mash agar, a 7.5 L multistage colony measured 29.5 mm, 12.5 mm above the group median. This was matched by similarly large deviations in the SC6 control on the same agar. Likewise, SC6 outliers were observed on LB, milk, and starch agars, where experimental group outliers did not exceed this baseline variability.

R2A agar showed no outliers in any group, indicating low inherent variability on this medium.

Overall, pellet size measurements varied by medium, with *S. clavuligerus* SC6 often displaying the same or greater variability than experimental groups. This suggests that most observed outliers are the result of normal medium-specific variation rather than strain-specific or fermentation dependent differences.

That said, a small number of exceptions were observed where experimental groups displayed outliers on specific agars without any SC6 outliers. These included 1 L multistage outliers on ISP4, instant mash, and L3M9, and 7.5 L multistage outliers on L3M9 and LB. While infrequent and isolated, these cases may represent genuine deviations from baseline morphology and merit further investigation.

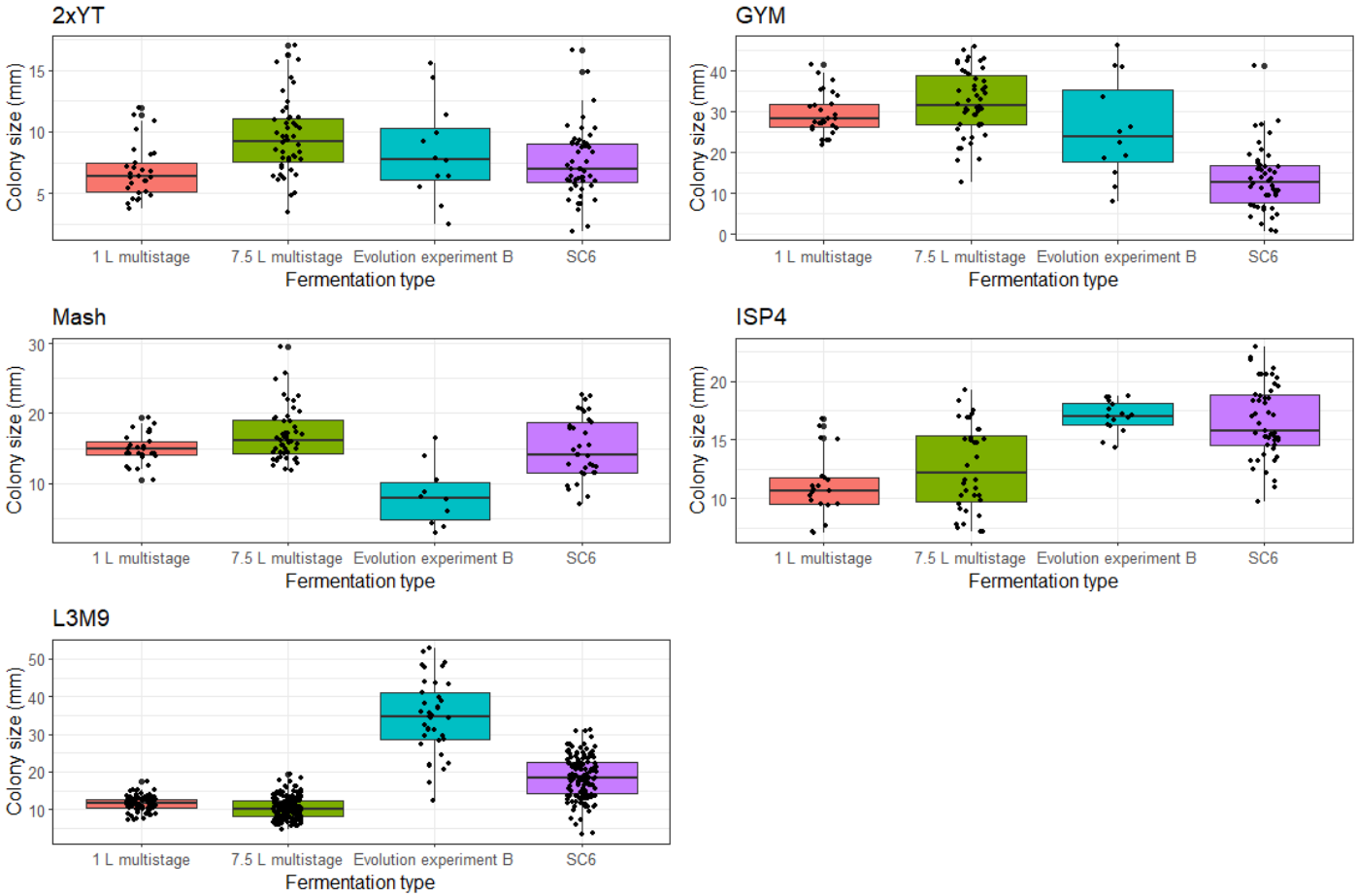


Figure 6.7. Colony size variation across fermentation groups on different agars. Colonies selected by replica plating from different fermentations were grown on various agars and grouped by fermentation type. Colony sizes were measured using ImageJ and compared to the parental strain *S. clavuligerus* SC6 on each agar. Outliers were identified as values >3 standard deviations from the group mean. P values for differences between groups on each agar are reported in Table 6.1.

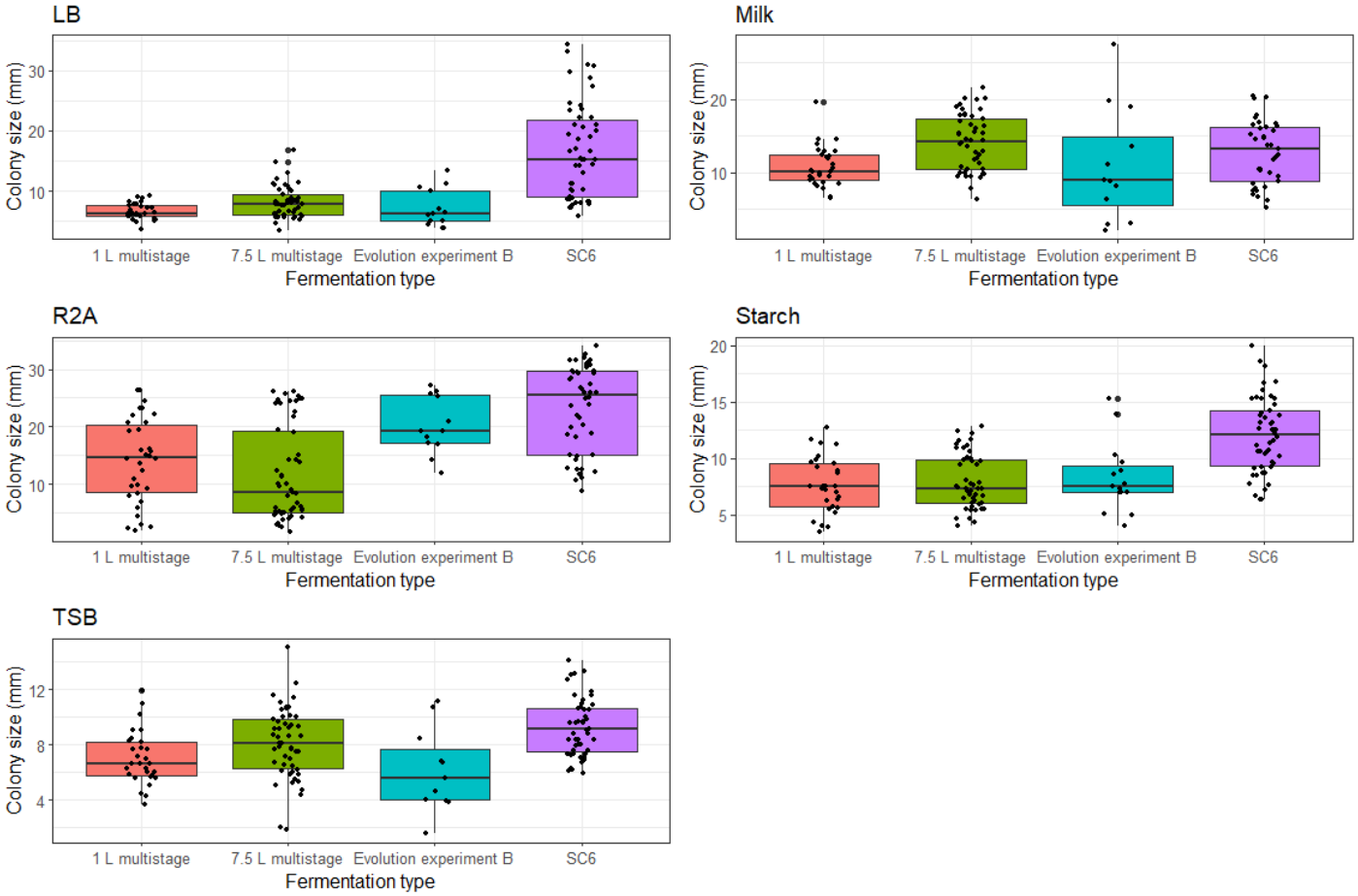


Figure 6.8. Colony size variation across fermentation groups on different agars. Colonies selected by replica plating from different fermentations were grown on various agars and grouped by fermentation type. Colony sizes were measured using ImageJ and compared to the parental strain *S. clavuligerus* SC6 on each agar. Outliers were identified as values >3 standard deviations from the group mean. P values for differences between groups on each agar are reported in Table 6.1.

Table 6.1. P values from Tukey tests comparing colony sizes between fermentation types on different agars. ANOVA tests showed significant differences in colony size between groups across all agars. Post hoc Tukey tests were performed to compare each fermentation type to parental strain *S. clavuligerus* SC6 on each agar. Reported p values reflect these pairwise comparisons.

Agar	Group	p-value
2xYT	7.5 L multistage	2.98×10^{-4}
	1 L multistage	0.253
	Evolution experiment B	0.438
GYM	7.5 L multistage	4.10×10^{-14}
	1 L multistage	5.18×10^{-11}
	Evolution experiment B	1.59×10^{-3}
ISP4	1 L multistage	1.15×10^{-8}
	7.5 L multistage	1.01×10^{-6}
	Evolution experiment B	0.482
L3M9	7.5 L multistage	1.01×10^{-28}
	1 L multistage	2.07×10^{-28}
	Evolution experiment B	4.83×10^{-12}
LB	1 L multistage	7.80×10^{-11}
	7.5 L multistage	1.96×10^{-9}
	Evolution experiment B	2.09×10^{-5}
Mash	Evolution experiment B	6.16×10^{-4}
	7.5 L multistage	0.0493
	1 L multistage	0.879
Milk	1 L multistage	0.0331
	7.5 L multistage	0.0850
	Evolution experiment B	0.466
R2A	7.5 L multistage	6.51×10^{-10}
	1 L multistage	6.32×10^{-6}
	Evolution experiment B	0.146
Starch	1 L multistage	4.45×10^{-9}
	7.5 L multistage	1.23×10^{-8}
	Evolution experiment B	4.51×10^{-4}
TSB	1 L multistage	1.30×10^{-5}
	Evolution experiment B	0.00647
	7.5 L multistage	0.0129

6.3 Genotypic analysis of fermented colonies

Given that outliers appeared to exist within groups but not consistently across media, further genotypic and phenotypic analysis was conducted on a subset of colonies to explore whether any individual strains exhibited unique characteristics. Five colonies were selected based on their distinct morphologies on L3M9 agar (Figure 6.1): Colony 19 displayed a small colony phenotype while still sporulating; colony 24 had a large colony phenotype; Colony 56 exhibited poor sporulation compared to other colonies isolated from the 1 L fermentations; Colony 60 also exhibited poor sporulation; and Colony 82 showed precocious sporulation around the colony edge, coupled with a small colony phenotype. Colonies from evolution experiment B (wells 2–16) were excluded from this selection as their phenotypes and genotypes were characterised in the previous chapter. Of the selected colonies, two were from the 7.5 L multistage fermentations (colonies 19 and 24), and three were from the 1 L multistage fermentations (colonies 56, 60, and 82).

To characterise their genotypes, genomic DNA was extracted from all five colonies after 48 hours of growth in TSB. Whole-genome sequencing was performed using the Illumina NovaSeq X Plus platform with at PE150 chemistry (Novogene UK). Using this data, the barcode identity of each colony was confirmed:

- 7.5 L multistage fermentation - Colony 19 – JCb17
- 7.5 L multistage fermentation - Colony 24 – JCb7
- 1 L multistage fermentation - Colony 56 – JCb3
- 1 L multistage fermentation - Colony 60 – JCb17
- 1 L multistage fermentation - Colony 82 – JCb14

Mutations were identified using breseq as previously described (Section 5.4), using *S. clavuligerus* SC6 as the reference genome. Mutations found in the parental strain control strains sequenced in this study and were treated as background and excluded from comparisons.

Overall, the five selected colonies had fewer mutations than those identified in the strains from evolution experiment B (Table 6.2). Notably, one colony from the 1 L fermentation was identical to the progenitor strain and the second colony (colony 82) exhibited only the loss of pSCL3. The three colonies from the 7.5 L fermentation, two (colonies 19 and 24) had lost pSCL3, although for one of them this was the only mutation detected. Additional mutations were found in genes encoding an ABC transporter (SCLAV_SC2_06140, Δ7 bp), DNA polymerase III subunit β (SCLAV_SC2_14090, C→G, Ser→Cys), alanine racemase (important for the conversion of L- to D-alanine for incorporation into the cell wall (Tassoni et al., 2017; SCLAV_SC2_18995, C→G, Val→Leu), ABC transporter periplasmic-binding protein YtfQ (SCLAV_SC2_24990, C→T, Ala→Val), and a non-coding region, 11bp upstream from L,D-transpeptidase family protein (SCLAV_SC2_14610, also involved in cell wall synthesis; Espaillat et al., 2024). While these genes are not part of a shared pathway or function, they are all involved in cellular resource allocation or stress adaptation. None are directly associated with central carbon metabolism. Notably with two ABC transporters being affected. It is also worth noting that within the 7.5 L fermentation, all barcoded strains maintained relatively constant abundance over time.

The potential impact of the non-synonymous SNPs was examined. The substitution in DNA polymerase III β (SCLAV_SC2_14090, Ser→Cys) does not fall within a

characterised catalytic motif of the β-clamp, and AlphaFold (Jumper et al., 2021) modelling did not provide high-confidence predictions for interpretation, its functional significance therefore remains unclear. The alanine racemase mutation (SCLAV_SC2_18995, Val→Leu) is a conservative change, and its position relative to the PLP-binding active site could not be determined with confidence, so the effect on enzyme function is uncertain. The YtfQ periplasmic-binding protein (SCLAV_SC2_24990, Ala→Val) mutation is located within the predicted binding protein domain, but outside well-defined conserved motifs, and again no confident structural predictions were obtained, leaving the impact unresolved. Finally, the tetratricopeptide repeat protein substitution (SCLAV_SC2_28310, Asp→Glu) is also conservative and, given limited annotation and low-confidence modelling, the functional consequences are unknown.

As loss of pSCL3 had previously been linked to potential differences in strain fitness, the correlation between pSCL3 loss and relative abundance in fermentation was examined. Colonies 19, 24, and 56 started at relative abundances of 2.75%, 5.04%, and 5.13%, respectively, and ended at 15.07%, 3.97%, and 4.97%, corresponding to $\log_2$ fold changes of 1.85, –0.38, and –0.008. These differences suggest that pSCL3 loss alone does not predictably confer a competitive advantage.

Within the 1 L fermentation, both colony 60 (no detected mutations) and colony 82 (pSCL3 loss only) showed declines in relative abundance over the course of fermentation: 6.2% to 5.2% (–0.25 log2FC) and 6.7% to 3.5% (–0.68 log2FC), respectively. This again suggests that pSCL3 loss in isolation does not strongly influence strain dynamics in this context.

Table 6.2. Mutations detected in genomes of colonies isolated from different fermentations. Five colonies selected from replica plating of multiple fermentations were whole genome sequenced. Mutations relative to the SC6 parental strain genome were identified using Breseq. Distinct mutations were detected in four of the five isolates.

Replicon	Position	Mutation	Gene
Colony 19 (7.5 L multistage fermentation)			
Chrm	1,335,051	Δ7 bp	ABC transporter integral membrane transport protein
Chrm	3,223,673	C→G	DNA polymerase III subunit β
pSCL3	1	Δ444,894 bp	-
Colony 24 (7.5 L multistage fermentation)			
pSCL3	1	Δ444,894 bp	-
Colony 56 (7.5 L multistage fermentation)			
Chrm	3,351,757	(GCTGCC)8→7	Noncoding
Chrm	4,321,528	C→G	Alanine racemase
Chrm	5,810,797	C→T	ABC transporter periplasmic-binding protein YtfQ
Colony 60 (1 L multistage fermentation)			
None detected			
Colony 82 (1 L multistage fermentation)			
pSCL3	1	Δ444,894 bp	-

Following the repeated loss of pSCL3 in several strains, the presence and relative copy number of all four plasmids were examined in previously sequenced strains. Illumina reads from parental strain strains (JCb6, JCb10, and JCb17) and the five selected colonies were mapped to the chromosome and four plasmids of the *S. clavuligerus* SC6 reference genome using Bowtie2 (Langmead and Salzberg, 2012). The average coverage, including zero-coverage bases, was calculated for each replicon using Bedtools (Quinlan and Hall, 2010). Plasmid coverage was then normalised to the chromosomal coverage for each strain. These normalised values were plotted, using the average of the three parental strains as the control (Figure 6.9). For downstream analysis, colonies from the same fermentation were grouped together.

An ANOVA showed no differences in copy number between groups for pSCL1, pSCL2, or pSCL4 ($p = 0.284$, 0.941, and 0.995, respectively). In contrast, pSCL3 displayed a reduced copy number in the fermentation-derived strains compared to the parent strain, although it is unlikely that the difference between strains is significant ($p = 0.083$). Among the plasmids, pSCL1 exhibited the greatest variation. Normalised coverage in parental strain, 1 L, and 7.5 L multistage fermentation strains was 6.63x, 19.8x, and 11.04x, respectively.

Read coverage of the plasmid pSCL2 remained consistent across groups, ranging from 9.2x to 9.9x. This contrasts with the previous chapter, where pSCL2 coverage dropped to 3.27x and 5.47x in strains that dominant the fermentation and those strains that become rare, respectively. Similarly, pSCL4 showed minimal variation between groups (1.25–1.26x), compared to 0.98x and 1.20x in dominant strains and rare strains from the previous chapter. As observed before, pSCL1 showed the highest variability in copy number across strains.

The loss of pSCL3 identified by breseq was confirmed by this coverage-based analysis. Normalised coverage for pSCL3 dropped from 1.56x in parental strain strains to 0.23x and 0.27x in the 1 L and 7.5 L fermentation-derived strains, respectively. These residual low levels of coverage are explained by the presence of pSCL3 in colonies 56 and 60, which retained the plasmid despite others having lost it entirely.

Additionally, JBrowse was used to detect any new junctions or large-scale deletions across replicons. Consistent with breseq analysis, no new junctions or large-scale deletion were detected within replicons for any strain.

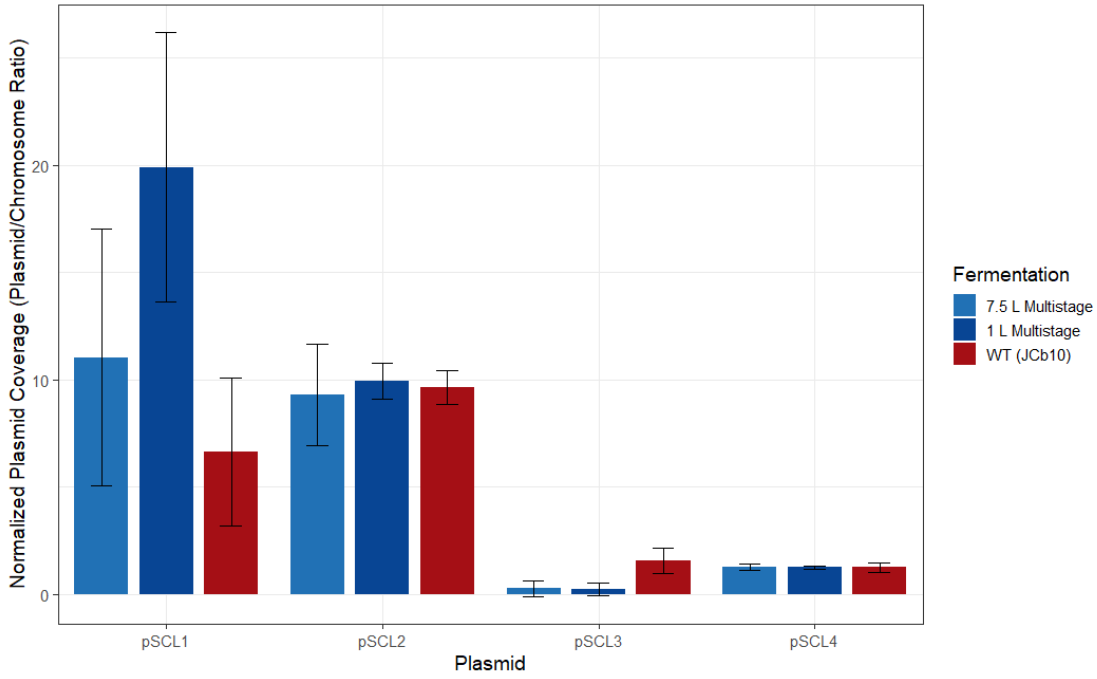


Figure 6.9. Plasmid copy number colonies selected by replica plating from different fermentations. Five colonies selected by replica plating from different fermentations were grown in TSB for 48 hours, genomic DNA was extracted and whole genome sequenced. Read coverage of plasmids pSCL1–4 was normalised to chromosome coverage to estimate plasmid copy number. Colonies were grouped by fermentation type, error bars indicate standard deviations. ANOVA tests showed no significant differences between groups for pSCL1, pSCL2, and pSCL4 ($p = 0.284$, 0.941, 0.995, respectively), while pSCL3 showed reduced copy number compared to the parental strain ($p = 0.083$).

6.4 Phenotypic analysis of strains from fermentations

After identifying genomic differences between the five selected strains and their parent strain, further phenotypic analyses were carried out to investigate possible genotype to phenotype links following fermentation. Although no mutations were found in the CA biosynthetic genes of these strains, previous transcriptomic data from evolution experiment B showed downregulation of CA genes post-adaptation. Measuring CA yield in these five colonies would therefore provide insight into whether similar phenotypic shifts also occurred here.

The five selected strains and parental strain JCb10 were grown for 48 hours in 50 mL TSB. CA titres were then calculated and normalised to dry weight (Figure 6.10). CA yield dropped in all strains compared to JCb10, which produced 14.00 µg/mg cell dry weight. In contrast, the mean titres for the 1 L and 7.5 L multistage fermentation strains were 7.81 and 11.23 µg/mg cell dry weight, respectively. An ANOVA indicated differences between groups ($p = 3.5 \times 10^{-4}$). A Tukey test identified a differences between the 1 L and 7.5 L strains ($p = 0.0012$), whereas differences between each group and JCb10 were not significant ($p = 0.0714$ and 0.4364 for 1 L and 7.5 L vs JCb10, respectively).

Outlier analysis of CA measurements was conducted using the 1.5xIQR method, treating each strain independently. No outliers were detected in any of the three groups. Shapiro–Wilk tests confirmed normality across all groups ($p > 0.05$), indicating that CA measurements were consistent and that no biologically unusual values were present.

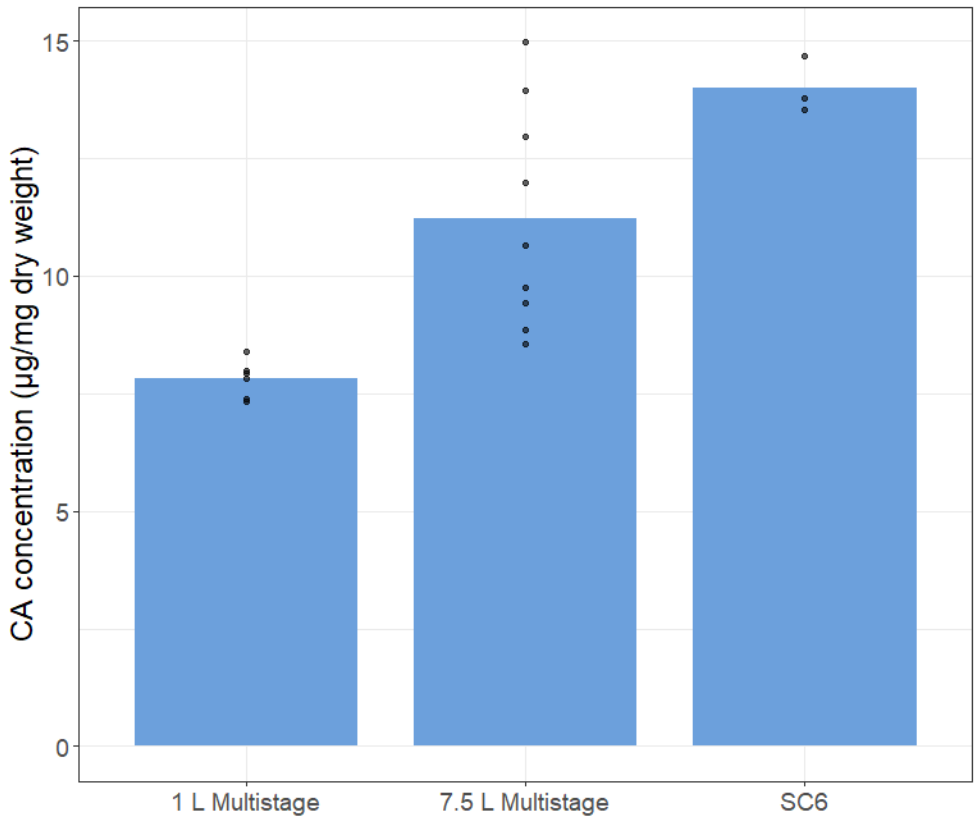


Figure 6.10 Clavulanic acid titre in colonies selected by replica plating from different fermentations. Five colonies selected by replica plating were grown in TSB for 48 hours, and clavulanic acid (CA) titre was measured. Colonies were grouped by fermentation type. Strains from 1 L and 7.5 L multistage fermentations showed reduced CA output compared to parental strain *S. clavuligerus* SC6. ANOVA indicated significant differences between groups ($p = 3.5 \times 10^{-4}$). A Tukey test showed a significant difference between 1 L and 7.5 L multistage strains ($p = 0.0012$), while comparisons to SC6 were less significant ($p = 0.0714$ and 0.4364 for 1 L and 7.5 L vs SC6, respectively). No outliers were detected within groups.

In addition to attempting to detect outliers both between and within groups using CA production, G6PDH activity was measured to understand levels of primary metabolism within strains, where G6PDH acts as branch point between two major carbon catabolic pathways of glycolysis and the pentose phosphate pathway. This was based on observations in the previous chapter, where modest differences in G6PDH activity were seen across evolved strains (Figure 5.11), alongside transcriptomic changes in genes involved in central carbon metabolism. The aim was to determine whether similar trends could be observed following multistage fermentations and to detect within group outliers.

The five selected strains and the parental JCb10 were grown in 50 mL TSB for 48 hours. G6PDH activity was then assessed using a glucose 6-phosphate dehydrogenase Activity Assay Kit (Fluorometric, Abcam), with values normalised to total protein concentration (Bradford assay), as recommended by the manufacturer.

Both 1 L and 7.5 L multistage fermentation strains showed increased G6PDH activity relative to the parental strain (Figure 6.11). *S. clavuligerus* SC6 parent strain displayed a mean activity of 13.24 nmol/min/mL/mg total protein, compared to 17.27 and 30.58 nmol/min/mL/mg for 1 L and 7.5 L strains, respectively These results are consistent with G6PDH activity found within the literature of ~20 nmol/min/mL/mg total protein ((Obanye et al., 1996). Additionally, these values are similar to those reported in the previous chapter for dominant and rare strains (25.18 and 16.33 nmol/min/mL/mg, respectively).

An ANOVA detected differences between groups ($P = 9.6 \times 1^{-4}$). Mann–Whitney tests returned P values of 0.108 (1 L vs 7.5 L), 1.000 (1 L vs SC6), and 0.0012 (7.5 L vs SC6),

indicating levels of G6PDH activity were elevated in strains derived from the 7.5 L multistage fermentation strains.

Outlier analysis using the 1.5xIQR method identified no outliers in any group. Shapiro–Wilk tests confirmed that G6PDH activity values followed a normal distribution, suggesting measurements were consistent within groups and that no biologically unusual values were present.

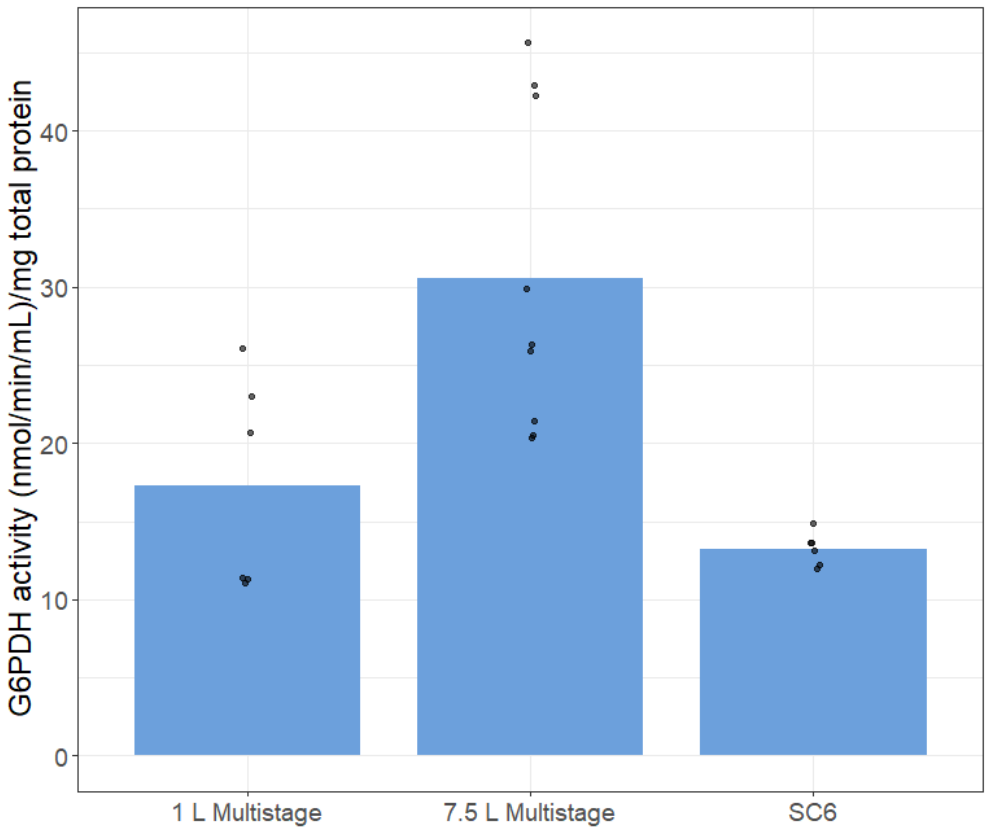


Figure 6.11. Glucose-6-phosphate dehydrogenase activity in colonies selected by replica plating from different fermentations. Five colonies selected by replica plating were grown in TSB for 48 hours, and glucose-6-phosphate dehydrogenase (G6PDH) activity was measured and normalised to total protein concentration (Bradford assay). Colonies were grouped by fermentation type. ANOVA detected differences between groups ($p = 9.6 \times 10^{-4}$). Mann–Whitney tests gave p values of 0.108 (1 L vs 7.5 L), 1.000 (1 L vs SC6), and 0.0012 (7.5 L vs SC6), indicating elevated G6PDH activity in 7.5 L strains. No outliers were detected within groups.

Given that CA yield and G6PDH activity assays did not identify any outliers within groups, specific growth rate, μ, were analysed to determine whether these could help detect outliers or provide further insight into phenotypic variability. Specific growth rates may also reflect competitive ability during fermentation, providing another potential link between genotype and phenotype.

The five selected strains and parental strain JCb10 were grown in 50 mL TSB for 48 hours, and backscatter measurements were taken using the aquilabiolabs Cell Growth Quantifier. The exponential growth phase was used to calculate Specific growth rate, which were then plotted (Figure 6.12).

As seen previously (Figure 5.12), both 1 L and 7.5 L multistage fermentation strains displayed lower Specific growth rate (μ) than the parental strain. The mean μ for both 1 L and 7.5 L strains was 2.1×10^{-3} h^{-1}, compared to 0.0082 h^{-1} for the parental JCb10. For context, the dominant and rare isolates in the previous chapter showed μ of 0.0031 and 0.0046 h^{-1}, respectively (Table 5.2).

An ANOVA confirmed significant differences between groups ($p = 7.36 \times 10^{-8}$). No significant difference was detected between the 1 L and 7.5 L strains ($p = 0.864$), but both groups exhibited an approximate 3.8-fold reduction in μ compared to the parental ($p = 0.001$).

Outlier analysis using the 1.5xIQR method identified no outliers in any group. Shapiro–Wilk tests indicated no significant deviation from normality, and boxplot inspection confirmed symmetric distributions. These results suggest that μ measurements were consistent within each group, with no biologically unusual values detected.

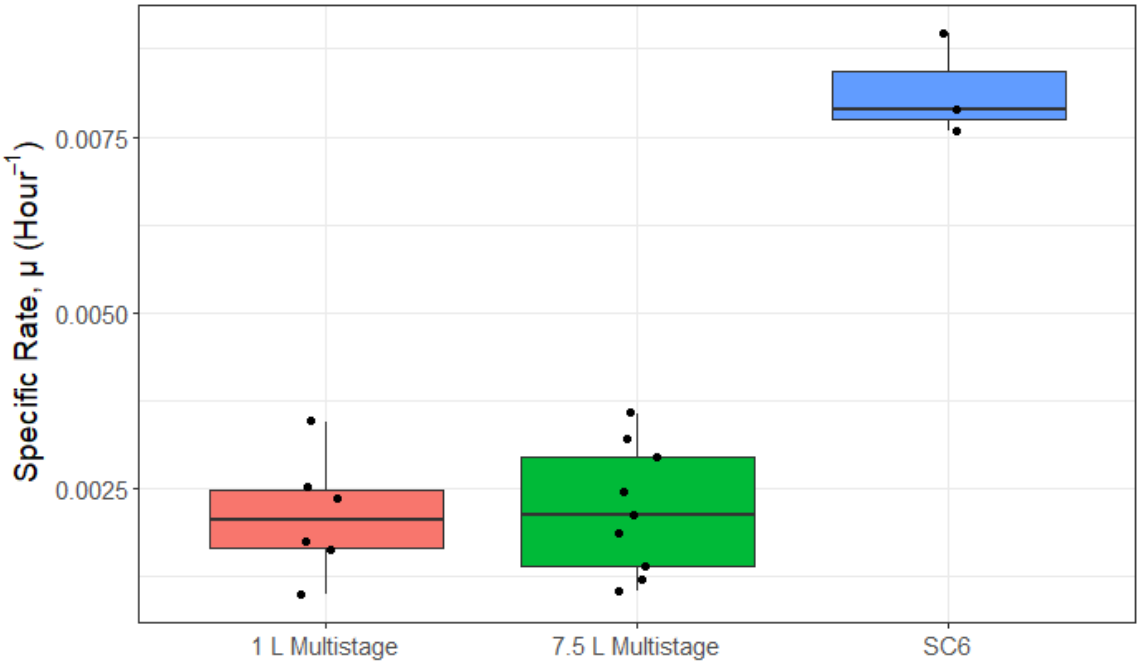


Fig 6.12. Specific growth rates of colonies selected by replica plating from different fermentations. Five colonies selected by replica plating were grown in TSB for 48 hours, and backscatter was measured using a cell growth quantifier to calculate specific growth rates (μ). Colonies were grouped by fermentation type. ANOVA indicated significant differences between groups ($p = 7.36 \times 10^{-8}$). Strains from 1 L and 7.5 L multistage fermentations showed similar specific growth rates ($p = 0.864$) but both were significantly lower, approximately 3.8-fold reduction, compared to parental *S. clavuligerus* SC6 ($p = 0.001$). No outliers were detected within groups.

As CA yield, G6PDH activity, and μ did not reveal outliers within groups, strains were grown in liquid culture and imaged to assess whether adaptation affected pellet morphology. The five selected strains and parental strain JCb10 were grown in 50 mL TSB for 48 hours, after which they were imaged using phase contrast microscopy. One representative image was selected for each strain (Figure 6.13). It was immediately apparent that strains from the 1 L and 7.5 L multistage fermentations produced larger pellets than the parental strain, consistent with colonies from evolution experiment B.

Hyphae were observed extending from the edge of each pellet, consistent with typical *Streptomyces* morphology in liquid culture. Qualitative inspection found no obvious outliers between samples: although some pellets appeared larger or had associated free hyphae, no sample showed consistent or extreme deviation. Notably, while no pellet formation was observed during growth in CM5 medium, these same strains formed pellets when grown in TSB medium, indicating a medium-dependent switch in phenotype.

To assess pellet size more quantitatively, ImageJ was used to measure pellet diameters (Figure 6.14). Strains were grouped by fermentation type. Parental strain S. *clavuligerus* SC6 had the smallest average pellet diameter (128 μm), while the 1 L and 7.5 L multistage fermentation strains had average diameters of 205 μm and 154 μm, representing 1.6-fold and 1.2-fold increases, respectively. These values were markedly lower than the average 1,001 μm diameter observed for dominant strains from evolution experiment B in the previous chapter.

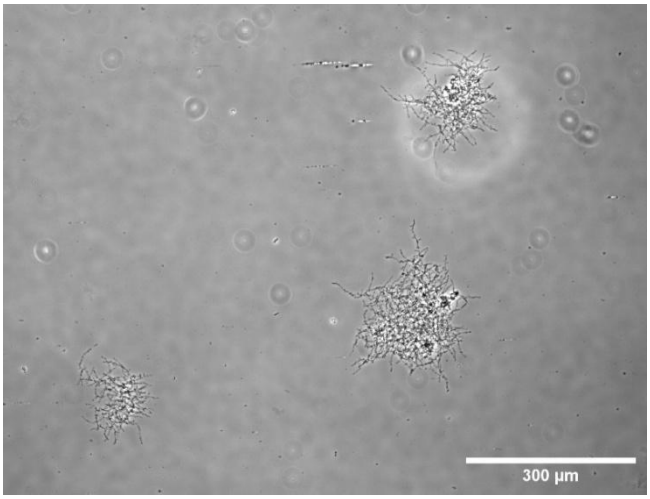
An ANOVA confirmed significant differences between groups ($p = 1.52 \times 10^{-6}$). Mann–Whitney tests indicated significant differences between 1 L and 7.5 L strains (p =

0.00021) and between 1 L and parental strain SC6 ($p = 3.00 \times 10^{-6}$), but no significant difference between 7.5 L and SC6 (P = 0.197).

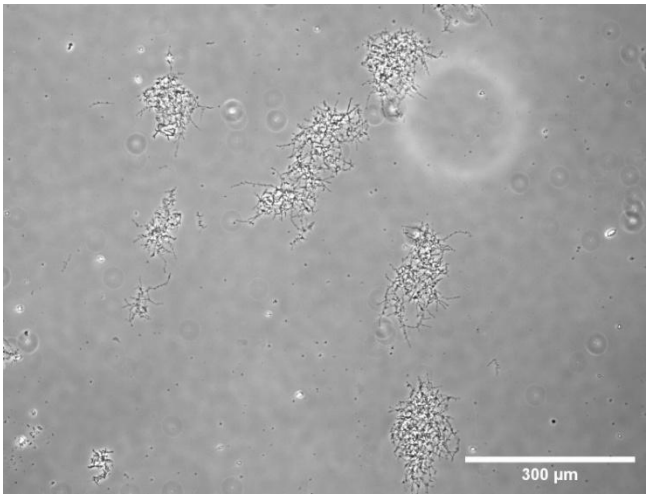
Outlier analysis (1.5xIQR) revealed no outliers in the parental strain group, while the 7.5 L group showed four outliers (399.3, 345.5, 393.6, and 326.5 μm), and the 1 L group showed three (463.9, 508.2, and 371.9 μm). Shapiro–Wilk tests showed that pellet size distributions for both 1 L ($P = 9.6 \times 10^{-5}$) and 7.5 L ($P = 4.6 \times 10^{-6}$) groups deviated significantly from normality, whereas parental strain SC6 did not (P = 0.54).

These results suggest that pellet morphology differs both between and within groups. Parental strain pellets were small and tightly distributed, while multistage fermentation strains showed increased pellet size and greater variability. Some pellets in these groups exceeded the internal variability expected for normally distributed data. While the nature of pellet sampling prevented re-isolation of specific outlier colonies, these findings suggest that the multistage strains contain a mixture of phenotypes and that pellet formation is influenced by both growth conditions and strain background.

7.5 L multistage



1 L multistage



JCb10 (parental)

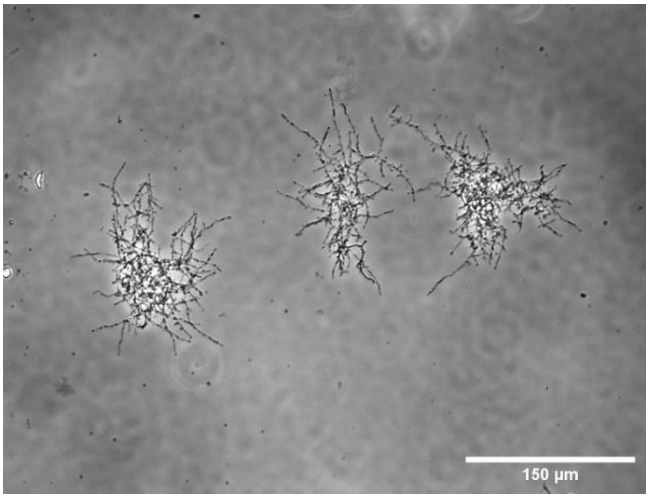


Figure 6.13. Pellet morphology of colonies selected by replica plating from different fermentations. Five colonies selected by replica plating from different fermentations were grown in TSB for 48 hours. Pellets were imaged by phase contrast microscopy. One representative image is shown for each fermentation type.

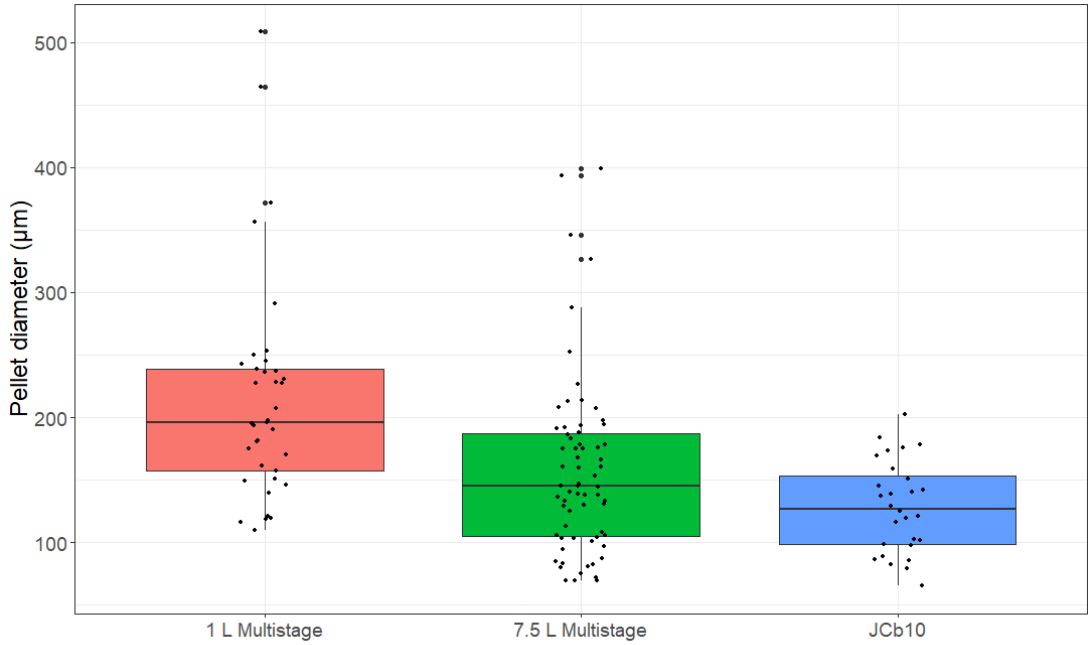


Figure 6.14. Pellet size distribution of colonies selected by replica plating from different fermentations. Five colonies selected by replica plating from different fermentations were grown in TSB for 48 hours, and pellet diameters were measured from phase contrast images using ImageJ. Colonies were grouped by fermentation type. Parental strain *S. clavuligerus* SC6 showed the smallest average pellet diameter (128 µm), with 1.6-fold and 1.2-fold increases observed for 1 L and 7.5 L multistage fermentation strains, respectively. ANOVA detected significant differences between groups ($p = 1.52 \times 10^{-6}$). Mann–Whitney tests showed significant differences between 1 L and 7.5 L strains ($p = 0.00021$) and between 1 L and SC6 ($p = 3.00 \times 10^{-6}$), but no significant difference between 7.5 L and SC6 (P = 0.197). Some outliers were detected within groups.

.

6.5 Further phenotypic analysis of strains from fermentations

As no consistent outliers were identified within any group in Section 6.4, attempts to link genotype to phenotype in the 1 L and 7.5 L multistage fermentation strains were not possible. To improve the likelihood of detecting phenotypic outliers within a group, the number of colonies analysed was increased. Given that the 1 L multistage fermentation showed greater variability in population dynamics than the 7.5 L fermentation (chapter 4), 10 colonies were randomly selected from a 1 L multistage fermentation for further analysis. Increasing the number of replicates aimed to improve the chances of identifying a consistent outlier and potentially linking phenotype to genotype within this fermentation group.

Each of the 10 colonies was grown for 48 hours in 50 mL TSB, then transferred to CM5 medium for clavulanic acid production assays. CM5 was chosen because it matched the original fermentation conditions, potentially increasing the likelihood of revealing phenotypic differences that would not be observed in TSB. However, as cultures were grown in CM5 rather than TSB, clavulanic acid yield could not be normalised to dry weight.

The average clavulanic acid titre was 54.81 µg/mL (Figure 6.15). Statistical analysis using the Wilcoxon signed-rank test and Grubbs' test detected no outliers (P = 1.00 for both tests). Thus, both visually and statistically, no evidence of phenotypic outliers was found within these 10 colonies.

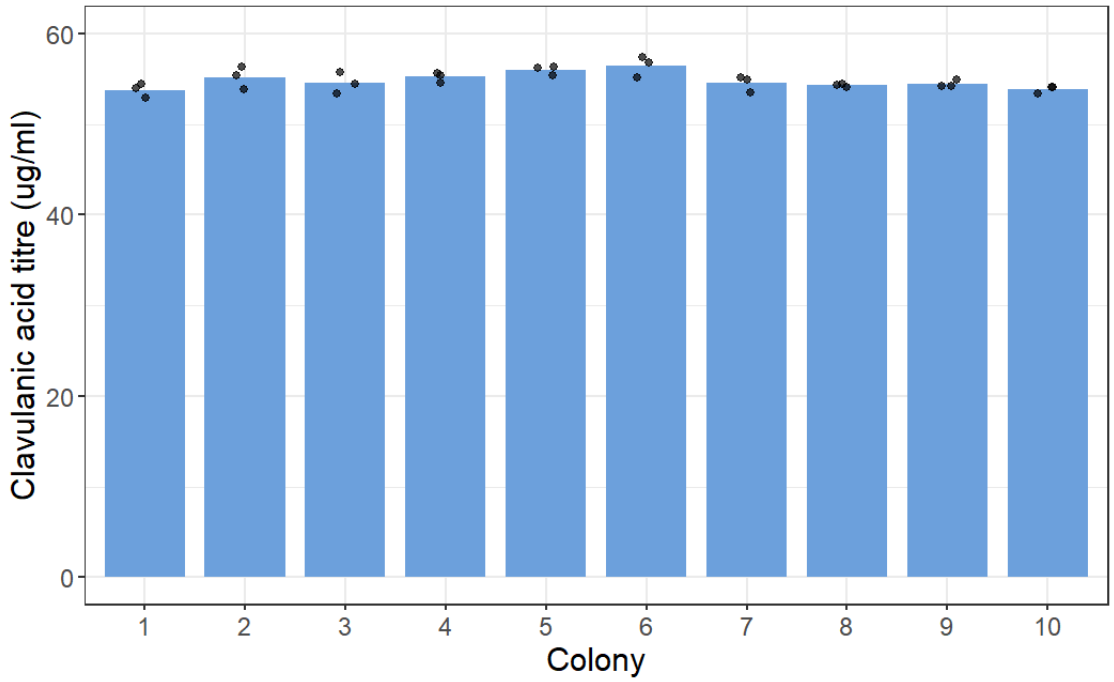


Figure 6.15. Clavulanic acid titre of randomly selected colonies from a 1 L multistage fermentation. Ten colonies randomly selected from a 1 L multistage fermentation were grown in TSB for 48 hours, and clavulanic acid (CA) titre was measured. The average CA titre was 54.81 µg/mL. Wilcoxon signed-rank and Grubbs tests detected no outliers (p = 1 for both tests).

As no outliers were detected within the 10 colonies using a CA assay, further analysis was performed to identify any phenotypic outliers. A bioactivity assay was therefore designed, which would also provide additional insight into the CA production of these strains and detect any differences in production of other bioactive metabolites. The assay was based on previous work (Gallardo, 2020).

The 10 randomly selected colonies from the 1 L multistage fermentation were grown as confluent lawns on L3M9 agar for 10 days. Triplicate agar plugs from each lawn were transferred onto soft agar seeded with *Micrococcus luteus* at an OD_{600} of 0.001. The plates were incubated for a further 24 hours, after which photographs were taken (example plates shown in Figure 6.16). Each plate also included a negative control (L3M9 agar plug) and a positive control (Oxoid 30 µg amoxicillin/clavulanic acid susceptibility disc).

The diameter of the inhibition zones was measured for each plug, and the average inhibition zone for each colony was plotted (Figure 6.17). An ANOVA test detected no significant differences between colonies ($P = 0.100$). The average zone of inhibition was 1.89 cm, with the smallest and largest means being 1.47 cm and 2.23 cm, respectively.

As neither this bioactivity assay nor the earlier CA assay revealed a consistent phenotypic outlier among the 10 colonies, no further analyses were conducted on this group.

Figure 6.16. Bioactivity assay of colonies from a 1 L multistage fermentation against *M. luteus*. Ten colonies randomly selected from a 1 L multistage fermentation were grown as confluent lawns on agar for 10 days. Triplicate agar plugs from each lawn were transferred onto soft agar seeded with *M. luteus* and incubated for a further 24 hours. Three representative plates are shown, each containing (top to bottom) an L3M9 agar plug, three test plugs, and a 30 µg amoxicillin/clavulanic acid susceptibility disc (Oxoid) as control.

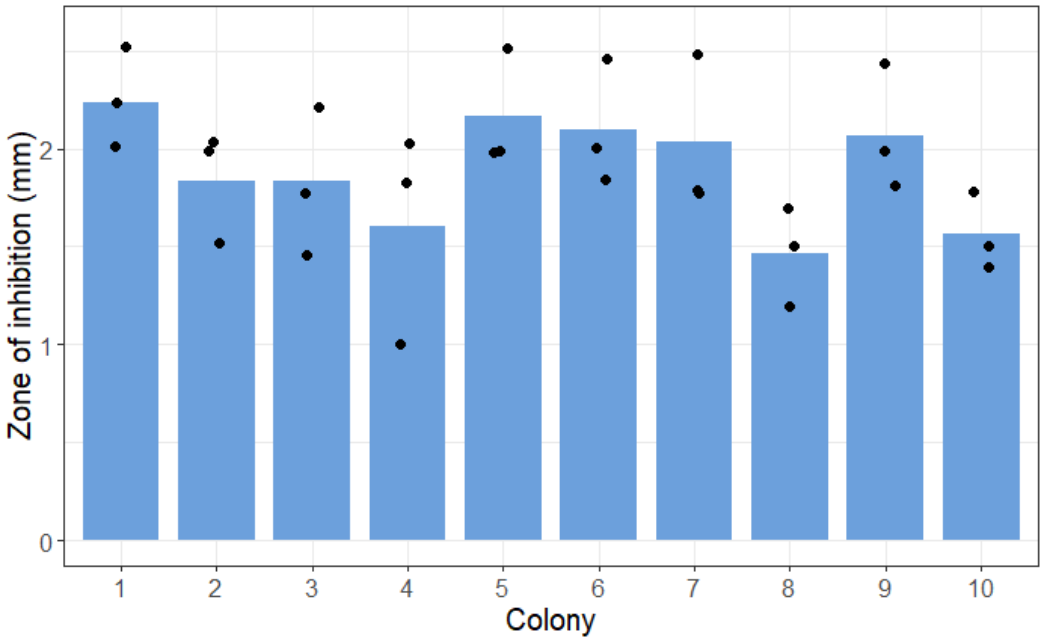


Figure 6.17. Zones of inhibition for lawns of *Micrococcus luteus* do not significantly differ between strains in bioactivity assay. Plates from the M. luteus bioactivity assay (Figure 6.16) were analysed by measuring the inhibition zones of each triplicate agar plug. Individual measurements are shown as black points, with averages for each isolate indicated by blue bars. ANOVA detected no significant differences between groups ($p = 0.100$).

6.6 Going back to GSK fermentations

As no phenotypic variants could be identified within the 1 L or 7.5 L multistage fermentations, further genotypic analysis was conducted on additional isolates from the 7.5 L fermentation. This would expand upon the three previously sequenced isolates (Table 6.2) and help assess the diversity of mutations arising during industrial fermentations. The original three genomes already displayed variability in both the number and types of mutations detected, so sequencing more isolates would provide a broader picture of the genetic changes occurring under these conditions.

To do this, the 7.5 L multistage fermentation was plated on to L3M9 agar (Figure 4.15 A3), and five additional colonies were randomly selected. These isolates were grown in TSB for 48 hours, and genomic DNA was extracted and sequenced using the Illumina NovaSeq X Plus at PE150 (Novogene UK). FASTQ analysis revealed the corresponding barcodes for each isolate as follows:

- 7.5 L multistage fermentation - Colony 1 – JCb17
- 7.5 L multistage fermentation - Colony 2 – JCb1
- 7.5 L multistage fermentation - Colony 3 – JCb9
- 7.5 L multistage fermentation - Colony 4 – JCb2
- 7.5 L multistage fermentation - Colony 5 – JCb2

Breseq was used to identify genetic differences between these strains and the *S. clavuligerus* SC6 reference genome. The mutational profile of the parental strain strains was used as a baseline, and all further mutations were reported relative to this.

All mutations detected in these further five strains were nonsynonymous in nature (Table 6.3). Three of the five new isolates had lost pSCL3, as observed previously. In total, five

out of eight sequenced isolates from this fermentation lacked pSCL3. However, in contrast to the earlier isolates, one of the newly sequenced colonies had a genotype identical to the parental strain.

One mutation was identified in an ATP-dependent DNA helicase UvrD1, an important enzyme in DNA repair (Chadda et al., 2022; SCLAV_SC2_18110, (AGCCCGG)7→8). Notably, two isolates had mutations in ABC transporter proteins. One of these mutations was unique (SCLAV_SC2_05800, +C). However, the second ABC transporter protein mutation has been seen exactly within the previously sequenced isolates (SCLAV_SC2_06140, Δ7 bp). This is of interest since one of the previously sequenced colonies (colony 19, JCb17) exhibit exactly the same mutation. Although colony 1 here is also JCb17, the mutational profiles differed. In the earlier isolate, mutations were seen in the ABC transporter, DNA polymerase III subunit β, and a loss of pSCL3. In the newly selected strains with the same barcode, only the ABC transporter mutation was detected, and occurred in over 99.5% of reads mapping to that region, highlighting the intra-strain heterogeneity. Generally ABC transporters are important due to the broad range of subsrates they transport across the membrane (Rees et al., 2009). Additionally in isolate 3 a AfsR/SARP family transcriptional regulator mutation occurred (SCLAV_SC2_14845 ; Table 6.3) which itself is important for the regulation of antibiotic biosynthesis and interestingly contains a tetratricopeptide repeat (TPR) domain (Wang et al., 2024) as seen in a previously mutated gene (Table 5.1) .

As before, none of the mutations identified were in genes directly linked to central carbon metabolism but are involved in transport and regulation. However, as previously noted, these mutations do not appear to confer a fitness advantage during fermentation over the measured timescales. The 7.5 L multistage fermentation remained stable in terms of

population dynamics, with no strain becoming dominant (Figures 4.14 and 4.15). The $\log_2$ fold changes for colonies 1–5 was 1.85, 0.85, 0.28, and –0.27 (colonies 4 and 5), respectively.

Table 6.3. Mutations detected in additional isolates from a 7.5 L multistage fermentation. Five colonies isolated from a 7.5 L multistage fermentation were whole genome sequenced. Mutations relative to the SC6 parental strain genome were identified using Breseq. Distinct mutations were detected in four of the five isolates.

Replicon	Position	Mutation	Gene
Colony 1			
Chromosome	1,335,051	Δ7 bp	ABC transporter integral membrane transport protein
Colony 2			
Chromosome	1,256,545	+C	ABC transporter ATP-binding protein
pSCL3	1	Δ444,894 bp	-
Colony 3			
Chromosome	3,404,243	(CCGTTGGTG)4→5	AfsR/SARP family transcriptional regulator
pSCL3	1	Δ444,894 bp	-
Colony 4			
None detected			
Colony 5			
Chromosome	4,119,689	(AGCCCGG)7→8	ATP-dependent DNA helicase UvrD1
pSCL3	1	Δ444,894 bp	-

6.7 Summary

Compared to the work conducted within chapter 5, using the model of experimental evolution, fermentations at a more industrially relevant timescale do not reliably result in changes to cell phenotype. This chapter evaluated whether strain-level differences arising from fermentation could be detected through colony morphology. Here, colony size, pigmentation, and sporulation patterns were assessed across 94 isolates from multiple fermentation conditions, including a 1 L multistage fermentation, a 7.5 L multistage fermentation, and evolution experiment B. The aim was to determine whether phenotypic traits could be linked with genotype in strains following fermentation.

Colonies were replica plated onto L3M9 agar and a panel of 19 different nutrient agars. Clear group-level differences were observed. For example, colonies from evolution experiment B were visibly larger and more pigmented, while those from the 7.5 L fermentation often showed reduced sporulation. Within-group variability was high, and in many cases, the range of colony sizes and morphologies overlapped substantially between fermentation conditions. Additionally, a range of mutations were seen within different strains within the 7.5 L fermentation, seemingly with none affecting strain fitness within the fermentation, despite key regulatory and stress response genes being affected.

To establish a baseline for natural morphological variation, an isogenic stock of *S. clavuligerus* SC6 was aliquoted into 48 wells and grown on the same set of agars. These controls showed a similar degree of variability to the experimental groups. For example, differences in colony size, colony edge smoothness, and sporulation were common, even in genetically identical SC6 colonies grown on the same plate. These observations were consistent across several media types, particularly on GYM and L3M9 where sporulation and pigment production were more readily expressed. In contrast, LB and

starch agars produced small colonies with limited morphological differentiation, likely due to nutrient limitation or poor growth. Nonetheless, the differences within parental strain colonies make confidently determining phenotypic variations between fermented colonies difficult.

Quantitative analysis of colony size using ImageJ supported these observations. While some colonies were flagged as outliers based on interquartile range cutoffs, nearly all of these fell within the variability already observed in the SC6 control. This suggests that most phenotypic differences seen in strains do not exceed the baseline noise of clonal populations. In a small number of cases, such as a few large colonies from the studied 7.5 L fermentation, phenotypes exceeded the variability seen in controls and may reflect genuine adaptive differences.

In summary, colony morphology reflects both environmental and genetic inputs but lacks the resolution or reliability needed to distinguish strain type on its own. Even among strains that underwent significant genotypic change (e.g. JCb10 from evolution experiment B), morphology was not sufficiently distinct to enable identification without molecular methods. These results highlight the limits of morphology as a screening tool in *S. clavuligerus* and reinforce the importance of pairing phenotypic assessments with barcode sequencing, especially in industrially relevant timescale. While colony morphology can still be informative when used carefully and in the right context, it should not be relied upon as a primary tool for monitoring population structure in *Streptomyces* fermentations.

Chapter 7 Discussion

7.1 General discussion

Industrial strain development is a constant process aimed at improving fermentation yields of valuable compounds. For example, GSK routinely generates new *S. clavuligerus* strains for 100,000 L fermentations to increase clavulanic acid production (Huckle & Kendrew, Personal Communication). A central challenge is that even closely related strains differ in performance, meaning population structure directly affects productivity. In this work, competition assays showed fitness variation even between isogenic strains (Figures 4.18–4.20), with evolved isolates consistently outcompeting the parental strain in CM5 but to differing extents. This uneven success suggests division of labour within populations, where some subpopulations specialise in costly functions such as antibiotic production while others prioritise growth (Westhoff et al., 2020; Zhang et al., 2020). Although this can benefit the collective, it also risks skewing population structure and undermining yield stability under industrial conditions (Harms et al., 2017; Hoskisson and Hobbs, 2005).

Long-term evolution studies in other systems illustrate how quickly adaptation reshapes populations. The *E. coli* LTEE shows predictable fitness gains and occasional novel traits such as citrate utilisation under constant conditions (Lenski, 2023, 2017; Lenski et al., 1991). Additionally, the *Streptomyces* LTEE demonstrates a reduction in pelleting, improved relative fitness and reduced secondary metabolite production after 1000 generations (Munnoch et al., 2025). Barcoded lineage tracking in yeast similarly highlights rapid, environment-dependent selection (Larcombe et al., 2023). *Streptomyces* species are predisposed to such dynamics given their large G+C-rich linear chromosomes, unstable chromosome ends, and mobile plasmids, all of which promote rearrangements and deletions (Birch et al., 1991; Zhang et al., 2022). In *S.*

clavuligerus, the megaplasmid pSCL4 interacts closely with the chromosome due to essential replication functions being present (Algora and Herron, 2019; Mohit, 2023) although wider genome plasticity is apparent (Munnoch et al., in prep).

Against this background, the current study applied Bar-seq to track strain-level dynamics across media, scales, and timescales, alongside a 152-generation evolution experiment with genome and transcriptome profiling. Several consistent patterns emerged. First, short-term stability depended strongly on media and scale. TSB fermentations became unstable with scale-up, whereas multistage industrial conditions (S2A–CM5) remained stable across 50 mL, 1 L, and 7.5 L fermentations (Figure 4.21). Most variability arose during the seed stage of multistage fermentations (Figures 4.13 and 4.16), which likely shapes downstream population structure. This has not previously been demonstrated in *Streptomyces* industrial fermentations, representing a novel finding. However, a reduced seed train within *E. coli* has been shown to positively impact 'process economics' (Okonkowski et al., 2005). Second, during long-term evolution in CM5, barcoded populations diverged, with different strains becoming dominant, rare strains persisting, and declines in clavulanic acid titres across all populations (Figure 5.1). These shifts were linked to specific mutations and altered transcription (Section 5.4 and 5.5). Third, phenotypic traits such as colony morphology and sporulation varied across evolved isolates, but they were unreliable indicators of underlying genomic or productivity changes (Section 5.6). This difficulty in using an altered genotype to predict fitness or secondary metabolite production is well documented and has been observed in strains evolved within the *E. coli* LTEE (Favate et al., 2023).

Together, these findings show that population dynamics in *S. clavuligerus* fermentations are shaped by both short-term process conditions and longer-term evolutionary change.

Media composition and scale can stabilise or destabilise strain frequencies, while prolonged adaptation reduces productivity despite apparently stable population structures. Phenotypic screening alone cannot capture these dynamics. Bar-seq therefore provides a link between genotype, phenotype, and productivity, offering a practical approach for monitoring strain stability in industrial fermentations.

7.2 Bar-seq can be used to reliably track fermentations

Bar-seq was adapted for use in *S. clavuligerus* using 20 isogenic strains, each marked with a unique 20 bp barcode integrated using pSET152. Integration at the ϕC31 attB site was stable (Figure 3.7), barcodes were sufficiently distinct (Figure 3.8), and mixed inputs were recovered at expected ratios from spores or DNA (Figures 3.14–3.15). Thus, Bar-seq provides a quantitative measure of strain relative abundance at a resolution and throughput beyond plating, flow cytometry, or strain-specific qPCR (Delvigne and Goffin, 2014; Hernández et al., 2020; Qamer et al., 2003). Conceptually, Bar-seq relates to signature tagged mutagenesis (STM) and Tn-seq (Saenz and Dehio, 2005; van Opijnen et al., 2009), but the neutral barcodes avoid mutational confounders and suit *Streptomyces*, where transformation and conjugation of complex molecular tools can be limiting for new systems (Krysenko, 2025). This therefore extends barcoded lineage tracking, established in yeast, into bacteria (Larcombe et al., 2023).

Critically, Bar-seq captured dynamics invisible to phenotype-only methods. Post-fermentation isolates exhibited limited variation in colony morphology and growth (Chapter 6). However, population composition shifted over time. Bar-seq therefore aids the link between genotype and population dynamics in mixed cultures. Industrially, it

enables head-to-head comparison of candidate strains inside the actual competitive environment (Wehrs et al., 2020).

Limitations of the technique are inherent. Bar-seq tracks frequencies of strains, not what causes them to change frequencies. Furthermore, intra-strain mutants within one barcoded strain are masked into a single trajectory (Table 6.3). Detection of very rare lineages (<1%) is also limited by depth, and DNA from dead cells can, in principle, contribute reads (Amar et al., 2021). The number of barcoded strains also constrains coverage of diversity. These limitations may be mitigated through deeper sequencing, larger barcode sets, further whole-genome sequencing of lineages, and targeted single-colony WGS to recover intra-lineage heterogeneity. However, Bar-seq is a robust, high-resolution tool for monitoring *Streptomyces* fermentations and for allowing competition into routine strain evaluation, in agreement with previous studies (Ferrari et al., 2021; Robinson et al., 2014; Wehrs et al., 2020).

7.3 Scale-up of TSB fermentation leads to instability

TSB fermentations at 50 mL remained stable over 48–72 h (Figures 3.17, 4.1). Instability emerged at 1 L over five days (Figure 4.2). Here, several lineages declined towards extinction while others rose, with different ‘winners’ in each replicate. That replicate specificity indicates strong batch-to-batch variability not seen at 50 mL.

Mechanistically, larger shake flasks intensify heterogeneity. Oxygen transfer and mixing are poorer than in small volumes, generating micro-environments with different oxygen and substrate availabilities (Garcia-Ochoa and Gomez, 2009; Rastädter et al., 2023; Xia et al., 2021). Such gradients can amplify small fitness differences over multi-day runs, particularly in filamentous systems where mycelial pellet size and fragmentation alter

mass transfer (Dragosits and Mattanovich, 2013). Inoculation and early growth also impose founder effects. For example, small sampling variance or slight early growth advantages can set divergent trajectories that persist (Hagan et al., 2024). Prior work shows similar divergence under oxygen-limited or spatially structured conditions in other microbes (Reisman, 1993; Zhou et al., 2018).

This means TSB is not a reliable media to use when scaled at larger volumes. Uncontrolled gradients and stochasticity drive instability that will confound reproducibility and mask true strain performance. The observed population dynamics motivated the focus on multistage industrial conditions designed to reduce heterogeneity (Huckle & Kendrew, Personal Communication). The contrast between 1 L TSB and industrial media is evidence that larger media volumes do not themselves cause instability but is how the system is scaled and what medium is used that are decisive, as has been previously seen (Viana Marques et al., 2018; Wehrs et al., 2020).

7.4 Scale-up of multi-stage industrial fermentations does not cause instability

Multistage fermentations remained stable at 50 mL, 1 L, and 7.5 L scales (Figures 4.3–4.9). In some instances, 1 L was more stable than 50 mL (however a greater number of replicates are needed to validate this) consistent with the idea that process control and media formulation, not volume, are the greatest influences on population stability.

CM5 is formulated for balanced nutrient delivery and buffered conditions that support secondary metabolism without rapid exhaustion of key resources (Huckle & Kendrew, Personal Communication). Stirred tanks provide controlled agitation and aeration, reducing spatial heterogeneity (Rastädter et al., 2023; Schmidt, 2005; Xia et al., 2021). Together, these features suppress the local selective sweeps common in rich media that

promoted fast growth such as TSB. Slower, more even growth reduces fitness differentials, limiting the chance that one lineage finds a unique niche and runs away.

This answers a key question for GSK, does population instability explain declining CA titres at production scale? The findings seen here suggested it does not. Under conditions closely mimicking industrial fermentation, strain relative abundance is stable (Figures 4.14 and 4.15), so variability in CA likely reflects physiology (e.g. regulation and stress response) or process problems rather than a non-producing strain becoming dominant within the fermentation. The stability observed also validates 1 L and 7.5 L pilots as reliable predictors of 100,000 L performance, at least with respect to population structure.

7.5 Most instability in industrial fermentations arises from the seed (S2A) stage

Across multistage fermentations, Bar-seq revealed the seed stage (S2A) as the main source of variability, populations stabilised after transfer into CM5 (Figures 4.13 and 4.16). Therefore, instability largely originates early and is then carried forward rather than arises in production medium.

This is consistent with both biology and process. Germination is heterogeneous, and S2A selects for rapid biomass accumulation under different nutrients to CM5, so small differences at inoculation can be magnified during exponential expansion (Deindoerfer and Humphrey, 1959; Wentzel et al., 2012). Additionally, the transfer itself imposes a bottleneck, even a 10% inoculum can create sampling variance such that high-frequency seed lineages disproportionately found the CM5 culture causing a founder effect (Hagan et al., 2024). In population genetics, serial bottlenecks accelerate drift and selection, especially under high growth (Mu and Zhang, 2023). These data mirror this

where changes in population dynamics in S2A tended to persist, while CM5 rarely generated new 'winners' or 'losers'.

The industrial implication of this is a requirement to better control the seed culture. Strategies include adjusting S2A composition to avoid strong selection, shortening seed duration to limit generations (and hence selection/drift), and standardising inoculation procedures to reduce founder variance (Rastädter et al., 2023; Xia et al., 2021). Sampling more timepoints during seed and immediately post-transfer would better resolve early dynamics. A limitation of this study is that it focuses on barcoded SC6 derivatives. Assessing current production strains will test if the findings presented here are also relevant to current production strains. Nonetheless, managing selection within S2A is key to improve downstream reproducibility.

7.6 Industrial fermentation conditions can drive adaptation

Whole genome sequencing detected mutations arising within industrially relevant fermentations, sometimes with different mutations in separate colonies from the same barcode lineage (Table 6.3). Thus, adaptation can occur rapidly and heterogeneously even over short production-like timescales as seen in previous studies (Rugbjerg et al., 2018). Plasmid instability was also clear, notably frequent loss of pSCL3 (Tables 5.5, 6.9). Prior work documents large-scale rearrangements and plasmid loss in *S. clavuligerus* and related *Streptomyces* (Algora and Herron, 2019; Mohit, 2023; Pérez-Llarena et al., 1997; Petković et al., 2006) (Munnoch in prep), consistent with dynamic linear replicons and chromosome ends containing telomeric repeats (Kirby et al., 1975; Lorenzi et al., 2021; Mohit, 2023).

Most mutations mapped to regulatory and accessory genes rather than core biosynthetic loci (Tables 5.1, 6.2 and 6.3) demonstrating selection acts via stress management, metabolic allocation, and interaction with nutrient resources (Bibb, 2005). Genomic plasticity of this sort, including compensatory changes upon plasmid curing, has precedent in Actinomycetes (Schrempf, 1982).

The key finding here is that genetic adaptation can occur even when strain-level population dynamics appear stable (e.g. in CM5). Gross phenotypic assays in isolation will miss this finding. Periodic genome checks, linked to Bar-seq, provide data of emerging genotypes with potential process impact (Wehrs et al., 2020).

7.7 Phenotypic variation occurs but is difficult to map to genotype

Replica plating and biochemical assays showed variation in sporulation, colony morphology, and CA titre among isolates (Figures 5.10–5.14 and Figures 6.2-6.14), but most effects were modest and often within the baseline variability of the parental control (Figure 6.1). No clear hyper- or non-producing strains emerged. Two factors likely contributed to this. Firstly, phenotype was measured in rich media (TSB) rather than CM5 due to practical constraints, condition-specific traits may therefore be under-represented. Secondly the phenotypic assays completed here were relatively limited in number (titre, growth, morphology, selected enzyme tests), so metabolic or stress adaptations that shift fitness without obvious colony-level signatures may have been missed (Boruta, 2021).

Linking genotype to phenotype was equally as challenging. Few mutations were observed and tended to be associated to regulation rather than in CA biosynthesis genes as is consistent with previous findings (Bibb, 2005; Santamarta et al., 2011).

Streptomyces regulatory architecture is layered and pleiotropic. Single regulatory changes can influence dozens of targets spanning primary and secondary metabolism. Prior studies in Actinomycetes show small regulatory differences producing context-dependent changes in development and antibiotic output (Bibb, 2005; Hoskisson and Fernández-Martínez, 2018).

Transcriptomics did not necessarily implicate cyclic-di-GMP signalling (Section 5.5), a master regulator of growth, sporulation and secondary metabolism in the adaptation of barcoded strains (Gallagher et al., 2024, 2020; Tschowri et al., 2014). This pleiotropy involved with this gene explains why there may have been difficulty in linking and single genotypic change to a single phenotypic change. Multi-omics would aid this potential link under relevant conditions. Combining Bar-seq with transcriptomics, metabolomics and targeted biochemical assays in CM5 would connect regulatory state to function (Delneri, 2010; Payen et al., 2016). As a limitation, our phenotyping in rich media and limited assay breadth likely masked phenotype to genotype links present in production settings.

7.8 Evolution experiment shows non-parallel adaptation

Over 152 generations, replicates of an evolution experiment showed different outcomes (Figure 5.1). One was dominated by a single lineage, one ended with co-dominant strains, and one with no strain becoming dominant. This non-parallelism contrasts with frequent convergence in the *E. coli* LTEE, where similar mutations often arise across replicates (Blount et al., 2020; Lenski et al., 1991). The difference likely reflects *S. clavuligerus* complexity where multiple linear replicons, a large genome, and complex

global regulation expand the adaptive potential and promote divergent outcomes (Medema et al., 2010; Song et al., 2010b; Zhang et al., 2022).

Plasmid burden is a plausible method of adaptation here. Although plasmid-deficient strains have been engineered and growth impacts measured (Mohit, 2023), fitness benefits of specific plasmid losses (e.g. pSCL3) appear context-dependent in these data. No universal advantage of pSCL3 loss was seen, implying media/process and epistasis are important. However, with only 152 generations, it may be that early divergence is observed prior to any longer-term convergence that sometimes emerges in LTEEs and *Streptomyces* serial passage studies (Lind et al., 2015; McDonald et al., 2009; Munnoch et al., 2025). It is highly recommended that future work includes an extended evolution experiment using these barcoded strains to see if parallelism emerges between replicates.

For industrial purposes, the data seen here urges caution. A single, repeatable adaptive trajectory can not necessarily be predicted within fermentation. Monitoring should therefore be by replicate and over time, not one-off. Methodologically, extending duration and sampling frequency, and sequencing across all replicates (not just one as done here), will reveal whether parallelism increases with time or remains low due to persistent adaptation (Lenski et al., 1991).

7.9 Evolution experiment reveals genomic and transcriptomic changes, with transcriptomics dominating

Genomic analysis identified few mutations across evolved isolates, mainly in regulatory or metabolic genes, plus plasmid loss (Table 5.1). In contrast, transcriptomics revealed broad changes in stress responses, carbon metabolism, and secondary metabolism

(Figures 5.6–5.9). Thus, adaptation was primarily regulatory where small sequence changes (if any) yielded large expression shifts, consistent with regulatory control of development and metabolism in *Streptomyces* (Gallagher et al., 2024, 2020).

This fits classical trajectories where early fitness gains derive from tuning expression rather than gaining novel functions (Lenski, 2017). Evolved *S. clavuligerus* did not acquire new capabilities, such as citrate use within the *E. coli* LTEE (Blount et al., 2020), but appear to have reallocated resources and improved stress handling. Demonstrating an incremental optimisation within existing networks compared to broad metabolic shifts (Chater, 2001; McCormick and Flärdh, 2012). Practically, that means substantial physiological change can occur without many obvious genomic adaptations, demonstrating a more in-depth exploration into *Streptomyces* adaptation is needed in the future.

A key limitation to this study is scope. Transcriptomics was only applied to the evolution experiment, not to more industrially relevant bioreactor fermentations. Although justified by the absence of large lineage shifts in CM5, adding omic based studies to bioreactors would contextualise stability at the expression level (Delneri, 2010; Payen et al., 2016). More generally these data show genome sequencing alone risks underestimating adaptation. Regulatory networks and their effects on the rest of the cell must be tracked. Coupling Bar-seq with more frequent transcriptomics and targeted metabolomics would provide an early-warning system for shifts that precede titre changes.

7.10 Adaptations produce subtle phenotypic shifts but confer major competitive advantages

Despite substantial transcriptomic changes, evolved isolates differed only modestly from the ancestor in CA titres and μ when assayed in isolation (Figures 5.10–5.14). No mutations were detected in CA biosynthetic genes. Yet in direct competition they consistently outcompeted the ancestor in CM5 (not in TSB), revealing sizeable fitness advantages that monoculture tests masked (Figure 5.15).

Mechanistically, this is consistent with small improvements in resource allocation, stress tolerance, or timing of developmental transitions, changes that carry little penalty in monoculture but compound into clear relative fitness gains in mixed culture (Lenski, 2017). These are small 'uphill' steps on the fitness landscape via regulatory tuning rather than leaps to new phenotypes (McCandlish, 2011; Papkou et al., 2023). The pattern reflects reports of 'quasi-cheater' behaviour where lineages that slightly reduce costly secondary metabolism or shift timing to improve growth under production conditions (Rugbjerg et al., 2018). While isolates seen here are not non-producers, early changes to that end could still reduce net CA yield if such lineages dominate.

For industry this demonstrates monoculture assays are necessary but insufficient to truly investigate fermentation dynamics. Small differences in titre or growth are not reliable proxies for competitive success. Incorporating competition-based evaluation, using Bar-seq, provides a more faithful test of robustness under the exact selective forces present in production media (Finkel and Kolter, 1999).

7.11 Competition assays are essential to detect hidden fitness changes

Across both fermentation experiments and the evolution study, competition assays exposed differences that monoculture phenotypes did not (Figures 4.18–4.20). They showed lineages with similar growth curves or titres in isolation showed clear, repeatable relative fitness differences in mixed culture (Figures 4.17–4.19, 5.15 and 5.16). Because relative abundance integrates many components of fitness (μ, yield, stress responsiveness, resource efficiency), competition assays are a useful tool (Finkel and Kolter, 1999; Lenski et al., 1998).

For pipelines, the message is practical. Relying on titres, endpoint biomass, or colony morphology can miss adaptative phenotypic changes that only matter in ecological context. Competition assays act as a stress test under realistic selection, revealing otherwise undetectable advantages early. Coupling competition with Bar-seq gives lineage-resolved quantification at scale, enabling multiplexed head-to-head evaluation of candidate strains (Robinson et al., 2014; Wehrs et al., 2020). This is aligned with long-standing practice in experimental evolution where relative fitness, not absolute growth alone, defines adaptation (Lenski et al., 1998, 1991).

It would be ideal to embed small-scale, Bar-seq based competition into routine screening, especially in CM5 or other production-like media. Positive controls (parental vs parental) and replicate competitions should be included to quantify baseline noise and founder variance. A limitation is that competition outcomes can depend on initial frequency and inoculum structure, frequency-dependent designs and reciprocal frequency tests can address this (Ribeck and Lenski, 2015).

7.12 Future work

There are several clear next steps that would build on this work. The evolution experiment here ran for 152 generations (Figure 5.1), but *S. clavuligerus* has the capacity for further adaptation and longer experiments would likely reveal additional changes. Running the experiment for more generations would help to test whether adaptation continues through gradual regulatory shifts or if larger genomic changes eventually appear, as in the *E. coli* LTEE or the *Streptomyces* LTEE (Lenski, 2023; Munnoch et al., 2025).

It will also be important to repeat these experiments with strains that are more industrially relevant than SC6. This would allow direct comparison to current production strains and help determine whether the dynamics observed here are typical of industrial strains or specific to SC6. Running similar evolution experiments with model organisms such as *S. coelicolor* or *S. venezuelae* would also make it easier to link adaptation to existing datasets. This would also allow the study of the evolutionary dynamics across species.

Only one replicate population from the evolution experiment was studied with genomics or transcriptomics. Extending this analysis to the replicate experiments would give a more comprehensive view of how parallel these genomic and transcriptomic mutations are. As transcriptomics revealed far more about strain behaviour than genomics alone, other omics layers should also be explored. Proteomics, lipidomics and metabolomics could all identify changes in metabolism or physiology that were not apparent here.

Finally, given Bar-seq has been shown to be a powerful tool to study population dynamics, Bar-seq itself can now be applied much more broadly. It allows direct competition assays between strains that cannot be told apart by phenotype alone and

that metagenomics cannot separate. This could be used not only in *Streptomyces* fermentations, but also in experiments on phage–host interactions, fungal interactions, or mixed bacterial communities.

7.13 Conclusion

This thesis has shown that Bar-seq can be used to reliably track *S. clavuligerus* populations during fermentation and that it reveals dynamics invisible to traditional methods. Strain instability was found mainly in TSB and seed stages, whereas populations in production-relevant CM5 media remained stable even at scale, meaning instability is not the main driver of variable CA titres. Long-term experiments demonstrated that adaptation still occurs, primarily through subtle regulatory changes that gave strong fitness advantages without obvious phenotypic shifts. These results highlight the need for tools like Bar-seq, competition assays and integrated omics to properly understand and manage strain behaviour in industrial fermentations.

References

Abbate, E., Andrion, J., Apel, A., Biggs, M., Chaves, J., Cheung, K., Ciesla, A., Clark-ElSayed, A., Clay, M., Contridas, R., Fox, R., Hein, G., Held, D., Horwitz, A., Jenkins, S., Kalbarczyk, K., Krishnamurthy, N., Mirsiaghi, M., Noon, K., Rowe, M., Shepherd, T., Tarasava, K., Tarasow, T.M., Thacker, D., Villa, G., Yerramsetty, K., 2023. Optimizing the strain engineering process for industrial-scale production of bio-based molecules. J Ind Microbiol Biotechnol 50, kuad025. https://doi.org/10.1093/jimb/kuad025

Abraham, E.P., Chain, E., 1940. An Enzyme from Bacteria able to Destroy Penicillin. Nature 146, 837–837. https://doi.org/10.1038/146837a0

AbuSara, N.F., Piercey, B.M., Moore, M.A., Shaikh, A.A., Nothias, L.-F., Srivastava, S.K., Cruz-Morales, P., Dorrestein, P.C., Barona-Gómez, F., Tahlan, K., 2019. Comparative Genomics and Metabolomics Analyses of Clavulanic Acid-Producing Streptomyces Species Provides Insight Into Specialized Metabolism. Frontiers in Microbiology 10.

Ackermann, M., 2015. A functional perspective on phenotypic heterogeneity in microorganisms. Nat Rev Microbiol 13, 497–508. https://doi.org/10.1038/nrmicro3491

Ait, B.E., Parul, V., Lisa, S., Nathalie, G.-V., Cedric, J., Hans-Peter, K., Christophe, C., Yder, O., P, van W.G., 2015. Taxonomy, Physiology, and Natural Products of Actinobacteria. Microbiology and Molecular Biology Reviews 80, 1–43. https://doi.org/10.1128/MMBR.00019-15

Alam, K., Mazumder, A., Sikdar, S., Zhao, Y.-M., Hao, J., Song, C., Wang, Y., Sarkar, R., Islam, S., Zhang, Y., Li, A., 2022. Streptomyces: The biofactory of secondary metabolites. Front Microbiol 13, 968053. https://doi.org/10.3389/fmicb.2022.968053

Al-Bassam, M.M., Haist, J., Neumann, S.A., Lindenberg, S., Tschowri, N., 2018. Expression Patterns, Genomic Conservation and Input Into Developmental Regulation of the GGDEF/EAL/HD-GYP Domain Proteins in Streptomyces. Front. Microbiol. 9. https://doi.org/10.3389/fmicb.2018.02524

Algora, L., Herron, P., 2019. Genome dynamics of Streptomyces clavuligerus. Access Microbiology 1, 532. https://doi.org/s

Altschul, S.F., Gish, W., Miller, W., Myers, E.W., Lipman, D.J., 1990. Basic local alignment search tool. Journal of Molecular Biology 215, 403–410. https://doi.org/10.1016/S0022-2836(05)80360-2

Álvarez-Álvarez, R., Martínez-Burgo, Y., Rodríguez-García, A., Liras, P., 2017. Discovering the potential of S. clavuligerus for bioactive compound production: cross-talk between the chromosome and the pSCL4 megaplasmid. BMC Genomics 18, 907. https://doi.org/10.1186/s12864-017-4289-y

Amar, Y., Lagkouvardos, I., Silva, R.L., Ishola, O.A., Foesel, B.U., Kublik, S., Schöler, A., Niedermeier, S., Bleuel, R., Zink, A., Neuhaus, K., Schloter, M., Biedermann, T., Köberle, M., 2021. Pre-digest of unprotected DNA by Benzonase improves the representation of living skin bacteria and efficiently depletes host DNA. Microbiome 9, 123. https://doi.org/10.1186/s40168-021-01067-0

Bai, C., van Wezel, G.P., 2023. CUBIC: A Versatile Cumate-Based Inducible CRISPRi System in Streptomyces. ACS Synth Biol 12, 3143–3147. https://doi.org/10.1021/acssynbio.3c00464

Baichwal, V., Liu, D., Ames, G.F., 1993. The ATP-binding component of a prokaryotic traffic ATPase is exposed to the periplasmic (external) surface. Proc Natl Acad Sci U S A 90, 620–624. https://doi.org/10.1073/pnas.90.2.620

Baltz, R.H., 2011. Strain improvement in actinomycetes in the postgenomic era. Journal of Industrial Microbiology and Biotechnology 38, 657–666. https://doi.org/10.1007/s10295-010-0934-z

Bankevich, A., Nurk, S., Antipov, D., Gurevich, A.A., Dvorkin, M., Kulikov, A.S., Lesin, V.M., Nikolenko, S.I., Pham, S., Prjibelski, A.D., Pyshkin, A.V., Sirotkin, A.V., Vyahhi, N., Tesler, G., Alekseyev, M.A., Pevzner, P.A., 2012. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. J Comput Biol 19, 455–477. https://doi.org/10.1089/cmb.2012.0021

Baños, S., Pérez-Redondo, R., Koekman, B., Liras, P., 2009. Glycerol Utilization Gene Cluster in Streptomyces clavuligerus. Applied and Environmental Microbiology 75, 2991–2995. https://doi.org/10.1128/AEM.00181-09

Baran, A., Kwiatkowska, A., Potocki, L., 2023. Antibiotics and Bacterial Resistance-A Short Story of an Endless Arms Race. Int J Mol Sci 24, 5777. https://doi.org/10.3390/ijms24065777

Barka, E.A., Vatsa, P., Sanchez, L., Gaveau-Vaillant, N., Jacquard, C., Klenk, H.-P., Clément, C., Ouhdouch, Y., van Wezel, G.P., 2015. Taxonomy, Physiology, and Natural Products of Actinobacteria. Microbiology and Molecular Biology Reviews 80, 1–43. https://doi.org/10.1128/mmbr.00019-15

Barona-Gómez, F., Chevrette, M.G., Hoskisson, P.A., 2023. Chapter Six - On the evolution of natural product biosynthesis, in: Poole, R.K., Kelly, D.J. (Eds.), Advances in Microbial Physiology, Advances in Microbial Physiology. Academic Press, pp. 309–349. https://doi.org/10.1016/bs.ampbs.2023.05.001

Barrangou, R., Doudna, J.A., 2016. Applications of CRISPR technologies in research and beyond. Nat Biotechnol 34, 933–941. https://doi.org/10.1038/nbt.3659

Baskaran, B., Muthukumarasamy, A., 2017. Isolation, characterisation and enzymatic activity of Streptomyces sp. and its pH control during fermentation process. IET Syst Biol 11, 114–118. https://doi.org/10.1049/iet-syb.2016.0048

Batey, S.F.D., Greco, C., Hutchings, M.I., Wilkinson, B., 2020. Chemical warfare between fungus-growing ants and their pathogens. Current Opinion in Chemical Biology, Mechanistic Biology, Energy 59, 172–181. https://doi.org/10.1016/j.cbpa.2020.08.001

Belmar-Beiny, M.T., Thomas, C.R., 1991. Morphology and clavulanic acid production of Streptomyces clavuligerus: Effect of stirrer speed in batch fermentations. Biotechnol Bioeng 37, 456–462. https://doi.org/10.1002/bit.260370507

Bentley, S.D., Chater, K.F., Cerdeño-Tárraga, A.-M., Challis, G.L., Thomson, N.R., James, K.D., Harris, D.E., Quail, M.A., Kieser, H., Harper, D., Bateman, A., Brown, S., Chandra, G., Chen, C.W., Collins, M., Cronin, A., Fraser, A., Goble, A., Hidalgo, J., Hornsby, T., Howarth, S., Huang, C.-H., Kieser, T., Larke, L., Murphy, L., Oliver, K., O'Neil, S., Rabbinowitsch, E., Rajandream, M.-A., Rutherford, K., Rutter, S., Seeger, K., Saunders, D., Sharp, S., Squares, R., Squares, S., Taylor, K., Warren, T., Wietzorrek, A., Woodward, J., Barrell, B.G., Parkhill, J., Hopwood, D.A., 2002. Complete genome sequence of the model actinomycete Streptomyces coelicolor A3(2). Nature 417, 141–147. https://doi.org/10.1038/417141a

Berger, B., Waterman, M.S., Yu, Y.W., 2021. Levenshtein Distance, Sequence Comparison and Biological Database Search. IEEE Trans Inf Theory 67, 3287–3294. https://doi.org/10.1109/tit.2020.2996543

Bersanetti, P.A., Almeida, R.M.R.G., Barboza, M., Araújo, M.L.G.C., Hokka, C.O., 2005. Kinetic studies on clavulanic acid degradation. Biochemical Engineering Journal 23, 31–36. https://doi.org/10.1016/j.bej.2004.10.007

Bibb, M.J., 2005. Regulation of secondary metabolism in streptomycetes. Current Opinion in Microbiology 8, 208–215. https://doi.org/10.1016/j.mib.2005.02.016

Bierman, M., Logan, R., O'Brien, K., Seno, E.T., Nagaraja Rao, R., Schoner, B.E., 1992. Plasmid cloning vectors for the conjugal transfer of DNA from Escherichia coli to Streptomyces spp. Gene 116, 43–49. https://doi.org/10.1016/0378-1119(92)90627-2

Bignell, D.R.D., Kapil, T., Colvin, K., Jensen, S., Leskiw, B., 2005. Expression of ccaR, Encoding the Positive Activator of Cephamycin C and Clavulanic Acid Production in Streptomyces clavuligerus, Is Dependent on bldG. Antimicrobial Agents and Chemotherapy 49, 1529–1541. https://doi.org/10.1128/AAC.49.4.1529-1541.2005

Birch, A., Häusler, A., Rüttener, C., Hütter, R., 1991. Chromosomal deletion and rearrangement in Streptomyces glaucescens. J Bacteriol 173, 3531–3538. https://doi.org/10.1128/jb.173.11.3531-3538.1991

Bird, A.E., Bellis, J.M., Gasson, B.C., 1982. Spectrophotometric assay of clavulanic acid by reaction with imidazole. Analyst 107, 1241–1245. https://doi.org/10.1039/AN9820701241

Birke, A.S., 2020. Engineering Streptomyces clavuligerus for efficient growth on sustainable carbon sources [WWW Document].

Bizuye, A., Moges, F., Andualem, B., 2013. Isolation and screening of antibiotic producing actinomycetes from soils in Gondar town, North West Ethiopia. Asian Pac J Trop Dis 3, 375–381. https://doi.org/10.1016/S2222-1808(13)60087-0

Blattman, S.B., Jiang, W., Oikonomou, P., Tavazoie, S., 2020. Prokaryotic single-cell RNA sequencing by in situ combinatorial indexing. Nat Microbiol 5, 1192–1201. https://doi.org/10.1038/s41564-020-0729-6

Blighe, K., 2025. kevinblighe/EnhancedVolcano [WWW Document]. URL https://github.com/kevinblighe/EnhancedVolcano (accessed 9.6.25).

Blount, Z.D., Maddamsetti, R., Grant, N.A., Ahmed, S.T., Jagdish, T., Baxter, J.A., Sommerfeld, B.A., Tillman, A., Moore, J., Slonczewski, J.L., Barrick, J.E., Lenski, R.E., 2020. Genomic and phenotypic evolution of Escherichia coli in a novel citrate-only resource environment. eLife 9, e55414. https://doi.org/10.7554/eLife.55414

Bökel, C., 2008. EMS Screens, in: Dahmann, C. (Ed.), Drosophila: Methods and Protocols. Humana Press, Totowa, NJ, pp. 119–138. https://doi.org/10.1007/978-1-59745-583-1_7

Boruta, T., 2021. A bioprocess perspective on the production of secondary metabolites by Streptomyces in submerged co-cultures. World J Microbiol Biotechnol 37, 171. https://doi.org/10.1007/s11274-021-03141-z

Botas, A., Pérez-Redondo, R., Rodríguez-García, A., Álvarez-Álvarez, R., Yagüe, P., Manteca, A., Liras, P., 2018. ArgR of Streptomyces coelicolor Is a Pleiotropic Transcriptional Regulator: Effect on the Transcriptome, Antibiotic Production, and Differentiation in Liquid Cultures. Front Microbiol 9, 361. https://doi.org/10.3389/fmicb.2018.00361

Brethauer, S., Held, M., Panke, S., 2008. Clavulanic Acid Decomposition Is Catalyzed by the Compound Itself and by Its Decomposition Products. Journal of Pharmaceutical Sciences 97, 3451–3455. https://doi.org/10.1002/jps.21225

Bull, A.T., Asenjo, J.A., Goodfellow, M., Gómez-Silva, B., 2016. The Atacama Desert: Technical Resources and the Growing Importance of Novel Microbial Diversity. Annu Rev Microbiol 70, 215–234. https://doi.org/10.1146/annurev-micro-102215-095236

Butler, M.J., Bruheim, P., Jovetic, S., Marinelli, F., Postma, P.W., Bibb, M.J., 2002. Engineering of primary carbon metabolism for improved antibiotic production in Streptomyces lividans. Appl Environ Microbiol 68, 4731–4739. https://doi.org/10.1128/AEM.68.10.4731-4739.2002

Cerveny, L., Straskova, A., Dankova, V., Hartlova, A., Ceckova, M., Staud, F., Stulik, J., 2013. Tetratricopeptide Repeat Motifs in the World of Bacterial Pathogens: Role in Virulence Mechanisms. Infection and Immunity 81, 629–635. https://doi.org/10.1128/iai.01035-12

Chadda, A., Jensen, D., Tomko, E.J., Ruiz Manzano, A., Nguyen, B., Lohman, T.M., Galburt, E.A., 2022. Mycobacterium tuberculosis DNA repair helicase UvrD1 is activated by redox-dependent dimerization via a 2B domain cysteine. Proc Natl Acad Sci U S A 119, e2114501119. https://doi.org/10.1073/pnas.2114501119

Chater, K.F., 2016. Recent advances in understanding Streptomyces. F1000Res 5, 2795. https://doi.org/10.12688/f1000research.9534.1

Chater, K.F., 2001. Regulation of sporulation in Streptomyces coelicolor A3(2): a checkpoint multiplex? Current Opinion in Microbiology 4, 667–673. https://doi.org/10.1016/S1369-5274(01)00267-3

Chater, K.F., Chandra, G., 2008. The use of the rare UUA codon to define "Expression Space" for genes involved in secondary metabolism, development and environmental adaptation in Streptomyces. The Journal of Microbiology 46, 1–11. https://doi.org/10.1007/s12275-007-0233-1

Chevrette, M.G., Gutiérrez-García, K., Selem-Mojica, N., Aguilar-Martínez, C., Yañez-Olvera, A., Ramos-Aboites, H.E., Hoskisson, P.A., Barona-Gómez, F., 2020. Evolutionary dynamics of natural product biosynthesis in bacteria. Nat. Prod. Rep. 37, 566–599. https://doi.org/10.1039/C9NP00048H

Chin, C.-S., Alexander, D.H., Marks, P., Klammer, A.A., Drake, J., Heiner, C., Clum, A., Copeland, A., Huddleston, J., Eichler, E.E., Turner, S.W., Korlach, J., 2013. Nonhybrid, finished microbial genome assemblies from long-read SMRT sequencing data. Nat Methods 10, 563–569. https://doi.org/10.1038/nmeth.2474

Cho, H.S., Jo, J.C., Shin, C.-H., Lee, N., Choi, J.-S., Cho, B.-K., Roe, J.-H., Kim, C.-W., Kwon, H.J., Yoon, Y.J., 2019. Improved production of clavulanic acid by reverse engineering and overexpression of the regulatory genes in an industrial Streptomyces clavuligerus strain. Journal of Industrial Microbiology and Biotechnology 46, 1205–1215. https://doi.org/10.1007/s10295-019-02196-0

Choulet, F., Aigle, B., Gallois, A., Mangenot, S., Gerbaud, C., Truong, C., Francou, F.-X., Fourrier, C., Guérineau, M., Decaris, B., Barbe, V., Pernodet, J.-L., Leblond, P., 2006. Evolution of the Terminal Regions of the Streptomyces Linear Chromosome. Molecular Biology and Evolution 23, 2361–2369. https://doi.org/10.1093/molbev/msl108

Collis, A.J., Crowhurst, N., Honiker, K.J., Kendrew, S.G., 2021. Streptomyces clavuligerus. WO2021004912A1.

Crater, J.S., Lievense, J.C., 2018. Scale-up of industrial microbial processes. FEMS Microbiol Lett 365, fny138. https://doi.org/10.1093/femsle/fny138

Crits-Christoph, A., Diamond, S., Butterfield, C.N., Thomas, B.C., Banfield, J.F., 2018. Novel soil bacteria possess diverse genes for secondary metabolite biosynthesis. Nature 558, 440–444. https://doi.org/10.1038/s41586-018-0207-y

Cruz-Morales, P., Vijgenboom, E., Iruegas-Bocardo, F., Girard, G., Yáñez-Guerra, L.A., Ramos-Aboites, H.E., Pernodet, J.-L., Anné, J., van Wezel, G.P., Barona-Gómez, F., 2013. The Genome Sequence of Streptomyces lividans 66 Reveals a Novel tRNA-Dependent Peptide Biosynthetic System within a Metal-Related Genomic Island. Genome Biology and Evolution 5, 1165–1175. https://doi.org/10.1093/gbe/evt082

Dagva, O., Thibessard, A., Lorenzi, J.-N., Labat, V., Piotrowski, E., Rouhier, N., Myllykallio, H., Leblond, P., Bertrand, C., 2024. Correction of non-random mutational biases along a linear bacterial chromosome by the mismatch repair endonuclease NucS. Nucleic Acids Research 52, 5033–5047. https://doi.org/10.1093/nar/gkae132

Dalbey, R.E., Wang, P., van Dijl, J.M., 2012. Membrane Proteases in the Bacterial Protein Secretion and Quality Control Pathway. Microbiology and Molecular Biology Reviews 76, 311–330. https://doi.org/10.1128/mmbr.05019-11

Deatherage, D.E., Barrick, J.E., 2014. Identification of mutations in laboratory evolved microbes from next-generation sequencing data using breseq. Methods Mol Biol 1151, 165–188. https://doi.org/10.1007/978-1-4939-0554-6_12

Deindoerfer, F.H., Humphrey, A.E., 1959. Design of Multistage Systems for Simple Fermentation Processes. Ind. Eng. Chem. 51, 809–812. https://doi.org/10.1021/ie50595a023

Delneri, D., 2010. Barcode technology in yeast: application to pharmacogenomics. FEMS Yeast Research 10, 1083–1089. https://doi.org/10.1111/j.1567-1364.2010.00676.x

Delvigne, F., Goffin, P., 2014. Microbial heterogeneity affects bioprocess robustness: dynamic single-cell analysis contributes to understanding of microbial populations. Biotechnol J 9, 61–72. https://doi.org/10.1002/biot.201300119

Delvigne, F., Zune, Q., Lara, A.R., Al-Soud, W., Sørensen, S.J., 2014. Metabolic variability in bioprocessing: implications of microbial phenotypic heterogeneity. Trends in Biotechnology 32, 608–616. https://doi.org/10.1016/j.tibtech.2014.10.002

Den Hengst, C.D., Tran, N.T., Bibb, M.J., Chandra, G., Leskiw, B.K., Buttner, M.J., 2010. Genes essential for morphological development and antibiotic production in Streptomyces coelicolor are targets of BldD during vegetative growth. Molecular Microbiology 78, 361–379. https://doi.org/10.1111/j.1365-2958.2010.07338.x

Diesh, C., Stevens, G.J., Xie, P., De Jesus Martinez, T., Hershberg, E.A., Leung, A., Guo, E., Dider, S., Zhang, J., Bridge, C., Hogue, G., Duncan, A., Morgan, M., Flores, T., Bimber, B.N., Haw, R., Cain, S., Buels, R.M., Stein, L.D., Holmes, I.H., 2023. JBrowse 2: a modular genome browser with views of synteny and structural variation. Genome Biology 24, 74. https://doi.org/10.1186/s13059-023-02914-z

Dragosits, M., Mattanovich, D., 2013. Adaptive laboratory evolution – principles and applications for biotechnology. Microbial Cell Factories 12, 64. https://doi.org/10.1186/1475-2859-12-64

Duncan, K.R., Haltli, B., Gill, K.A., Correa, H., Berrué, F., Kerr, R.G., 2015. Exploring the diversity and metabolic potential of actinomycetes from temperate marine sediments from Newfoundland, Canada. Journal of Industrial Microbiology and Biotechnology 42, 57–72. https://doi.org/10.1007/s10295-014-1529-x

Ensign, J.C., 1978. Formation, properties, and germination of actinomycete spores. Annu Rev Microbiol 32, 185–219. https://doi.org/10.1146/annurev.mi.32.100178.001153

Espaillat, A., Alvarez, L., Torrens, G., ter Beek, J., Miguel-Ruano, V., Irazoki, O., Gago, F., Hermoso, J.A., Berntsson, R.P.-A., Cava, F., 2024. A distinctive family of L,D-transpeptidases catalyzing L-Ala-mDAP crosslinks in Alpha- and Betaproteobacteria. Nat Commun 15, 1343. https://doi.org/10.1038/s41467-024-45620-5

Favate, J.S., Skalenko, K.S., Chiles, E., Su, X., Yadavalli, S.S., Shah, P., 2023. Linking genotypic and phenotypic changes in the E. coli long-term evolution experiment using metabolomics. eLife 12, RP87039. https://doi.org/10.7554/eLife.87039

Feeney, M.A., Newitt, J.T., Addington, E., Algora-Gallardo, L., Allan, C., Balis, L., Birke, A.S., Castaño-Espriu, L., Charkoudian, L.K., Devine, R., Gayrard, D., Hamilton, J., Hennrich, O., Hoskisson, P.A., Keith-Baker, M., Klein, J.G., Kruasuwan, W., Mark, D.R., Mast, Y., McHugh, R.E., McLean, T.C., Mohit, E., Munnoch, J.T., Murray, J., Noble, K., Otani, H., Parra, J., Pereira, C.F., Perry, L., Pintor-Escobar, L., Pritchard, L., Prudence, S.M.M., Russell, A.H., Schniete, J.K., Seipke, R.F., Sélem-Mojica, N., Undabarrena, A., Vind, K., van Wezel, G.P., Wilkinson, B., Worsley, S.F., Duncan, K.R., Fernández-Martínez, L.T., Hutchings, M.I., 2022. ActinoBase: tools and protocols for researchers working on Streptomyces and other filamentous actinobacteria. Microbial Genomics 8, 000824. https://doi.org/10.1099/mgen.0.000824

Ferguson, N.L., Peña-Castillo, L., Moore, M.A., Bignell, D.R.D., Tahlan, K., 2016. Proteomics analysis of global regulatory cascades involved in clavulanic acid production and morphological development in Streptomyces clavuligerus. J Ind Microbiol Biotechnol 43, 537–555. https://doi.org/10.1007/s10295-016-1733-y

Ferrari, S., Beretta, S., Jacob, A., Cittaro, D., Albano, L., Merelli, I., Naldini, L., Genovese, P., 2021. BAR-Seq clonal tracking of gene-edited cells. Nature Protocols 16, 2991–3025. https://doi.org/10.1038/s41596-021-00529-x

Finkel, S.E., Kolter, R., 1999. Evolution of microbial diversity during prolonged starvation. Proceedings of the National Academy of Sciences 96, 4023–4027. https://doi.org/10.1073/pnas.96.7.4023

Fischer, M., Schmidt, C., Falke, D., Sawers, R.G., 2012. Terminal reduction reactions of nitrate and sulfate assimilation in *Streptomyces coelicolor* A3(2): identification of genes encoding nitrite and sulfite reductases. Research in Microbiology 163, 340–348. https://doi.org/10.1016/j.resmic.2012.05.004

Fleming, A., 1929. On the Antibacterial Action of Cultures of a Penicillium, with Special Reference to their Use in the Isolation of B. influenzæ. Br J Exp Pathol 10, 226–236.

Fleming, A., Wright, A.E., 1997. On a remarkable bacteriolytic element found in tissues and secretions. Proceedings of the Royal Society of London. Series B, Containing Papers of a Biological Character 93, 306–317. https://doi.org/10.1098/rspb.1922.0023

Flint, H.J., Scott, K.P., Duncan, S.H., Louis, P., Forano, E., 2012. Microbial degradation of complex carbohydrates in the gut. Gut Microbes 3, 289–306. https://doi.org/10.4161/gmic.19897

Gallagher, K.A., Schumacher, M.A., Bush, M.J., Bibb, M.J., Chandra, G., Holmes, N.A., Zeng, W., Henderson, M., Zhang, H., Findlay, K.C., Brennan, R.G., Buttner, M.J., 2020. c-di-GMP Arms an Anti-σ to Control Progression of Multicellular

Differentiation in Streptomyces. Mol Cell 77, 586-599.e6. https://doi.org/10.1016/j.molcel.2019.11.006
Gallagher, K.A., Tschowri, N., Brennan, R.G., Schumacher, M.A., Buttner, M.J., 2024. How c-di-GMP controls progression through the *Streptomyces* life cycle. Current Opinion in Microbiology 80, 102516. https://doi.org/10.1016/j.mib.2024.102516
Gallardo, 2020. Investigating the genome complexity of Streptomyces clavuligerus. University of Strathclyde.
Gao, C., Hindra, Mulder, D., Yin, C., Elliot, M.A., 2012. Crp Is a Global Regulator of Antibiotic Production in Streptomyces. mBio 3, 10.1128/mbio.00407-12. https://doi.org/10.1128/mbio.00407-12
Garcia-Ochoa, F., Gomez, E., 2009. Bioreactor scale-up and oxygen transfer rate in microbial processes: An overview. Biotechnology Advances 27, 153–176. https://doi.org/10.1016/j.biotechadv.2008.10.006
Gaurav, A., Bakht, P., Saini, M., Pandey, S., Pathania, R., 2023. Role of bacterial efflux pumps in antibiotic resistance, virulence, and strategies to discover novel efflux pump inhibitors. Microbiology (Reading) 169, 001333. https://doi.org/10.1099/mic.0.001333
Glauert, A.M., Hopwood, D.A., 1959. A Membranous Component of the Cytoplasm in Streptomyces coelicolor. The Journal of Biophysical and Biochemical Cytology 6, 515–516. https://doi.org/10.1083/jcb.6.3.515
Gohain, A., Manpoong, C., Saikia, R., De Mandal, S., 2020. Chapter 9 - Actinobacteria: diversity and biotechnological applications, in: De Mandal, S., Bhatt, P. (Eds.), Recent Advancements in Microbial Diversity. Academic Press, pp. 217–231. https://doi.org/10.1016/B978-0-12-821265-3.00009-8
Gohl, D.M., Vangay, P., Garbe, J., MacLean, A., Hauge, A., Becker, A., Gould, T.J., Clayton, J.B., Johnson, T.J., Hunter, R., Knights, D., Beckman, K.B., 2016. Systematic improvement of amplicon marker gene methods for increased accuracy in microbiome studies. Nat Biotechnol 34, 942–949. https://doi.org/10.1038/nbt.3601
Gomez, M., Doukhan, L., Nair, G., Smith, I., 1998. sigA is an essential gene in Mycobacterium smegmatis. Molecular Microbiology 29, 617–628. https://doi.org/10.1046/j.1365-2958.1998.00960.x
Gómez, R., Schnabel, I., Garrido, J., 1988. Pellet growth and citric acid yield of *Aspergillus niger* 110. Enzyme and Microbial Technology 10, 188–191. https://doi.org/10.1016/0141-0229(88)90086-5
Gomez-Escribano, J.P., Algora Gallardo, L., Bozhüyük, K.A.J., Kendrew, S.G., Huckle, B.D., Crowhurst, N.A., Bibb, M.J., Collis, A.J., Micklefield, J., Herron, P.R., Wilkinson, B., 2021. Genome editing reveals that pSCL4 is required for chromosome linearity in Streptomyces clavuligerus. Microbial Genomics 7, 000669. https://doi.org/10.1099/mgen.0.000669
Gómez-Ríos, D., Gómez-Gaona, L.M., Ramírez-Malule, H., 2024. Clavulanic Acid Overproduction: A Review of Environmental Conditions, Metabolic Fluxes, and Strain Engineering in Streptomyces clavuligerus. Fermentation 10, 526. https://doi.org/10.3390/fermentation10100526
Gómez-Ríos, D., Ramírez-Malule, H., Neubauer, P., Junne, S., Ríos-Estepa, R., Ochoa, S., 2022. Tuning of fed-batch cultivation of Streptomyces clavuligerus for enhanced Clavulanic Acid production based on genome-scale dynamic modeling. Biochemical Engineering Journal 108534. https://doi.org/10.1016/j.bej.2022.108534

Grant, S.G., Jessee, J., Bloom, F.R., Hanahan, D., 1990. Differential plasmid rescue from transgenic mouse DNAs into Escherichia coli methylation-restriction mutants. Proceedings of the National Academy of Sciences 87, 4645–4649. https://doi.org/10.1073/pnas.87.12.4645

Green, S., Eyres, G.T., Agyei, D., Kebede, B., 2024. Solid-state fermentation: Bioconversions and impacts on bioactive and nutritional compounds in oats. Compr Rev Food Sci Food Saf 23, e70070. https://doi.org/10.1111/1541-4337.70070

Gregory, M.A., Till, R., Smith, M.C.M., 2003. Integration site for Streptomyces phage ϕBT1 and development of site-specific integrating vectors. Journal of Bacteriology 185, 5320–5323. https://doi.org/10.1128/JB.185.17.5320-5323.2003

Gu, D., Tang, S., Liu, C., He, D., Tian, J., Yang, Y., 2024. Optimization of liquid fermentation conditions for Coprinus comatus to enhance antioxidant activity. Prep Biochem Biotechnol 54, 830–837. https://doi.org/10.1080/10826068.2023.2297703

Hackl, S., Bechthold, A., 2015. The Gene bldA, a Regulator of Morphological Differentiation and Antibiotic Production in Streptomyces. Archiv der Pharmazie 348, 455–462. https://doi.org/10.1002/ardp.201500073

Hagan, T., Ding, G., Buchmann, G., Oldroyd, B.P., Gloag, R., 2024. Serial founder effects slow range expansion in an invasive social insect. Nat Commun 15, 3608. https://doi.org/10.1038/s41467-024-47894-1

Haginaka, J., Nakagawa, T., Uno, T., 1981. Stability of Clavulanic Acid in Aqueous Solutions. Chemical & Pharmaceutical Bulletin 29, 3334–3341. https://doi.org/10.1248/cpb.29.3334

Harmand, N., Gallet, R., Martin, G., Lenormand, T., 2018. Evolution of bacteria specialization along an antibiotic dose gradient. Evol Lett 2, 221–232. https://doi.org/10.1002/evl3.52

Harms, H., König, G.M., Schäberle, T.F., 2017. Production of Antimicrobial Compounds by Fermentation, in: Sass, P. (Ed.), Antibiotics: Methods and Protocols. Springer, New York, NY, pp. 49–61. https://doi.org/10.1007/978-1-4939-6634-9_3

Hartman, Z., Henrikson, E.N., Hartman, P.E., Cebula, T.A., 1994. Molecular models that may account for nitrous acid mutagenesis in organisms containing double-stranded DNA. Environmental and Molecular Mutagenesis 24, 168–175. https://doi.org/10.1002/em.2850240305

Hempel, A.M., Wang, S.-B., Letek, M., Gil, J.A., Flärdh, K., 2008. Assemblies of DivIVA mark sites for hyphal branching and can establish new zones of cell wall growth in Streptomyces coelicolor. Journal of Bacteriology 190, 7579–7583. https://doi.org/10.1128/JB.00839-08

Hernández, I., Sant, C., Martínez, R., Fernández, C., 2020. Design of Bacterial Strain-Specific qPCR Assays Using NGS Data and Publicly Available Resources and Its Application to Track Biocontrol Strains. Front. Microbiol. 11. https://doi.org/10.3389/fmicb.2020.00208

Higgens, C.E., Kastner, R.E., 1971. Streptomyces clavuligerus sp. nov., a -Lactam Antibiotic Producer. International Journal of Systematic Bacteriology 21, 326–331. https://doi.org/10.1099/00207713-21-4-326

Hobbs, G., Frazer, C.M., Gardner, D.C.J., Cullum, J.A., Oliver, S.G., 1989. Dispersed growth of Streptomyces in liquid culture. Applied Microbiology and Biotechnology 31, 272–277. https://doi.org/10.1007/BF00258408

Hodgson, D.A., 2000. Primary metabolism and its control in streptomycetes: a most unusual group of bacteria. Adv Microb Physiol 42, 47–238. https://doi.org/10.1016/s0065-2911(00)42003-5

Hodgson, D.A., 1982. Glucose Repression of Carbon Source Uptake and Metabolism in Streptomyces coelicolor A3(2) and its Perturbation in Mutants Resistant to 2-Deoxyglucose. Microbiology 128, 2417–2430. https://doi.org/10.1099/00221287-128-10-2417

Holmes, N.A., Walshaw, J., Leggett, R.M., Thibessard, A., Dalton, K.A., Gillespie, M.D., Hemmings, A.M., Gust, B., Kelemen, G.H., 2013. Coiled-coil protein Scy is a key component of a multiprotein assembly controlling polarized growth in Streptomyces. Proceedings of the National Academy of Sciences of the United States of America 110. https://doi.org/10.1073/pnas.1210657110

Homerová, D., Ševčíková, B., Sprušanský, O., Kormanec, J., 2000. Identification of DNA-binding proteins involved in regulation of expression of the Streptomyces aureofaciens sigF gene, which encodes sporulation sigma factor σF. Microbiology 146, 2919–2928. https://doi.org/10.1099/00221287-146-11-2919

Hoskisson, P.A., Barona-Gómez, F., Rozen, D.E., 2024. Phenotypic heterogeneity in *Streptomyces* colonies. Current Opinion in Microbiology 78, 102448. https://doi.org/10.1016/j.mib.2024.102448

Hoskisson, P.A., Fernández-Martínez, L.T., 2018. Regulation of specialised metabolites in Actinobacteria – expanding the paradigms. Environmental Microbiology Reports 10, 231–238. https://doi.org/10.1111/1758-2229.12629

Hoskisson, P.A., Hobbs, G., 2005. Continuous culture--making a comeback? Microbiology (Reading) 151, 3153–3159. https://doi.org/10.1099/mic.0.27924-0

Hoskisson, P.A., Hobbs, G., Sharples, G.P., 2000. Response of Micromonospora echinospora (NCIMB 12744) spores to heat treatment with evidence of a heat activation phenomenon. Letters in Applied Microbiology 30, 114–117. https://doi.org/10.1046/j.1472-765x.2000.00680.x

Hoskisson, P.A., Seipke, R.F., 2020. Cryptic or Silent? The Known Unknowns, Unknown Knowns, and Unknown Unknowns of Secondary Metabolism. mBio 11, e02642-20. https://doi.org/10.1128/mBio.02642-20

Hwang, S., Lee, N., Choe, D., Lee, Y., Kim, W., Jeong, Y., Cho, S., Palsson, B.O., Cho, B.-K., 2021. Elucidating the Regulatory Elements for Transcription Termination and Posttranscriptional Processing in the Streptomyces clavuligerus Genome. mSystems 6, 10.1128/msystems.01013-20. https://doi.org/10.1128/msystems.01013-20

Hwang, S., Lee, N., Jeong, Y., Lee, Y., Kim, W., Cho, S., Palsson, B.O., Cho, B.-K., 2019. Primary transcriptome and translatome analysis determines transcriptional and translational regulatory elements encoded in the Streptomyces clavuligerus genome. Nucleic Acids Research 47, 6114–6129. https://doi.org/10.1093/nar/gkz471

Ikehata, H., Ono, T., 2011. The Mechanisms of UV Mutagenesis. Journal of Radiation Research 52, 115–125. https://doi.org/10.1269/jrr.10175

Ives, P.R., Bushell, M.E., 1997. Manipulation of the physiology of clavulanic acid production in Streptomyces clavuligerus. Microbiology (Reading) 143 (Pt 11), 3573–3579. https://doi.org/10.1099/00221287-143-11-3573

Jensen, S., Elder, K., Aidoo, K., Paradkar, A., 2000. Enzymes Catalyzing the Early Steps of Clavulanic Acid Biosynthesis Are Encoded by Two Sets of Paralogous Genes inStreptomyces clavuligerus. Antimicrobial Agents and Chemotherapy 44, 720–726. https://doi.org/10.1128/AAC.44.3.720-726.2000

Jeyachandran, S., Vibhute, P., Kumar, D., Ragavendran, C., 2023. Random mutagenesis as a tool for industrial strain improvement for enhanced production of antibiotics: a review. Mol Biol Rep 51, 19. https://doi.org/10.1007/s11033-023-08975-4

Jnawali, H.N., Lee, H.C., Sohng, J.K., 2010. Enhancement of clavulanic acid production by expressing regulatory genes in gap gene deletion mutant of Streptomyces clavuligerus NRRL3585. J Microbiol Biotechnol 20, 146–152.

Jones, S.E., Elliot, M.A., 2018. 'Exploring' the regulation of Streptomyces growth and development. Current Opinion in Microbiology 42, 25–30. https://doi.org/10.1016/j.mib.2017.09.009

Jones, S.E., Ho, L., Rees, C.A., Hill, J.E., Nodwell, J.R., Elliot, M.A., 2017. Streptomyces exploration is triggered by fungal interactions and volatile signals. eLife 6, e21738. https://doi.org/10.7554/eLife.21738

Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., Tunyasuvunakool, K., Bates, R., Žídek, A., Potapenko, A., Bridgland, A., Meyer, C., Kohl, S.A.A., Ballard, A.J., Cowie, A., Romera-Paredes, B., Nikolov, S., Jain, R., Adler, J., Back, T., Petersen, S., Reiman, D., Clancy, E., Zielinski, M., Steinegger, M., Pacholska, M., Berghammer, T., Bodenstein, S., Silver, D., Vinyals, O., Senior, A.W., Kavukcuoglu, K., Kohli, P., Hassabis, D., 2021. Highly accurate protein structure prediction with AlphaFold. Nature 596, 583–589. https://doi.org/10.1038/s41586-021-03819-2

Jyothikumar, V., Tilley, E.J., Wali, R., Herron, P.R., 2008. Time-Lapse Microscopy of Streptomyces coelicolor Growth and Sporulation. Applied and Environmental Microbiology 74, 6774–6781. https://doi.org/10.1128/AEM.01233-08

Katsuya, F., Sonchita, B., Stuart, C., Linda, S., Di, W., Jessica, B., Masood, K.-M., Klas, F., Nora, A., 2013. Dynamic gradients of an intermediate filament-like cytoskeleton are recruited by a polarity landmark during apical growth. Proceedings of the National Academy of Sciences 110, E1889–E1897. https://doi.org/10.1073/pnas.1305358110

Kiepas, A.B., Hoskisson, P.A., Pritchard, L., 2023. 16S rRNA phylogeny and clustering is not a reliable proxy for genome-based taxonomy in Streptomyces. 16S rRNA phylogeny and clustering is not a reliable proxy for genome-based taxonomy in Streptomyces 1–31. https://doi.org/10.1101/2023.08.15.553377

Kieser, T., Bibb, M.J., Buttner, M.J., Chater, K.F., Hopwood, D.A., 2000. Practical Streptomyces genetics. John Innes Foundation, Norwich.

Kim, I.-K., Lee, C.-J., Kim, M.-K., Kim, J.-M., Kim, J.-H., Yim, H.-S., Cha, S.-S., Kang, S.-O., 2006. Crystal structure of the DNA-binding domain of BldD, a central regulator of aerial mycelium formation in Streptomyces coelicolor A3(2). Molecular Microbiology 60, 1179–1193. https://doi.org/10.1111/j.1365-2958.2006.05176.x

Kirby, R., Wright, L.F., Hopwood, D.A., 1975. Plasmid-determined antibiotic synthesis and resistance in Streptomyces coelicolor. Nature 254, 265–267. https://doi.org/10.1038/254265a0

Knowles, J.R., 1985. Penicillin resistance: the chemistry of .beta.-lactamase inhibition. Accounts of Chemical Research 18, 97–104. https://doi.org/10.1021/ar00112a001

Kormanec, J., Rezuchova, B., Homerova, D., Csolleiova, D., Sevcikova, B., Novakova, R., Feckova, L., 2019. Recent achievements in the generation of stable genome alterations/mutations in species of the genus Streptomyces. Applied

Microbiology and Biotechnology 103, 5463–5482. https://doi.org/10.1007/s00253-019-09901-0
Krysenko, S., 2025. Current Approaches for Genetic Manipulation of Streptomyces spp.—Key Bacteria for Biotechnology and Environment. BioTech (Basel) 14, 3. https://doi.org/10.3390/biotech14010003
Kuhstoss, S., Richardson, M.A., Rao, R.N., 1991. Plasmid cloning vectors that integrate site-specifically in Streptomyces spp. Gene 97, 143–146. https://doi.org/10.1016/0378-1119(91)90022-4
Kumar, V., Ahluwalia, V., Saran, S., Kumar, J., Patel, A.K., Singhania, R.R., 2021. Recent developments on solid-state fermentation for production of microbial secondary metabolites: Challenges and solutions. Bioresource Technology 323, 124566. https://doi.org/10.1016/j.biortech.2020.124566
Kyung, Y.S., Hu, W.-S., Sherman, D.H., 2001. Analysis of temporal and spatial expression of the CcaR regulatory element in the cephamycin C biosynthetic pathway using green fluorescent protein. Molecular Microbiology 40, 530–541. https://doi.org/10.1046/j.1365-2958.2001.02386.x
Labeda, D.P., Goodfellow, M., Brown, R., Ward, A.C., Lanoot, B., Vanncanneyt, M., Swings, J., Kim, S.-B., Liu, Z., Chun, J., Tamura, T., Oguchi, A., Kikuchi, T., Kikuchi, H., Nishii, T., Tsuji, K., Yamaguchi, Y., Tase, A., Takahashi, M., Sakane, T., Suzuki, K.I., Hatano, K., 2012. Phylogenetic study of the species within the family Streptomycetaceae. Antonie Van Leeuwenhoek 101, 73–104. https://doi.org/10.1007/s10482-011-9656-0
Langmead, B., Salzberg, S.L., 2012. Fast gapped-read alignment with Bowtie 2. Nat Methods 9, 357–359. https://doi.org/10.1038/nmeth.1923
Larcombe, D.E., Bohovych, I.M., Pradhan, A., Ma, Q., Hickey, E., Leaves, I., Cameron, G., Avelar, G.M., Assis, L.J. de, Childers, D.S., Bain, J.M., Lagree, K., Mitchell, A.P., Netea, M.G., Erwig, L.P., Gow, N.A.R., Brown, A.J.P., 2023. Glucose-enhanced oxidative stress resistance—A protective anticipatory response that enhances the fitness of Candida albicans during systemic infection. PLOS Pathogens 19, e1011505. https://doi.org/10.1371/journal.ppat.1011505
Larcombe, D.E., Braes, R.E., Croxford, J.T., Wilson, J.W., Figurski, D.H., Hoskisson, P.A., 2024. Sequence and origin of the Streptomyces intergenetic-conjugation helper plasmid pUZ8002. Access Microbiol 6, 000808.v3. https://doi.org/10.1099/acmi.0.000808.v3
Latoscha, A., Drexler, D.J., Al-Bassam, M.M., Bandera, A.M., Kaever, V., Findlay, K.C., Witte, G., Tschowri, N., 2020. c-di-AMP hydrolysis by the phosphodiesterase AtaC promotes differentiation of multicellular bacteria. Proceedings of the National Academy of Sciences 117, 7392–7400. https://doi.org/10.1073/pnas.1917080117
Latoscha, A., Wörmann, M.E., Tschowri, N., 2019. Nucleotide second messengers in Streptomyces. Microbiology 165, 1153–1165. https://doi.org/10.1099/mic.0.000846
Latta, O., Bechthold, A., 2022. C-di-GMP and Its Role in Regulation of Natural Products Production. J Cell Signal Volume 3, 79–90. https://doi.org/10.33696/Signaling.3.069
Lawlor, E.J., Baylis, H.A., Chater, K.F., 1987. Pleiotropic morphological and antibiotic deficiencies result from mutations in a gene encoding a tRNA-like product in Streptomyces coelicolor A3(2). Genes & Development 1, 1305–1310. https://doi.org/10.1101/gad.1.10.1305

Lee, S.-H., Jang, Y.-B., Choi, Y., Lee, Y., Shin, B.N., Lee, H.-S., Lee, J.-S., Bahn, Y.-S., 2023. Adenylyl-Sulfate Kinase (Met14)-Dependent Cysteine and Methionine Biosynthesis Pathways Contribute Distinctively to Pathobiological Processes in Cryptococcus neoformans. Microbiol Spectr 11, e00685-23. https://doi.org/10.1128/spectrum.00685-23

Lemoine, A., Delvigne, F., Bockisch, A., Neubauer, P., Junne, S., 2017. Tools for the determination of population heterogeneity caused by inhomogeneous cultivation conditions. Journal of Biotechnology 251, 84–93. https://doi.org/10.1016/j.jbiotec.2017.03.020

Lenski, R.E., 2023. Revisiting the Design of the Long-Term Evolution Experiment with Escherichia coli. J Mol Evol 91, 241–253. https://doi.org/10.1007/s00239-023-10095-3

Lenski, R.E., 2017. Experimental evolution and the dynamics of adaptation and genome evolution in microbial populations. ISME J 11, 2181–2194. https://doi.org/10.1038/ismej.2017.69

Lenski, R.E., Mongold, J.A., Sniegowski, P.D., Travisano, M., Vasi, F., Gerrish, P.J., Schmidt, T.M., 1998. Evolution of competitive fitness in experimental populations of E. coli: what makes one genotype a better competitor than another? Antonie Van Leeuwenhoek 73, 35–47. https://doi.org/10.1023/a:1000675521611

Lenski, R.E., Rose, M.R., Simpson, S.C., Tadler, S.C., 1991. Long-Term Experimental Evolution in Escherichia coli. I. Adaptation and Divergence During 2,000 Generations. The American Naturalist 138, 1315–1341. https://doi.org/10.1086/285289

Lerminiaux, N.A., Cameron, A.D.S., 2018. Horizontal transfer of antibiotic resistance genes in clinical environments. Canadian Journal of Microbiology 65, 34–44. https://doi.org/10.1139/cjm-2018-0275

Lewis, K., 2013. Platforms for antibiotic discovery. Nat Rev Drug Discov 12, 371–387. https://doi.org/10.1038/nrd3975

Li, C., Teng, X., Peng, H., Yi, X., Zhuang, Y., Zhang, S., Xia, J., 2020. Novel scale-up strategy based on three-dimensional shear space for animal cell culture. Chemical Engineering Science 212, 115329. https://doi.org/10.1016/j.ces.2019.115329

Li, J., Jaitzig, J., Lu, P., Süssmuth, R.D., Neubauer, P., 2015. Scale-up bioprocess development for production of the antibiotic valinomycin in Escherichia coli based on consistent fed-batch cultivations. Microb Cell Fact 14, 83. https://doi.org/10.1186/s12934-015-0272-y

Li, R., Townsend, C.A., 2006. Rational strain improvement for enhanced clavulanic acid production by genetic engineering of the glycolytic pathway in Streptomyces clavuligerus. Metabolic Engineering 8, 240–252. https://doi.org/10.1016/j.ymben.2006.01.003

Li, W., Sharma, M., Kaur, P., 2014. The DrrAB Efflux System of Streptomyces peucetius Is a Multidrug Transporter of Broad Substrate Specificity. J Biol Chem 289, 12633–12646. https://doi.org/10.1074/jbc.M113.536136

Lilic, M., Holmes, N.A., Bush, M.J., Marti, A.K., Widdick, D.A., Findlay, K.C., Choi, Y.J., Froom, R., Koh, S., Buttner, M.J., Campbell, E.A., 2023. Structural basis of dual activation of cell division by the actinobacterial transcription factors WhiA and WhiB. Proceedings of the National Academy of Sciences 120, e2220785120. https://doi.org/10.1073/pnas.2220785120

Lind, P.A., Farr, A.D., Rainey, P.B., 2015. Experimental evolution reveals hidden diversity in evolutionary pathways. Elife 4, e07074. https://doi.org/10.7554/eLife.07074
Liras, P., Martín, J.F., 2021. Streptomyces clavuligerus: The Omics Era. Journal of industrial microbiology & biotechnology 48. https://doi.org/10.1093/jimb/kuab072
Liu, R., Deng, Z., Liu, T., 2018. *Streptomyces* species: Ideal chassis for natural product discovery and overproduction. Metabolic Engineering, Metabolic Engineering Host Organism Special Issue 50, 74–84. https://doi.org/10.1016/j.ymben.2018.05.015
Llarena, M., Llama, M.J., Serra, J.L., 2006. Purification and properties of NrtC and NrtD, the ATP-binding subunits of the ABC nitrate/nitrite transporter of Phormidium laminosum. Biochim Biophys Acta 1760, 1819–1826. https://doi.org/10.1016/j.bbagen.2006.08.006
Llor, C., Bjerrum, L., 2014. Antimicrobial resistance: risk associated with antibiotic overuse and initiatives to reduce the problem. Therapeutic Advances in Drug Safety 5, 229–241. https://doi.org/10.1177/2042098614554919
Locke, J.C.W., Elowitz, M.B., 2009. Using movies to analyse gene circuit dynamics in single cells. Nat Rev Microbiol 7, 383–392. https://doi.org/10.1038/nrmicro2056
López-Agudelo, V.A., Gómez-Ríos, D., Ramirez-Malule, H., 2021. Clavulanic Acid Production by Streptomyces clavuligerus: Insights from Systems Biology, Strain Engineering, and Downstream Processing. https://doi.org/10.3390/antibiotics
Lorenzi, J.N., Lespinet, O., Leblond, P., Thibessard, A., 2021. Subtelomeres are fast-evolving regions of the streptomyces linear chromosome. Microbial Genomics 7. https://doi.org/10.1099/MGEN.0.000525
Lorusso, A.B., Carrara, J.A., Barroso, C.D.N., Tuon, F.F., Faoro, H., 2022. Role of Efflux Pumps on Antimicrobial Resistance in Pseudomonas aeruginosa. Int J Mol Sci 23, 15779. https://doi.org/10.3390/ijms232415779
Love, M.I., Huber, W., Anders, S., 2014. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biology 15, 550. https://doi.org/10.1186/s13059-014-0550-8
Luo, J., Meng, Z., Xu, X., Wang, L., Zhao, K., Zhu, X., Qiao, Q., Ge, Y., Mao, L., Cui, L., 2022. Systematic benchmarking of nanopore Q20+ kit in SARS-CoV-2 whole genome sequencing. Front. Microbiol. 13. https://doi.org/10.3389/fmicb.2022.973367
Luo, Y., Melhem, S., Feelisch, M., Chatre, L., Morton, N.M., Dolga, A.M., van Goor, H., 2025. Thiosulphate sulfurtransferase: Biological roles and therapeutic potential. Redox Biol 82, 103595. https://doi.org/10.1016/j.redox.2025.103595
Ma, X., Ke, T., Mao, P., Jin, X., Ma, L., He, G., 2008. The mutagenic properties of BrdUTP in a random mutagenesis process. Mol Biol Rep 35, 663–667. https://doi.org/10.1007/s11033-007-9137-8
MacNeil, D.J., Gewain, K.M., Ruby, C.L., Dezeny, G., Gibbons, P.H., MacNeil, T., 1992. Analysis of Streptomyces avermitilis genes required for avermectin biosynthesis utilizing a novel integration vector. Gene 111, 61–68. https://doi.org/10.1016/0378-1119(92)90603-M
Maiti, P.K., Das, S., Sahoo, P., Mandal, S., 2020. Streptomyces sp SM01 isolated from Indian soil produces a novel antibiotic picolinamycin effective against multi drug resistant bacterial strains. Scientific Reports 10, 10092. https://doi.org/10.1038/s41598-020-66984-w

Manna, F., Martin, G., Lenormand, T., 2011. Fitness Landscapes: An Alternative Theory for the Dominance of Mutation. Genetics 189, 923–937. https://doi.org/10.1534/genetics.111.132944
Mannan, A.A., Darlington, A.P.S., Tanaka, R.J., Bates, D.G., 2025. Design principles for engineering bacteria to maximise chemical production from batch cultures. Nat Commun 16, 279. https://doi.org/10.1038/s41467-024-55347-y
Manteca, Á., Yagüe, P., 2018. Streptomyces differentiation in liquid cultures as a trigger of secondary metabolism. Antibiotics 7. https://doi.org/10.3390/antibiotics7020041
Mark, D.R., Tucker, N.P., Herron, P.R., 2024. Chromosome architecture as a determinant for biosynthetic diversity in Micromonospora. Microbial Genomics 10, 001313. https://doi.org/10.1099/mgen.0.001313
Martinez, J.L., Baquero, F., 2000. Mutation Frequencies and Antibiotic Resistance. Antimicrob Agents Chemother 44, 1771–1777. https://doi.org/10.1128/aac.44.7.1771-1777.2000
Mayfield, C.I., Williams, S.T., Ruddick, S.M., Hatfield, H.L., 1972. Studies on the ecology of actinomycetes in soil IV. Observations on the form and growth of streptomycetes in soil. Soil Biology and Biochemistry 4, 79–91. https://doi.org/10.1016/0038-0717(72)90045-4
Mazodier, P., Petter, R., Thompson, C., 1989. Intergeneric conjugation between Escherichia coli and Streptomyces species. Journal of Bacteriology 171, 3583–3585. https://doi.org/10.1128/jb.171.6.3583-3585.1989
McCandlish, D.M., 2011. VISUALIZING FITNESS LANDSCAPES. Evol 65, 1544–1558. https://doi.org/10.1111/j.1558-5646.2011.01236.x
McCormick, J.R., Flärdh, K., 2012. Signals and regulators that govern Streptomyces development. FEMS Microbiology Reviews 36, 206–231. https://doi.org/10.1111/j.1574-6976.2011.00317.x
McDonald, M.J., Gehrig, S.M., Meintjes, P.L., Zhang, X.-X., Rainey, P.B., 2009. Adaptive divergence in experimental populations of Pseudomonas fluorescens. IV. Genetic constraints guide evolutionary trajectories in a parallel adaptive radiation. Genetics 183, 1041–1053. https://doi.org/10.1534/genetics.109.107110
McNaughton, A.L., Roberts, H.E., Bonsall, D., de Cesare, M., Mokaya, J., Lumley, S.F., Golubchik, T., Piazza, P., Martin, J.B., de Lara, C., Brown, A., Ansari, M.A., Bowden, R., Barnes, E., Matthews, P.C., 2019. Illumina and Nanopore methods for whole genome sequencing of hepatitis B virus (HBV). Scientific Reports 9, 7081. https://doi.org/10.1038/s41598-019-43524-9
Medema, M.H., Trefzer, A., Kovalchuk, A., van den Berg, M., Müller, U., Heijne, W., Wu, L., Alam, M.T., Ronning, C.M., Nierman, W.C., Bovenberg, R.A.L., Breitling, R., Takano, E., 2010. The Sequence of a 1.8-Mb Bacterial Linear Plasmid Reveals a Rich Evolutionary Reservoir of Secondary Metabolic Pathways. Genome Biology and Evolution 2, 212–224. https://doi.org/10.1093/gbe/evq013
Min, J., Lee, C.W., Gu, M.B., 2003. Gamma-radiation dose-rate effects on DNA damage and toxicity in bacterial cells. Radiat Environ Biophys 42, 189–192. https://doi.org/10.1007/s00411-003-0205-8
Minges, A., Ciupka, D., Winkler, C., Höppner, A., Gohlke, H., Groth, G., 2017. Structural intermediates and directionality of the swiveling motion of Pyruvate Phosphate Dikinase. Sci Rep 7, 45389. https://doi.org/10.1038/srep45389
Mohit, E., 2023. Investigating streptomyces clavuligerus linear replicons for improved clavulanic acid production. University of Strathclyde.

Moore, J.M., Bradshaw, E., Seipke, R.F., Hutchings, M.I., McArthur, M., 2012. Use and discovery of chemical elicitors that stimulate biosynthetic gene clusters in Streptomyces bacteria. Methods Enzymol 517, 367–385. https://doi.org/10.1016/B978-0-12-404634-4.00018-8

Mu, X., Zhang, F., 2023. Diverse mechanisms of bioproduction heterogeneity in fermentation and their control strategies. J Ind Microbiol Biotechnol 50, kuad033. https://doi.org/10.1093/jimb/kuad033

Munita, J.M., Arias, C.A., 2016. Mechanisms of Antibiotic Resistance. Microbiology spectrum 4, 10.1128/microbiolspec.VMBF-0016–2015. https://doi.org/10.1128/microbiolspec.VMBF-0016-2015

Munnoch, J.T., Vahtokari, S., Kerr, L., Hoskisson, P.A., 2025. Adaptive radiation during long-term experimental evolution of the multicellular bacterium, Streptomyces. https://doi.org/10.1101/2025.08.20.671327

Murray, C.J.L., Ikuta, K.S., Sharara, F., Swetschinski, L., Aguilar, G.R., Gray, A., Han, C., Bisignano, C., Rao, P., Wool, E., Johnson, S.C., Browne, A.J., Chipeta, M.G., Fell, F., Hackett, S., Haines-Woodhouse, G., Hamadani, B.H.K., Kumaran, E.A.P., McManigal, B., Achalapong, S., Agarwal, R., Akech, S., Albertson, S., Amuasi, J., Andrews, J., Aravkin, A., Ashley, E., Babin, F.-X., Bailey, F., Baker, S., Basnyat, B., Bekker, A., Bender, R., Berkley, J.A., Bethou, A., Bielicki, J., Boonkasidecha, S., Bukosia, J., Carvalheiro, C., Castañeda-Orjuela, C., Chansamouth, V., Chaurasia, S., Chiurchiù, S., Chowdhury, F., Donatien, R.C., Cook, A.J., Cooper, B., Cressey, T.R., Criollo-Mora, E., Cunningham, M., Darboe, S., Day, N.P.J., Luca, M.D., Dokova, K., Dramowski, A., Dunachie, S.J., Bich, T.D., Eckmanns, T., Eibach, D., Emami, A., Feasey, N., Fisher-Pearson, N., Forrest, K., Garcia, C., Garrett, D., Gastmeier, P., Giref, A.Z., Greer, R.C., Gupta, V., Haller, S., Haselbeck, A., Hay, S.I., Holm, M., Hopkins, S., Hsia, Y., Iregbu, K.C., Jacobs, J., Jarovsky, D., Javanmardi, F., Jenney, A.W.J., Khorana, M., Khusuwan, S., Kissoon, N., Kobeissi, E., Kostyanev, T., Krapp, F., Krumkamp, R., Kumar, A., Kyu, H.H., Lim, C., Lim, K., Limmathurotsakul, D., Loftus, M.J., Lunn, M., Ma, J., Manoharan, A., Marks, F., May, J., Mayxay, M., Mturi, N., Munera-Huertas, T., Musicha, P., Musila, L.A., Mussi-Pinhata, M.M., Naidu, R.N., Nakamura, T., Nanavati, R., Nangia, S., Newton, P., Ngoun, C., Novotney, A., Nwakanma, D., Obiero, C.W., Ochoa, T.J., Olivas-Martinez, A., Olliaro, P., Ooko, E., Ortiz-Brizuela, E., Ounchanum, P., Pak, G.D., Paredes, J.L., Peleg, A.Y., Perrone, C., Phe, T., Phommasone, K., Plakkal, N., Ponce-de-Leon, A., Raad, M., Ramdin, T., Rattanavong, S., Riddell, A., Roberts, T., Robotham, J.V., Roca, A., Rosenthal, V.D., Rudd, K.E., Russell, N., Sader, H.S., Saengchan, W., Schnall, J., Scott, J.A.G., Seekaew, S., Sharland, M., Shivamallappa, M., Sifuentes-Osornio, J., Simpson, A.J., Steenkeste, N., Stewardson, A.J., Stoeva, T., Tasak, N., Thaiprakong, A., Thwaites, G., Tigoi, C., Turner, C., Turner, P., Doorn, H.R. van, Velaphi, S., Vongpradith, A., Vongsouvath, M., Vu, H., Walsh, T., Walson, J.L., Waner, S., Wangrangsimakul, T., Wannapinij, P., Wozniak, T., Sharma, T.E.M.W.Y., Yu, K.C., Zheng, P., Sartorius, B., Lopez, A.D., Stergachis, A., Moore, C., Dolecek, C., Naghavi, M., 2022. Global burden of bacterial antimicrobial resistance in 2019: a systematic analysis. The Lancet 399, 629–655. https://doi.org/10.1016/S0140-6736(21)02724-0

Nesbitt, N.M., Arora, D.P., Johnson, R.A., Boon, E.M., 2015. Modification of a bi-functional diguanylate cyclase-phosphodiesterase to efficiently produce cyclic diguanylate monophosphate. Biotechnol Rep (Amst) 7, 30–37. https://doi.org/10.1016/j.btre.2015.04.008

Netolitzky, D.J., Wu, X., Jensen, S.E., Roy, K.L., 1995. Giant linear plasmids of beta-lactam antibiotic producing Streptomyces. FEMS Microbiol Lett 131, 27–34. https://doi.org/10.1016/0378-1097(95)00230-3

Neubauer, P., Lin, H.Y., Mathiszik, B., 2003. Metabolic load of recombinant protein production: inhibition of cellular capacities for glucose uptake and respiration after induction of a heterologous gene in Escherichia coli. Biotechnol Bioeng 83, 53–64. https://doi.org/10.1002/bit.10645

Nielsen, J., 1997. Physiological Engineering Aspects of Penicillium Chrysogenum. World Scientific.

Nieminen, L., Webb, S., Smith, M.C.M., Hoskisson, P.A., 2013a. A Flexible Mathematical Model Platform for Studying Branching Networks: Experimentally Validated Using the Model Actinomycete, Streptomyces coelicolor. PLOS ONE 8, e54316-.

Nieminen, L., Webb, S., Smith, M.C.M., Hoskisson, P.A., 2013b. A Flexible Mathematical Model Platform for Studying Branching Networks: Experimentally Validated Using the Model Actinomycete, Streptomyces coelicolor. PLOS ONE 8, e54316. https://doi.org/10.1371/journal.pone.0054316

Nieselt, K., Battke, F., Herbig, A., Bruheim, P., Wentzel, A., Jakobsen, Ø.M., Sletta, H., Alam, M.T., Merlo, M.E., Moore, J., Omara, W.A., Morrissey, E.R., Juarez-Hermosillo, M.A., Rodríguez-García, A., Nentwich, M., Thomas, L., Iqbal, M., Legaie, R., Gaze, W.H., Challis, G.L., Jansen, R.C., Dijkhuizen, L., Rand, D.A., Wild, D.L., Bonin, M., Reuther, J., Wohlleben, W., Smith, M.C., Burroughs, N.J., Martín, J.F., Hodgson, D.A., Takano, E., Breitling, R., Ellingsen, T.E., Wellington, E.M., 2010. The dynamic architecture of the metabolic switch in Streptomyces coelicolor. BMC Genomics 11, 10. https://doi.org/10.1186/1471-2164-11-10

Obanye, A.I.C., Hobbs, G., Gardner, D.C.J., Oliver, S.G., 1996. Correlation between carbon flux through the pentose phosphate pathway and production of the antibiotic methylenomycin in Streptomyces coelicolor A3(2). Microbiology 142, 133–137. https://doi.org/10.1099/13500872-142-1-133

O'Keefe, 2016. 1250T Clavulanic Acid Expansion Project [WWW Document]. Gardiner & Theobald. URL https://www.gardiner.com/projects/1250t-clavulanic-acid-expansion-project (accessed 6.9.25).

Okonkowski, J., Kizer-Bentley, L., Listner, K., Robinson, D., Chartrain, M., 2005. Development of a robust, versatile, and scalable inoculum train for the production of a DNA vaccine. Biotechnol Prog 21, 1038–1047. https://doi.org/10.1021/bp040041p

O'Neill, J., 2016. The Review on Antimicrobial Resistance (AMR).

Pagès, J.-M., James, C.E., Winterhalter, M., 2008. The porin and the permeating antibiotic: a selective diffusion barrier in Gram-negative bacteria. Nat Rev Microbiol 6, 893–903. https://doi.org/10.1038/nrmicro1994

Papkou, A., Garcia-Pastor, L., Escudero, J.A., Wagner, A., 2023. A rugged yet easily navigable fitness landscape. Science 382, eadh3860. https://doi.org/10.1126/science.adh3860

Paradkar, A., 2013. Clavulanic acid production by Streptomyces clavuligerus: biogenesis, regulation and strain improvement. The Journal of Antibiotics 66, 411–420. https://doi.org/10.1038/ja.2013.26

Patro, R., Duggal, G., Love, M.I., Irizarry, R.A., Kingsford, C., 2017. Salmon provides fast and bias-aware quantification of transcript expression. Nat Methods 14, 417–419. https://doi.org/10.1038/nmeth.4197

Paul, G.C., Thomas, C.R., 1998. Characterisation of mycelial morphology using image analysis, in: Schügerl, K. (Ed.), Relation Between Morphology and Process Performances. Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 1–59. https://doi.org/10.1007/BFb0102278

Pawar, S.B., 2018. Computational fluid dynamics (CFD) analysis of airlift bioreactor: effect of draft tube configurations on hydrodynamics, cell suspension, and shear rate. Bioprocess Biosyst Eng 41, 31–45. https://doi.org/10.1007/s00449-017-1841-8

Payen, C., Sunshine, A.B., Ong, G.T., Pogachar, J.L., Zhao, W., Dunham, M.J., 2016. High-Throughput Identification of Adaptive Mutations in Experimentally Evolved Yeast Populations. PLOS Genetics 12, e1006339-.

Pérez-Llarena, F.J., Liras, P., Rodríguez-García, A., Martín, J.F., 1997. A regulatory gene (ccaR) required for cephamycin and clavulanic acid production in Streptomyces clavuligerus: amplification results in overproduction of both beta-lactam compounds. Journal of Bacteriology 179, 2053–2059. https://doi.org/10.1128/jb.179.6.2053-2059.1997

Perkins, 2025. When was beer invented? [WWW Document]. Live Science. URL https://www.livescience.com/archaeology/when-was-beer-invented (accessed 6.9.25).

Petković, H., Cullum, J., Hranueli, D., Hunter, I.S., Perić-Concha, N., Pigac, J., Thamchaipenet, A., Vujaklija, D., Long, P.F., 2006. Genetics of Streptomyces rimosus, the oxytetracycline producer. Microbiology and molecular biology reviews : MMBR 70, 704–728. https://doi.org/10.1128/MMBR.00004-06

Polz, M.F., Cavanaugh, C.M., 1998. Bias in Template-to-Product Ratios in Multitemplate PCR. Applied and Environmental Microbiology 64, 3724–3730. https://doi.org/10.1128/AEM.64.10.3724-3730.1998

Prestinaci, F., Pezzotti, P., Pantosti, A., 2015. Antimicrobial resistance: a global multifaceted phenomenon. Pathogens and Global Health 109, 309–318. https://doi.org/10.1179/2047773215Y.0000000030

Price, M.N., Wetmore, K.M., Waters, R.J., Callaghan, M., Ray, J., Liu, H., Kuehl, J.V., Melnyk, R.A., Lamson, J.S., Suh, Y., Carlson, H.K., Esquivel, Z., Sadeeshkumar, H., Chakraborty, R., Zane, G.M., Rubin, B.E., Wall, J.D., Visel, A., Bristow, J., Blow, M.J., Arkin, A.P., Deutschbauer, A.M., 2018. Mutant phenotypes for thousands of bacterial genes of unknown function. Nature 557, 503–509. https://doi.org/10.1038/s41586-018-0124-0

Qamer, S., Sandoe, J.A.T., Kerr, K.G., 2003. Use of Colony Morphology To Distinguish Different Enterococcal Strains and Species in Mixed Culture from Clinical Specimens. J Clin Microbiol 41, 2644–2646. https://doi.org/10.1128/JCM.41.6.2644-2646.2003

Quinlan, A.R., Hall, I.M., 2010. BEDTools: a flexible suite of utilities for comparing genomic features. Bioinformatics 26, 841–842. https://doi.org/10.1093/bioinformatics/btq033

Ramamurthy, T., Ghosh, A., Chowdhury, G., Mukhopadhyay, A.K., Dutta, S., Miyoshi, S., 2022. Deciphering the genetic network and programmed regulation of antimicrobial resistance in bacterial pathogens. Front Cell Infect Microbiol 12, 952491. https://doi.org/10.3389/fcimb.2022.952491

Ramirez-Malule, H., Junne, S., Nicolás Cruz-Bournazou, M., Neubauer, P., Ríos-Estepa, R., 2018. Streptomyces clavuligerus shows a strong association between TCA cycle intermediate accumulation and clavulanic acid biosynthesis. Applied

Microbiology and Biotechnology 102, 4009–4023. https://doi.org/10.1007/s00253-018-8841-8
Rand, J.M., Pisithkul, T., Clark, R.L., Thiede, J.M., Mehrer, C.R., Agnew, D.E., Campbell, C.E., Markley, A.L., Price, M.N., Ray, J., Wetmore, K.M., Suh, Y., Arkin, A.P., Deutschbauer, A.M., Amador-Noguez, D., Pfleger, B.F., 2017. A metabolic pathway for catabolizing levulinic acid in bacteria. Nature Microbiology 2, 1624–1634. https://doi.org/10.1038/s41564-017-0028-z
Rastädter, K., Wurm, D.J., Spadiut, O., Quehenberger, J., 2023. kLa based scale-up cultivation of the extremophilic archaeon Sulfolobus acidocaldarius: from benchtop to pilot scale. Front. Bioeng. Biotechnol. 11. https://doi.org/10.3389/fbioe.2023.1160012
Rateb, M.E., Hallyburton, I., Houssen, W.E., Bull, A.T., Goodfellow, M., Santhanam, R., Jaspars, M., Ebel, R., 2013. Induction of diverse secondary metabolites in Aspergillus fumigatus by microbial co-culture. RSC Adv. 3, 14444–14450. https://doi.org/10.1039/C3RA42378F
Rawson, T.M., Moore, L.S.P., Zhu, N., Ranganathan, N., Skolimowska, K., Gilchrist, M., Satta, G., Cooke, G., Holmes, A., 2020. Bacterial and Fungal Coinfection in Individuals With Coronavirus: A Rapid Review To Support COVID-19 Antimicrobial Prescribing. Clinical Infectious Diseases 71, 2459–2468. https://doi.org/10.1093/cid/ciaa530
Reading, C., Cole, M., 1977. Clavulanic acid: a beta-lactamase-inhiting beta-lactam from Streptomyces clavuligerus. Antimicrob Agents Chemother 11, 852–857. https://doi.org/10.1128/AAC.11.5.852
Reasoner, D.J., Geldreich, E.E., 1985. A new medium for the enumeration and subculture of bacteria from potable water. Appl Environ Microbiol 49, 1–7. https://doi.org/10.1128/aem.49.1.1-7.1985
Rees, D.C., Johnson, E., Lewinson, O., 2009. ABC transporters: the power to change. Nat Rev Mol Cell Biol 10, 218–227. https://doi.org/10.1038/nrm2646
Reisman, H.B., 1993. Problems in scale-up of biotechnology production processes. Crit Rev Biotechnol 13, 195–253. https://doi.org/10.3109/07388559309041319
Ribeck, N., Lenski, R.E., 2015. Modeling and quantifying frequency-dependent fitness in microbial populations with cross-feeding interactions. Evolution 69, 1313–1320. https://doi.org/10.1111/evo.12645
Ríos-Fernández, P., Caicedo-Montoya, C., Ríos-Estepa, R., 2024. Genomic Diversity of Streptomyces clavuligerus: Implications for Clavulanic Acid Biosynthesis and Industrial Hyperproduction. Int J Mol Sci 25, 10992. https://doi.org/10.3390/ijms252010992
Robinson, D.G., Chen, W., Storey, J.D., Gresham, D., 2014. Design and analysis of bar-seq experiments. G3: Genes, Genomes, Genetics 4, 11–18. https://doi.org/10.1534/g3.113.008565
Rodríguez-García, A., Barreiro, C., Santos-Beneit, F., Sola-Landa, A., Martín, J.F., 2007. Genome-wide transcriptomic and proteomic analysis of the primary response to phosphate limitation in Streptomyces coelicolor M145 and in a ΔphoP mutant. PROTEOMICS 7, 2410–2429. https://doi.org/10.1002/pmic.200600883
Romero-Rodríguez, A., Rocha, D., Ruiz-Villafan, B., Tierrafría, V., Rodríguez-Sanoja, R., Segura-González, D., Sánchez, S., 2016. Transcriptomic analysis of a classical model of carbon catabolite regulation in Streptomyces coelicolor. BMC Microbiol 16, 77. https://doi.org/10.1186/s12866-016-0690-y

Rugbjerg, P., Myling-Petersen, N., Porse, A., Sarup-Lytzen, K., Sommer, M.O.A., 2018. Diverse genetic error modes constrain large-scale bio-based production. Nature Communications 9. https://doi.org/10.1038/s41467-018-03232-w
Rugbjerg, P., Sommer, M.O.A., 2019. Overcoming genetic heterogeneity in industrial fermentations. Nature Biotechnology 37, 869–876. https://doi.org/10.1038/s41587-019-0171-6
Rutledge, P.J., Challis, G.L., 2015. Discovery of microbial natural products by activation of silent biosynthetic gene clusters. Nat Rev Microbiol 13, 509–523. https://doi.org/10.1038/nrmicro3496
Sabouraud, 1896. La question des teignes. Annales de Dermatologie 3rd series.
Saenz, H.L., Dehio, C., 2005. Signature-tagged mutagenesis: technical advances in a negative selection method for virulence gene identification. Current Opinion in Microbiology, Antimicrobials / Edited by Malcolm Page and Christopher T Walsh · Genomics / Edited by Stephan C Schuster and Gerhard Gottschalk 8, 612–619. https://doi.org/10.1016/j.mib.2005.08.013
Saloni, 2024. What was the Golden Age of Antibiotics, and how can we spark a new one?
Sambrook, J., Fritsch, E.F., Maniatis, T., 1989. Molecular Cloning: A Laboratory Manual. Cold Spring Harbor Laboratory.
Sandoval, J.M., Arenas, F.A., Vásquez, C.C., 2011. Glucose-6-Phosphate Dehydrogenase Protects Escherichia coli from Tellurite-Mediated Oxidative Stress. PLOS ONE 6, e25573. https://doi.org/10.1371/journal.pone.0025573
Santamarta, I., López-García, M.T., Kurt, A., Nárdiz, N., Álvarez-Álvarez, R., Pérez-Redondo, R., Martín, J.F., Liras, P., 2011. Characterization of DNA-binding sequences for CcaR in the cephamycin–clavulanic acid supercluster of Streptomyces clavuligerus. Molecular Microbiology 81, 968–981. https://doi.org/10.1111/j.1365-2958.2011.07743.x
Saudagar, P.S., Survase, S.A., Singhal, R.S., 2008. Clavulanic acid: A review. Biotechnology Advances 26, 335–351. https://doi.org/10.1016/j.biotechadv.2008.03.002
Schatz, A., Bugle, E., Waksman, S.A., 1944. Streptomycin, a Substance Exhibiting Antibiotic Activity Against Gram-Positive and Gram-Negative Bacteria.∗†. Proceedings of the Society for Experimental Biology and Medicine 55, 66–69. https://doi.org/10.3181/00379727-55-14461
Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T., Preibisch, S., Rueden, C., Saalfeld, S., Schmid, B., Tinevez, J.-Y., White, D.J., Hartenstein, V., Eliceiri, K., Tomancak, P., Cardona, A., 2012. Fiji: an open-source platform for biological-image analysis. Nat Methods 9, 676–682. https://doi.org/10.1038/nmeth.2019
Schmidt, F.R., 2005. Optimization and scale up of industrial fermentation processes. Appl Microbiol Biotechnol 68, 425–435. https://doi.org/10.1007/s00253-005-0003-0
Schneider, C.A., Rasband, W.S., Eliceiri, K.W., 2012. NIH Image to ImageJ: 25 years of image analysis. Nat Methods 9, 671–675. https://doi.org/10.1038/nmeth.2089
Schniete, J.K., Pablo, C.-M., Nelly, S.-M., Fernández-Martínez, L., Hunter, I.S., Francisco, B.-G., Hoskisson, P.A., Yup, L.S., 2018. Expanding Primary Metabolism Helps Generate the Metabolic Robustness To Facilitate Antibiotic Biosynthesis in Streptomyces. mBio 9, e02283-17. https://doi.org/10.1128/mBio.02283-17

Schniete, J.K., Reumerman, R., Kerr, L., Tucker, N.P., Hunter, I.S., Herron, P.R., Hoskisson, P.A., 2020. Differential transcription of expanded gene families in central carbon metabolism of Streptomyces coelicolor A3(2). Access Microbiol 2, acmi000122. https://doi.org/10.1099/acmi.0.000122

Schrempf, H., 1982. Plasmid loss and changes within the chromosomal DNA of Streptomyces reticuli. J Bacteriol 151, 701–707. https://doi.org/10.1128/jb.151.2.701-707.1982

Schumacher, M.A., Wörmann, M.E., Henderson, M., Salinas, R., Latoscha, A., Al-Bassam, M.M., Singh, K.S., Barclay, E., Gunka, K., Tschowri, N., 2022. Allosteric regulation of glycogen breakdown by the second messenger cyclic di-GMP. Nat Commun 13, 5834. https://doi.org/10.1038/s41467-022-33537-w

Seemann, T., 2014. Prokka: rapid prokaryotic genome annotation. Bioinformatics 30, 2068–2069. https://doi.org/10.1093/bioinformatics/btu153

Shin, C.-H., Cho, H.S., Won, H.-J., Kwon, H.J., Kim, C.-W., Yoon, Y.J., 2021. Enhanced production of clavulanic acid by improving glycerol utilization using reporter-guided mutagenesis of an industrial Streptomyces clavuligerus strain. Journal of Industrial Microbiology and Biotechnology 48, kuab004. https://doi.org/10.1093/jimb/kuab004

Shirling, E.B., Gottlieb, D., 1966. Methods for characterization of Streptomyces species1. International Journal of Systematic and Evolutionary Microbiology 16, 313–340. https://doi.org/10.1099/00207713-16-3-313

Shokralla, S., Gibson, J.F., Nikbakht, H., Janzen, D.H., Hallwachs, W., Hajibabaei, M., 2014. Next-generation DNA barcoding: using next-generation sequencing to enhance and accelerate DNA barcode capture from single specimens. Molecular Ecology Resources 14, 892–901. https://doi.org/10.1111/1755-0998.12236

Sievers, F., Higgins, D.G., 2014. Clustal Omega. Current Protocols in Bioinformatics 48, 3.13.1-3.13.16. https://doi.org/10.1002/0471250953.bi0313s48

Sigle, S., Ladwig, N., Wohlleben, W., Muth, G., 2015. Synthesis of the spore envelope in the developmental life cycle of Streptomyces coelicolor. International Journal of Medical Microbiology 305, 183–189. https://doi.org/10.1016/j.ijmm.2014.12.014

Silberstein, S., 1924. Zur Frage der salvarsanresistenten Lues. Arch. f. Dermat. 147, 116–130. https://doi.org/10.1007/BF01828193

Smith, M.A., Bertrand, C., Crosby, K., Eveleigh, E.S., Fernandez-Triana, J., Fisher, B.L., Gibbs, J., Hajibabaei, M., Hallwachs, W., Hind, K., Hrcek, J., Huang, D.-W., Janda, M., Janzen, D.H., Li, Y., Miller, S.E., Packer, L., Quicke, D., Ratnasingham, S., Rodriguez, J., Rougerie, R., Shaw, M.R., Sheffield, C., Stahlhut, J.K., Steinke, D., Whitfield, J., Wood, M., Zhou, X., 2012. Wolbachia and DNA Barcoding Insects: Patterns, Potential, and Problems. PLOS ONE 7, e36514-.

Som, N.F., Heine, D., Holmes, N.A., Munnoch, J.T., Chandra, G., Seipke, R.F., Hoskisson, P.A., Wilkinson, B., Hutchings, M.I., 2017. The Conserved Actinobacterial Two-Component System MtrAB Coordinates Chloramphenicol Production with Sporulation in Streptomyces venezuelae NRRL B-65442. Frontiers in Microbiology 8.

Song, J., Jordan, F., 2012. Inter-chain Acetyl Transfer in the E2 Component of Bacterial Pyruvate Dehydrogenase Suggests a Model with Different Roles for Each Chain in a Trimer of the Homooligomeric Component. Biochemistry 51, 2795–2803. https://doi.org/10.1021/bi201614n

Song, J.Y., Jensen, S.E., Lee, K.J., 2010a. Clavulanic acid biosynthesis and genetic manipulation for its overproduction. Applied Microbiology and Biotechnology 88, 659–669. https://doi.org/10.1007/s00253-010-2801-2
Song, J.Y., Jeong, H., Yu, D.S., Fischbach, M.A., Park, H.-S., Kim, J.J., Seo, J.-S., Jensen, S.E., Oh, T.K., Lee, K.J., Kim, J.F., 2010b. Draft genome sequence of Streptomyces clavuligerus NRRL 3585, a producer of diverse secondary metabolites. J Bacteriol 192, 6317–6318. https://doi.org/10.1128/JB.00859-10
Spaans, S.K., Weusthuis, R.A., Van Der Oost, J., Kengen, S.W.M., 2015. NADPH-generating systems in bacteria and archaea. Front. Microbiol. 6. https://doi.org/10.3389/fmicb.2015.00742
Spasic, J., Mandic, M., Djokic, L., Nikodinovic-Runic, J., 2018. Streptomyces spp. in the biocatalysis toolbox. Applied Microbiology and Biotechnology 102, 3513–3536. https://doi.org/10.1007/s00253-018-8884-x
Spero, M.A., Aylward, F.O., Currie, C.R., Donohue, T.J., 2015. Phylogenomic Analysis and Predicted Physiological Role of the Proton-Translocating NADH:Quinone Oxidoreductase (Complex I) Across Bacteria. mBio 6, 10.1128/mbio.00389-15. https://doi.org/10.1128/mbio.00389-15
Stefan, C.P., Hall, A.T., Graham, A.S., Minogue, T.D., 2022. Comparison of Illumina and Oxford Nanopore Sequencing Technologies for Pathogen Detection from Clinical Matrices Using Molecular Inversion Probes. The Journal of Molecular Diagnostics 24, 395–405. https://doi.org/10.1016/j.jmoldx.2021.12.005
Stekel, D., 2018. First report of antimicrobial resistance pre-dates penicillin. Nature 562, 192–192. https://doi.org/10.1038/d41586-018-06983-0
Stevens, B.M., Creed, T.B., Reardon, C.L., Manter, D.K., 2023. Comparison of Oxford Nanopore Technologies and Illumina MiSeq sequencing with mock communities and agricultural soil. Sci Rep 13, 9323. https://doi.org/10.1038/s41598-023-36101-8
Stincone, A., Prigione, A., Cramer, T., Wamelink, M.M.C., Campbell, K., Cheung, E., Olin-Sandoval, V., Grüning, N.-M., Krüger, A., Alam, M.T., Keller, M.A., Breitenbach, M., Brindle, K.M., Rabinowitz, J.D., Ralser, M., 2015. The return of metabolism: biochemistry and physiology of the pentose phosphate pathway. Biol Rev Camb Philos Soc 90, 927–963. https://doi.org/10.1111/brv.12140
Stroud, J.T., Ratcliff, W.C., 2025. Long-term studies provide unique insights into evolution. Nature 639, 589–601. https://doi.org/10.1038/s41586-025-08597-9
Tassoni, R., van der Aart, L.T., Ubbink, M., van Wezel, G.P., Pannu, N.S., 2017. Structural and functional characterization of the alanine racemase from *Streptomyces coelicolor* A3(2). Biochemical and Biophysical Research Communications 483, 122–128. https://doi.org/10.1016/j.bbrc.2016.12.183
Taurino, C., Frattini, L., Marcone, G.L., Gastaldo, L., Marinelli, F., 2011. Actinoplanes teichomyceticus ATCC 31121 as a cell factory for producing teicoplanin. Microb Cell Fact 10, 82. https://doi.org/10.1186/1475-2859-10-82
Thacker, T., 2024. India starts making key API clavulanic acid. The Economic Times.
Thompson, M.G., Valencia, L.E., Blake-Hedges, J.M., Cruz-Morales, P., Velasquez, A.E., Pearson, A.N., Sermeno, L.N., Sharpless, W.A., Benites, V.T., Chen, Y., Baidoo, E.E.K., Petzold, C.J., Deutschbauer, A.M., Keasling, J.D., 2019. Omics-driven identification and elimination of valerolactam catabolism in Pseudomonas putida KT2440 for increased product titer. Metabolic Engineering Communications 9, e00098. https://doi.org/10.1016/j.mec.2019.e00098

Thompson, W., Shah, S., Wordley, V., Edwards, D., 2022. Understanding the impact of COVID-19 on dental antibiotic prescribing across England: “it was a minefield.” Br Dent J 233, 653–658. https://doi.org/10.1038/s41415-022-5104-y

Tong, Y., Whitford, C.M., Blin, K., Jørgensen, T.S., Weber, T., Lee, S.Y., 2020. CRISPR–Cas9, CRISPRi and CRISPR-BEST-mediated genetic manipulation in streptomycetes. Nat Protoc 15, 2470–2502. https://doi.org/10.1038/s41596-020-0339-z

Tong, Y., Whitford, C.M., Robertsen, H.L., Blin, K., Jørgensen, T.S., Klitgaard, A.K., Gren, T., Jiang, X., Weber, T., Lee, S.Y., 2019. Highly efficient DSB-free base editing for streptomycetes with CRISPR-BEST. Proc Natl Acad Sci U S A 116, 20366–20375. https://doi.org/10.1073/pnas.1913493116

Torii, N., Nozaki, T., Masutani, M., Nakagama, H., Sugiyama, T., Saito, D., Asaka, M., Sugimura, T., Miki, K., 2003. Spontaneous mutations in the Helicobacter pylori rpsL gene. Mutat Res 535, 141–145. https://doi.org/10.1016/s1383-5718(02)00292-9

Toro, L., Pinilla, L., Avignone-Rossa, C., Ríos-Estepa, R., 2018. An enhanced genome-scale metabolic reconstruction of Streptomyces clavuligerus identifies novel strain improvement strategies. Bioprocess and Biosystems Engineering 41, 657–669. https://doi.org/10.1007/s00449-018-1900-9

Tschowri, N., Schumacher, M.A., Schlimpert, S., Chinnam, N.B., Findlay, K.C., Brennan, R.G., Buttner, M.J., 2014. Tetrameric c-di-GMP mediates effective transcription factor dimerization to control Streptomyces development. Cell 158, 1136–1147. https://doi.org/10.1016/j.cell.2014.07.022

van Bergeijk, D.A., Terlouw, B.R., Medema, M.H., van Wezel, G.P., 2020. Ecology and genomics of Actinobacteria: new concepts for natural product discovery. Nat Rev Microbiol 18, 546–558. https://doi.org/10.1038/s41579-020-0379-y

van Dissel, D., Claessen, D., Roth, M., van Wezel, G.P., 2015. A novel locus for mycelial aggregation forms a gateway to improved Streptomyces cell factories. Microb Cell Fact 14, 44. https://doi.org/10.1186/s12934-015-0224-6

van Dissel, D., Claessen, D., van Wezel, G.P., 2014. Chapter One - Morphogenesis of Streptomyces in Submerged Cultures, in: Sariaslani, S., Gadd, G.M. (Eds.), Advances in Applied Microbiology. Academic Press, pp. 1–45. https://doi.org/10.1016/B978-0-12-800259-9.00001-9

van Opijnen, T., Bodi, K.L., Camilli, A., 2009. Tn-seq: high-throughput parallel sequencing for fitness and genetic interaction studies in microorganisms. Nat Methods 6, 767–772. https://doi.org/10.1038/nmeth.1377

Veiter, L., Rajamanickam, V., Herwig, C., 2018. The filamentous fungal pellet—relationship between morphology and productivity. Appl Microbiol Biotechnol 102, 2997–3006. https://doi.org/10.1007/s00253-018-8818-7

Ventura, M., Canchaya, C., Tauch, A., Chandra, G., Fitzgerald, G.F., Chater, K.F., van Sinderen, D., 2007. Genomics of Actinobacteria: Tracing the Evolutionary History of an Ancient Phylum. Microbiology and Molecular Biology Reviews 71, 495–548. https://doi.org/10.1128/mmbr.00005-07

Viana Marques, D. de A., Machado, S.E.F., Carvalho Santos Ebinuma, V., Duarte, C. de A.L., Converti, A., Porto, A.L.F., 2018. Production of β-lactamase inhibitors by Streptomyces species. Antibiotics 7. https://doi.org/10.3390/antibiotics7030061

Waksman, S., 1958. My life with the microbe. Robert Hale.

Waksman, S.A., Henrici, A.T., 1943. The Nomenclature and Classification of the Actinomycetes. J Bacteriol 46, 337–341. https://doi.org/10.1128/jb.46.4.337-341.1943

Wang, Y., Yang, X., Yu, F., Deng, Z., Lin, S., Zheng, J., 2024. Structural and functional characterization of AfsR, an SARP family transcriptional activator of antibiotic biosynthesis in Streptomyces. PLOS Biology 22, e3002528. https://doi.org/10.1371/journal.pbio.3002528

Wang, Z., Zhang, M., Shi, X., Xiang, Q., 2017. Purification and Characterization of an ATPase GsiA from Salmonella enterica. Biomed Res Int 2017, 3076091. https://doi.org/10.1155/2017/3076091

Waschulin, V., Borsetto, C., James, R., Newsham, K.K., Donadio, S., Corre, C., Wellington, E., 2022. Biosynthetic potential of uncultured Antarctic soil bacteria revealed through long-read metagenomic sequencing. ISME J 16, 101–111. https://doi.org/10.1038/s41396-021-01052-3

Webb, J.P., Paiva, A.C., Rossoni, L., Alstrom-Moore, A., Springthorpe, V., Vaud, S., Yeh, V., Minde, D.-P., Langer, S., Walker, H., Hounslow, A., Nielsen, D.R., Larson, T., Lilley, K., Stephens, G., Thomas, G.H., Bonev, B.B., Kelly, D.J., Conradie, A., Green, J., 2022. Multi-omic based production strain improvement (MOB*psi*) for bio-manufacturing of toxic chemicals. Metabolic Engineering 72, 133–149. https://doi.org/10.1016/j.ymben.2022.03.004

Wehrs, M., Thompson, M.G., Banerjee, D., Prahl, J.-P., Morella, N.M., Barcelos, C.A., Moon, J., Costello, Z., Keasling, J.D., Shih, P.M., Tanjore, D., Mukhopadhyay, A., 2020. Investigation of Bar-seq as a method to study population dynamics of Saccharomyces cerevisiae deletion library during bioreactor cultivation. Microbial Cell Factories 19, 167. https://doi.org/10.1186/s12934-020-01423-z

Weisblum, B., 1995. Erythromycin resistance by ribosome modification. Antimicrob Agents Chemother 39, 577–585. https://doi.org/10.1128/AAC.39.3.577

Wentzel, A., Bruheim, P., Øverby, A., Jakobsen, Ø.M., Sletta, H., Omara, W.A.M., Hodgson, D.A., Ellingsen, T.E., 2012. Optimized submerged batch fermentation strategy for systems scale studies of metabolic switching in Streptomyces coelicolor A3(2). BMC Systems Biology 6, 59. https://doi.org/10.1186/1752-0509-6-59

Westhoff, S., Otto, S.B., Swinkels, A., Bode, B., van Wezel, G.P., Rozen, D.E., 2020. Spatial structure increases the benefits of antibiotic production in Streptomyces*. Evol 74, 179–187. https://doi.org/10.1111/evo.13817

Wetmore, K.M., Price, M.N., Waters, R.J., Lamson, J.S., He, J., Hoover, C.A., Blow, M.J., Bristow, J., Butland, G., Arkin, A.P., Deutschbauer, A., 2015. Rapid quantification of mutant fitness in diverse bacteria by sequencing randomly bar-coded transposons. mBio 6, e00306-00315. https://doi.org/10.1128/mBio.00306-15

Wietzorrek, A., Bibb, M., 1997. A novel family of proteins that regulates antibiotic production in streptomycetes appears to contain an OmpR-like DNA-binding fold. Molecular Microbiology 25, 1181–1184. https://doi.org/10.1046/j.1365-2958.1997.5421903.x

Williams, S.T., Shameemullah, M., Watson, E.T., Mayfield, C.I., 1972. Studies on the ecology of actinomycetes in soil—VI. The influence of moisture tension on growth and survival. Soil Biology and Biochemistry 4, 215–225. https://doi.org/10.1016/0038-0717(72)90014-4

World Health Organisation, 2024. 2023 Antibacterial agents in clinical and preclinical development.

World Health Organisation, 2021. Comprehensive Review of the WHO Global Action Plan on Antimicrobial Resistance.
Xia, J., Wang, G., Fan, M., Chen, M., Wang, Z., Zhuang, Y., 2021. Understanding the scale-up of fermentation processes from the viewpoint of the flow field in bioreactors and the physiological response of strains. Chinese Journal of Chemical Engineering 30, 178–184. https://doi.org/10.1016/j.cjche.2020.12.004
Xiao, Y., Bowen, C.H., Liu, D., Zhang, F., 2016. Exploiting nongenetic cell-to-cell variation for enhanced biosynthesis. Nature Chemical Biology 12, 339–344. https://doi.org/10.1038/nchembio.2046
Yagüe, P., Lopez-Garcia, M.T., Rioseras, B., Sanchez, J., Manteca, A., 2012. New insights on the development of Streptomyces and their relationships with secondary metabolite production. Curr Trends Microbiol 8, 65–73.
Yagüe, P., Rodríguez-García, A., López-García, M.T., Rioseras, B., Martín, J.F., Sánchez, J., Manteca, A., 2014. Transcriptomic Analysis of Liquid Non-Sporulating Streptomyces coelicolor Cultures Demonstrates the Existence of a Complex Differentiation Comparable to That Occurring in Solid Sporulating Cultures. PLOS ONE 9, e86296. https://doi.org/10.1371/journal.pone.0086296
Yanisch-Perron, C., Vieira, J., Messing, J., 1985. Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mpl8 and pUC19 vectors. Gene 33, 103–119. https://doi.org/10.1016/0378-1119(85)90120-9
Yin, P., Wang, Y.-H., Zhang, S.-L., Chu, J., Zhuang, Y.-P., Chen, N., Li, X.-F., Wu, Y.-B., 2008. Effect of mycelial morphology on bioreactor performance and avermectin production of *Streptomyces avermitilis* in submerged cultivations. Journal of the Chinese Institute of Chemical Engineers 39, 609–615. https://doi.org/10.1016/j.jcice.2008.04.008
Yoshikawa, S., Chikara, K.-I., Hashimoto, H., Mitsui, N., Shimosaka, M., Okazaki, M., 1995. Isolation and characterization of *Zygosaccharomyces rouxii* mutants defective in proton pumpout activity and salt tolerance. Journal of Fermentation and Bioengineering 79, 6–10. https://doi.org/10.1016/0922-338X(95)92735-U
Zhang, J.-W., Wang, R., Liang, X., Han, P., Zheng, Y.-L., Li, X.-F., Gao, D.-Z., Liu, M., Hou, L.-J., Dong, H.-P., 2023. Novel Gene Clusters for Natural Product Synthesis Are Abundant in the Mangrove Swamp Microbiome. Applied and Environmental Microbiology 89, e00102-23. https://doi.org/10.1128/aem.00102-23
Zhang, Z., Du, C., de Barsy, F., Liem, M., Liakopoulos, A., van Wezel, G.P., Choi, Y.H., Claessen, D., Rozen, D.E., 2020. Antibiotic production in Streptomyces is organized by a division of labor through terminal genomic differentiation. Science Advances 6, eaay5781. https://doi.org/10.1126/sciadv.aay5781
Zhang, Z., Shitut, S., Claushuis, B., Claessen, D., Rozen, D.E., 2022. Mutational meltdown of putative microbial altruists in Streptomyces coelicolor colonies. Nat Commun 13, 2266. https://doi.org/10.1038/s41467-022-29924-y
Zhou, Y., Zhou, X., Yu, D.-L., Sang, B., Feng, J.-T., Han, L.-R., Zhang, X., 2018. Optimization of Fermentation Conditions and Bench-Scale for Improvement of a Novel Glycoprotein GP-1 Production by Streptomyces kanasenisi ZX01. Molecules 23, 137. https://doi.org/10.3390/molecules23010137
Zimmermann, W., 1990. Degradation of lignin by bacteria. Journal of Biotechnology 13, 119–130. https://doi.org/10.1016/0168-1656(90)90098-V