School of Public Health and the School of Natural and Computing Sciences

Modelling the Composition and Structure of *Campylobacter jejuni* Biofilms

Paulina Agnieszka Dzianach

0000-0003-4202-9680

This thesis is presented for the Degree of Doctor of Philosophy (PhD) of Curtin University and the University of Aberdeen

August 2020

Declaration

I hereby declare that this thesis contains no material which has been accepted for the award of any other degree or diploma at any university or equivalent institution and that, to the best of my knowledge and belief, this thesis contains no material previously published or written by another person, except where due reference is made in the text of the thesis.

Publications included in this thesis:

Thesis chapter	**Publication title**	**Status**	**Nature and % of student contribution**	**Co-author names and % of co-author's contributions**
Chapter 2	The challenges of biofilm control and utilization – lessons from mathematical modelling	Published	90% of drafting and writing	1) Francisco Pérez-Reche, 5% of writing and review of the publication 2) Gary Dykes, 3% of writing and review of the publication 3) Norval Strachan, 1% of writing and review of the publication 4) Kenneth Forbes, 1% of writing and review of the publication

Contents

List of Figures

List of Tables

Glossary

ATCC	American type culture collection
CA	Cellular Automaton
CFU	Colony forming unit
DNAse	Deoxyribonuclease
ECM	Extracellular matrix
eDNA	Extracellular DNA
EPS	Exopolysaccharides
GWAS	Genome wide association study
IbM	Individual-based modelling
KMC	Kinetic Monte Carlo algorithm
MHB	Mueller Hinton Broth
NB2	Nutrient broth no. 2
OD	Optical density
PBS	Phosphate buffered saline
QS	Quorum sensing
RAxML	Randomized accelerated maximum likelihood
SEM	Scanning electron microscopy
SNP	Single nucleotide polymorphism
TSB	Triptone Soya Broth

Acknowledgements

I would like to thank my supervisors – Prof. Gary Dykes, Dr Francisco Pérez-Reche, Prof. Norval Strachan and Prof. Kenneth Forbes for their invaluable support, expertise, and for the time they have sacrificed to assist in the creation of this work. I would also like to thank Prof. Elizabeth Watkin for her help at the last stage of this project. My gratitude also extends to members of the research group. I thank all for their assistance and company, especially Amreeta Sarjit who helped me on countless occasions and has been a great friend to me from the minute I arrived on campus. I would also like to thank Dr Joshua Ravensdale for showing and explaining some microbiology techniques to me and for being a great lab leader. This work was supported by CIPRS and Research Stipend Scholarship, granted by the School of Natural and Computing Sciences at the University of Aberdeen and the Faculty of Health Sciences at Curtin University.

Abstract

The ability of bacteria to attach and grow on virtually any surface poses a tremendous challenge to industries and to human health. There are innumerable species of human pathogens capable of forming biofilms. One such pathogenic species is *Campylobacter jejuni* - a bacterium well known for causing foodborne illness around the globe. The goal of this research project was to study the effects of environmental and genetic factors on biofilm formation of *C. jejuni* with the use of mathematical modelling, experimental and bioinformatics techniques.

A novel mathematical model of biofilm formation was developed in order to reveal potential reasons about the mechanisms involved and their importance in the ability of *C. jejuni* to form biofilm communities under various environmental conditions. Through analysis of the output generated by our numerical simulations, we proposed arguments for some puzzling observations regarding *C. jejuni* biofilm formation which have been previously reported.

The effect of different media and atmospheric conditions on biofilm formation of *C. jejuni* ATCC33291 strain was also investigated experimentally. Growth was enhanced in microaerobic conditions compared to aerobic conditions in all media tested for this particular strain and media. Furthermore, our results suggest that it is more likely that composition of media along with the interactions between specific ingredients determine the extent of biofilm formation, rather than simply the level of nutrients, as has been previously postulated in literature.

A Genome wide association study (GWAS) was conducted to unveil potential genetic factors influencing biofilm formation ability of *C. jejuni* across its phylogeny. A number of genes and SNPs have been identified which may play a role in biofilm formation of *C. jejuni*. The results obtained may guide future investigations to understand the molecular mechanisms behind *C. jejuni* biofilm formation.

This study adds to the current knowledge of *C. jejuni* biofilm formation ability. The model presented in this work may be tuned further to include more processes or dimensions and its analysis may be extended to studies of how bacterial cell properties may affect biofilm formation. Our experimental results highlight the importance of interactions between particular ingredients of media on biofilm formation and challenge the previous idea that *C. jejuni* biofilm growth may be enhanced by low nutrient media, and thus may motivate further studies in this area. Finally, the GWAS study presented here suggests candidates for further analysis of the effect of gene manipulation on biofilm formation of *Campylobacter jejuni*. Recognition of genetic markers associated with biofilm forming ability of *C. jejuni* may

inform strategies designed to reduce the burden of this pathogen at food processing stages as well as propose targets for drug development to reduce its host colonisation capabilities.

Chapter 1 - Literature Review

1.1 Introduction

Aside from viruses, bacteria are the most numerous organisms on earth, and can survive in nearly all known habitats (1). One of the great challenges in the modern world is to minimize the spread of pathogenic bacteria to the human population. *Campylobacter jejuni* and *Campylobacter coli* are bacteria frequently found in the intestinal microbiota of farm birds and other domesticated animals such as pig, cattle or sheep, which are among most common causes of human food-borne disease (2, 3). While transmission to humans may occur from direct contact with infected animals or their faeces, or from ingestion of contaminated meat or dairy (4), vehicles of *C. jejuni* transmission among farm animals is also of importance. In particular, apart from contact with the animals around them, there has been evidence of *Campylobacter* transmission spreading among the farm animal population through ingestion of contaminated water (5), transmission by flying insects (6) or contaminated farm equipment (7).

In humans, it has been demonstrated that the infection with *Campylobacter jejuni* resulting in campylobacteriosis can occur after ingestion of approximately 500 organisms, with symptoms manifesting within 7 days of ingestion. Typically, the symptoms subside within a 7 day period without any medical treatment (8). Maintaining hydration and electrolyte balance is the recommended treatment for mild cases (9). Potential complications of the infection include Guillain-Barre syndrome, Miller Fisher syndrome or reactive arthritis (9). Compared to the mortality rate associated with infections from food-borne pathogens such as *Listeria monocytogenes* (~15%) or *Salmonella* (~0.15%) the mortality rate associated with *C. jejuni* infections is very low (~0.01%) (10). Although hospitalization rates relative to reported cases have been estimated to be quite high – around 14-34%, with rate of hospitalization highly dependent on the age of the infected individual (11), the true rate of hospitalization, relative to all human campylobacteriosis infections, is likely to be lower, as many mild cases are not reported – for example, an Infectious Intestinal Disease study of UK cases has reported that only about 1 in 10 cases of *Campylobacter* infections may be reported (12).

Although *C. jejuni* generally does not pose a serious threat to healthy individuals, it has been the predominant cause of food-borne disease in the EU since 2005. (10, 13). In 2018 alone, 246,571 cases of campylobacteriosis have been reported in Europe, compared to 91,857 salmonellosis cases (14). Australia has seen a considerable increase of campylobacteriosis cases over the last decade - 35,863 cases have been reported in 2019, compared to 16,106 cases in 2009 (15). The high incidence of *C. jejuni* infections place a substantial strain on the

economy and health of populations worldwide (4, 16, 17). The World Health Organisation (WHO) reported in 2018 *Campylobacter* to be in the top four causes of diarrheal disease, and furthermore, the most common bacterial cause of human gastroenteritis in the world (18). These facts illustrate that *C. jejuni* contamination problem not only has not been solved, but the urgency to address it is very high.

Bacterial infection through ingestion of contaminated food has been recognised to be a major cause of death and illness around the globe (19). The poultry industry is recognized as a significant risk for the spread of campylobacteriosis, specifically because *C. jejuni* easily spreads asymptomatically in chicken populations, causing subsequent contamination of water distributed to other farm animals and the contamination of meat intended for human consumption (7). Broiler chicken has been identified as the most common cause of human infection by *C. jejuni* (3, 10, 19). In an Australian study of retail meat samples collected weekly from 2016 to 2018, 85% of chicken samples tested positive for *Campylobacter* (20). In comparison, beef, lamb and pork retail samples tested positive for *Campylobacter* in 14%, 38% and 31% of samples, respectively (20). These data suggest a need for improvement in pathogen control at the different steps of the meat production process. One example of a solution which has been demonstrated to reduce the *C. jejuni* contamination of chickens is the introduction of fly screens in Danish broiler houses, which subsequently reduced the transfer of the pathogen within the flock (21). A follow-up study on the broiler houses on which the intervention was done has confirmed the long-term effectiveness of this method in reducing flock infection incidence (22). It should be noted, however, that the effectiveness of this strategy for a wider range of broiler houses is yet to be assessed (5).

Even though most *C. jejuni* strains are unable to grow outside of a host (23), they exhibit several survival mechanisms which allow *C. jejuni* to survive outside of its preferred environment (24). For example, it has been reported that *C. jejuni* can survive up to 7 months in water at low temperatures (24). Biofilm formation and attachment to surfaces or to existing biofilms of other species have been indicated as factors which may contribute to survival of *C. jejuni* outside of the host (25, 26), although the latter has been suggested more likely to contribute to *C. jejuni* prevalence in food-related environments, where conditions may be suboptimal for its growth (26). That being said, biofilm formation is still extensively studied for this species, as through these studies the physiology of *C. jejuni* as sessile cultures may be better understood, which may lead to improved control of this pathogen. One of important properties of biofilms is that they have been shown to increase the survival potential of bacteria exposed to environmental stress (27). It should also be recognised that *C. jejuni* may already be in a sessile state at the time of contamination of food surfaces, as this bacterium certainly does colonise guts of animals (28).

Due to its fastidious growth requirements and the difficulty in recovering it from the environment (29), the study of this organism *in situ* is relatively more challenging compared to organisms such as *Salmonella*, *E. coli* or *P. aeruginosa.* The use of e.g. mathematical and statistical models or analysis of whole genome reconstructions may aid research through reducing the amount of resources required to expand our knowledge on this pathogen. Specifically, this can be achieved by providing directions to experimental analysis or to provide focus on the key data which needs to be collected. Models, through simplifying reality, allow for some phenomena to be explained, which would have been much harder to establish by experiment and/or observation alone.

The models which have been applied in *C. jejuni* research may be grouped in terms of their resolution (Figure 1.1) as those which focus on the microorganisms themselves (i.e. biofilm formation models, predictive models, metabolic models), the individual host (animal models), or a population of hosts (epidemiological models). Biofilm formation models have been developed through extrapolating commonly observed patterns in order to provide a general framework through which *C. jejuni* biofilms can be more easily controlled. Namely, these models may inform the development of strategies inhibiting survival of *C. jejuni* colonies in biofilms (30). Predictive models have been employed in order to improve food safety at various stages of production by determining whether *C. jejuni* is capable of growth or survival in given conditions (defined by temperature, pH, flow, etc.), and thus whether *C. jejuni* prevalence may cause a food safety issue in these conditions (31). Metabolic modelling uncovers complex metabolic pathways and thus also cell-cell interactions (32). Metabolic models are recognised for their usefulness in biotechnology field and they are applied for design of new drugs and vaccines or for engineering of cells by changing their metabolism (33, 34). In the context of *C. jejuni*, these models have a potential to supplement microorganism level models (i.e. predictive and biofilm formation models) through their ability to predict cell physiology at the resolution of a single cell. Furthermore, metabolic models have a potential to help ease the disease burden caused by *C. jejuni* ingestion by their power to identify target proteins for drug or vaccine development (34), or by identifying factors which affect pathogen virulence (35). This in turn could aid *C. jejuni* research on an individual host level. In particular, identification of metabolic factors affecting virulence or ability to colonise the host may motivate further case studies in which animal models are employed – these are models in which animal subjects are used to study the disease *in vivo* (36). Finally, assessment of *C. jejuni* incidence and disease data at a host population level through the use of epidemiological models has the potential to identify most prominent sources of infection (3) , factors affecting the severity of illness (11) or risk of post infection complications (37), among many other useful information.

In the following sections, we present the types of models mentioned above which have been employed to improve our understanding of *C. jejuni*. The schematic diagram of the presented classes of models can be seen in Figure 1.1.

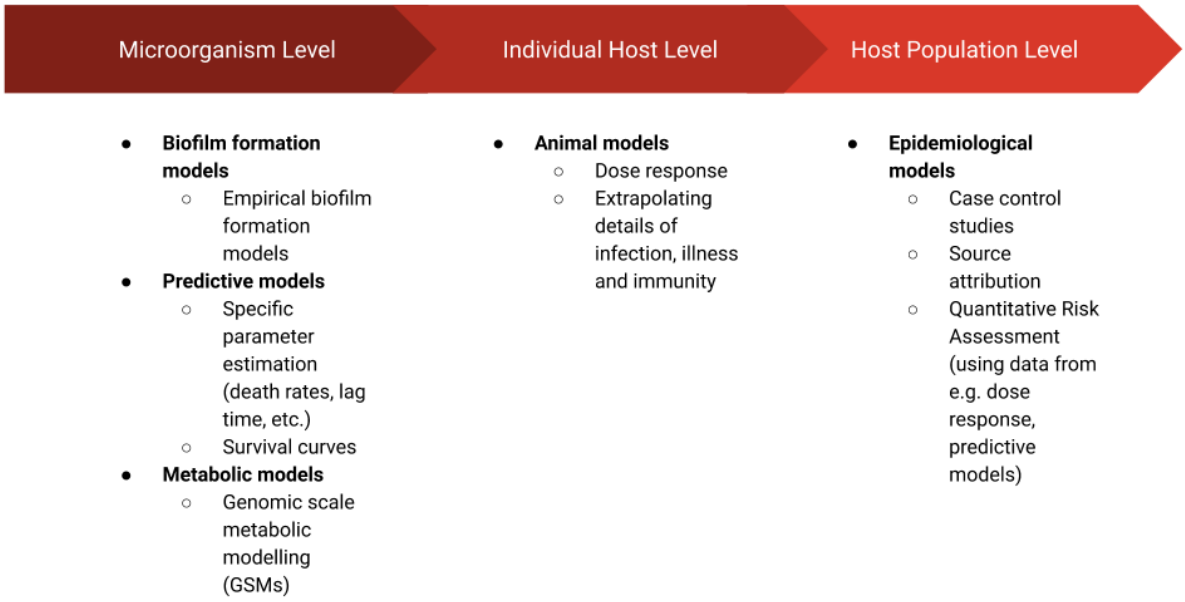

Figure 1.1 *Schematic diagram of the models discussed in this review, which have been employed to study Campylobacter jejuni species.*

1.2 Microorganism level models

1.2.1 Biofilm formation empirical models

The structure and composition of a mature biofilm forms a physical and chemical barrier which protects bacterial cells from harsh environmental conditions and antimicrobial agents. There is evidence suggesting an increased survival of *C. jejuni* biofilm cultures under adverse conditions as compared to the same type of cells in planktonic cultures (27, 30). For example, although most *C. jejuni* strains are not able to grow in aerobic conditions, it has been shown that in biofilms they survive significantly longer compared to planktonic cells under the same aerobic atmosphere (27). The increased length of survival of cells within the biofilm when exposed to atmospheric conditions may increase the chance of the bacteria being transferred to a more suitable environment in which it can grow, such as a living host. Furthermore, horizontal gene transfer, which may be particularly enhanced within biofilms due to proximity of individual cells, has been found to increase antimicrobial resistance of *C. jejuni* (38). This evidence suggests that the protective nature of biofilms may allow *C. jejuni* to colonize many different environments, and could explain why it is ubiquitous in the agricultural, food and medical sectors (27).

Biofilm formation by any bacterial species occurs in the following stages: surface attachment, microcolony formation, biofilm maturation and cell detachment and dispersal (39) (Figure 1.2). This is the simplest, general biofilm lifecycle description. The particular mechanisms which facilitate biofilm formation, however, vary between different species. For example, the composition of the extracellular matrix, mechanisms facilitating surface attachment, or responses to environmental factors, may differ (30). To create a more detailed model of biofilm formation, one must focus on the properties of the species of interest.

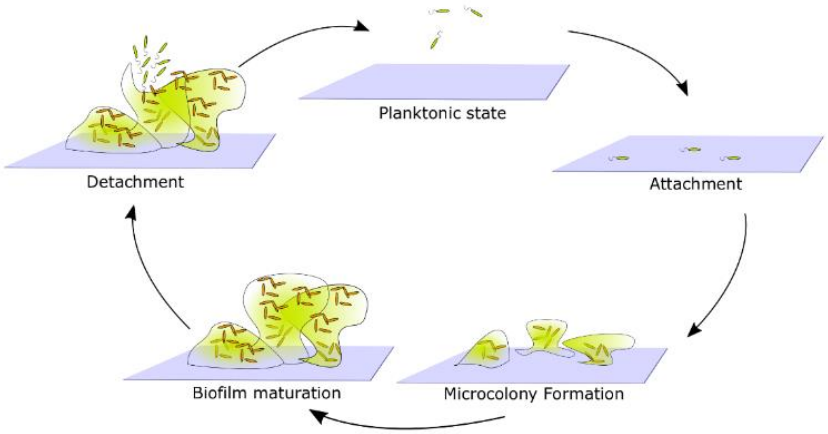

Figure 1.2 *General biofilm lifecycle. Cells in planktonic state migrate and attach to the surface. Attached cells form microcolonies by reproduction and generation of an extracellular products which together forms a biofilm matrix. Over time microcolonies begin to merge and a mature biofilm emerges. Eventually, some cells detach from the biofilm and return to the planktonic state.*

An extensive laboratory analysis which identified the pillars of *C. jejuni* biofilm formation under static conditions, resulted in the development of a general description of *C. jejuni* biofilms (30). For *C. jejuni*, adhesion is believed to be facilitated by flagella, since aflagellate mutant strains have an impaired ability to attach to surfaces (27, 30, 40, 41), unless the surface conditions are particularly favourable (25). The study of Svensson et al. (30) additionally revealed an association of *C. jejuni* biofilm maturation with bacterial lysis, which was later confirmed in another study (42). Confocal microscopy imaging of *C. jejuni* biofilms showed an abundancy of eDNA present in mature biofilms, which has also been confirmed in another study, where additionally lipids, proteins and polysaccharides were reported as other key constituents of the extracellular matrix (42). Biofilm formation was significantly reduced in the presence of DNAse I, however, no significant difference in surface attachment was observed, indicating that eDNA is not necessary for attachment of *C.*

jejuni to surfaces. On the other hand, in the conditions for which eDNA release and biofilm formation were enhanced (MHB with sodium deoxycholate), horizontal gene transfer, manifesting through recovery of colonies exhibiting combined antibiotic resistance of two parental strains which initiated the formation of the biofilm, was found to be increased. This property has also been later confirmed in another study (43). A replica of the model built on the collection of experimental evidence gathered by Svensson et al. (30) on *C. jejuni* biofilms can be found in Figure 1.3. In summary, the study concluded that biofilm formation of *C. jejuni* may be triggered by adverse environmental conditions and initial attachment is facilitated by flagella. Furthermore, as the biofilm matures, an abundance of eDNA is released, with evidence suggesting that this release is in significant part due to a lytic process. Finally, the study presented evidence of increased stress tolerance and horizontal gene exchange in well-formed biofilms which exhibited an abundancy of eDNA (30).

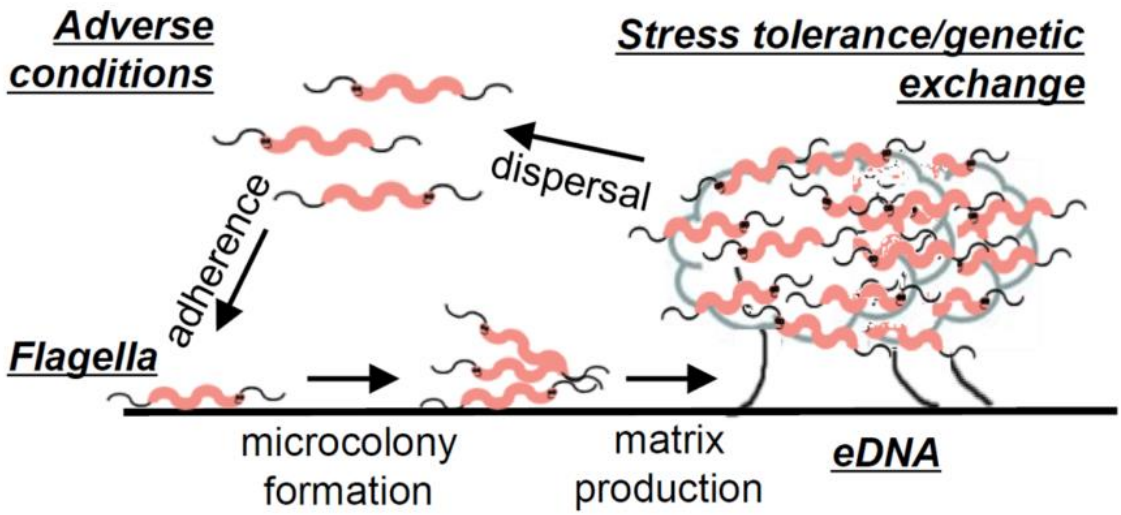

Figure 1.3 *Illustration of C. jejuni biofilm formation. Figure taken from "Flagella mediated adhesion and extracellular DNA release contribute to biofilm formation and stress tolerance of Campylobacter jejuni" by Svennson et al.*(30) *(reprinted under an open access license).*

C. jejuni has frequently been reported to take eDNA up in the horizontal gene transfer process. Its ability to successfully integrate eDNA into its existing genome has been suggested to account for the apparent genetic variation between *C. jejuni* strains (44). The process of horizontal gene transfer, specifically the binding of double-helix DNA strands onto the bacterial surface, followed by degradation of one of the strands into nucleotides and the integration of the other strand into the genome of the bacterial host, begs the question as to whether the nucleotides released could be utilized as a nutrient source for *C. jejuni*. The use of eDNA as a source or carbon, nitrogen and phosphate in nutrient limiting conditions has been confirmed for many other species of bacteria (45). In fact, a recent study on the

genes required for fitness of three *C. jejuni* strains, NCTC11168 and 81-176 among them, have identified genes for nucleotide metabolism and transport in their genome (46).

Apart from general properties of the structure, composition and physiology of *C. jejuni* biofilms, the relationship between *C. jejuni* biofilm formation and environmental conditions has also received a considerable amount of interest. Review of comparisons of biofilm formation when submerged in media and cultivated in aerobic conditions or microaerobic conditions in the available literature suggests that caution should be exercised when making postulates about the enhancement or inhibition of biofilm formation under aerobic conditions. In some studies, microaerobic conditions have been found to produce higher amounts of biofilms (40, 42), while in others, the opposite was the case (47). This may be partly attributed to specific properties of the media and particular strains, as a systematic study comparing biofilm formation in various media (MHB, Bolton and Brucella broths) and 8 strains revealed all possible types of effects of cultivation in aerobic conditions on biofilm formation (i.e. biofilm formation inhibited, enhanced or equivalent), which seemed to depend on the type of media used (48). Specifically, it has been suggested that presence of sodium bisulfite, which is an agent reducing levels of dissolved oxygen (DO) in Bolton and Brucella broths, may have played a role in biofilm formation being equivalent or higher in aerobic conditions compared to microaerobic conditions in these media, while in Mueller Hinton Broth, which lacks oxygen reducing ingredients, biofilm formation was either equivalent or lower in aerobic conditions (48). Apart from the effect of atmospheric oxygen on biofilm forming ability of *C. jejuni* there have also been some postulates made on the effect of nutrient levels on biofilm formation of *C. jejuni*. Namely, some studies suggested that lower nutrient media may promote biofilm formation of *C. jejuni* through comparisons of biofilm formation in nutrient low MHB with higher nutrient NB2, Bolton or Brucella Broths (40, 49). On the other hand, higher planktonic growth was obtained in a highly nutritious Triptone Soya broth with 0.6% yeast extract (TSBYE) compared to NB2 and MHB (50), begging to question as to how *C. jejuni* biofilms would perform when cultured in TSBYE, compared to MHB.

The qualitative biofilm formation models described in this section aided building a general picture of *Campylobacter jejuni* biofilms by finding common traits observed in biofilm assays of this species (e.g. eDNA as a major component of biofilms and lysis as an important process involved in biofilm formation, natural ability for horizontal gene transfer, flagella as an important structural component initiating surface attachment, etc.). These common traits might be used to inform the development of control measures for *Campylobacter jejuni* contamination through biofilms which would be potentially applicable for a wide range of *C. jejuni* strains. For example, since the results of biofilm matrix composition studies have

indicated that eDNA is a major component of *C. jejuni* biofilms, researchers have turned to studying the effect of *C. jejuni* biofilm treatment with a DNA disruptive enzyme (DNAse), which has shown to result in disruption of a number of *C. jejuni* strains (42, 43, 51). What could be of interest regarding this biofilm control strategy, is an assessment of how its effectiveness may vary between strains which normally produce DNAse and those which do not. It is our belief that apart from application of DNAse in biofilm control, many other research avenues may stem from a general description of *Campylobacter jejuni* biofilms. For instance, such descriptions may lead to development of mathematical models of *C. jejuni* biofilms, which may help to elucidate information which would be challenging to discover using experimental methods. Biofilm modelling using mathematical descriptions has already proven to be useful in answering particular questions related to areas in which biofilm formation is important, such as wastewater management, or the food and medical sectors (52). Such models commonly include computer simulations, which allow for testing hypotheses related to an occurrence of observed phenomena, or for prediction of biofilm formation for a wide range of possible scenarios (52), however, as far as we know, a specific mathematical model of *C. jejuni* biofilms has not been yet reported.

1.2.2 Predictive models for C. jejuni survival

Hazard analysis and critical control point (HACCP) principles are considered a cornerstone on which preventative strategies at all stages of food production are developed to ensure food safety (53). Predictive models are constituents in the process of following the HACCP principles through assessment of efficiency of interventions introduced at the slaughter and food processing stages, which aim to reduce pathogen incidence on food products (53). In particular, these models aim to assess how a variable (such as temperature, or a concentration of a biocide, for example) affects the observed reduction of the treated bacterial population.

In the case of *C. jejuni*, such predictive models generally focus on its elimination from chicken carcasses (53). These decontamination interventions can be grouped into three categories: physical interventions (hot water, steam irradiation, ultrasound, ultraviolet light, air chilling, freezing etc.), chemical interventions (organic acids, chorine, hypochlorite, sodium hypochlorite, sodium chlorite, chlorine dioxide, acidified sodium chlorite, monochloramine, cetylpyrydinium chloride, phosphate-based compounds, electrolyzed oxidizing water and ozonated water, etc.) and biological interventions (bacteriophages) (53).

A recent 2018 study which compared the effectiveness of several chemical interventions on reduction of *Campylobacter* and *Salmonella* incidence on chicken carcasses in a post chill decontamination tank, reported peracetic acid (PAA) and cetylpyridinium chloride (CPC) as

most effective methods compared to all other interventions considered, while chlorine and acidified sodium chlorite were found to be the least effective interventions (54). Although use of chemicals may be efficient in reducing microbial counts, there are concerns regarding the consumer and environmental safety of application of chemicals on food products. While in the USA many chemical decontamination methods are allowed, in the EU, only lactic acid up to 5% has been so far approved for use (55). Instead, physical treatments involving temperature or water are mostly applied in that area (56).

A study published this year proposed a promising method for treatment of carcasses with steam at 95 ºC and 120ºC for 3-5s (55). Although complete elimination of pathogens may be achieved with steam if applied for long enough, application for 10s has been previously shown to reduce the quality of meat. The method has been designed with consideration of the ease of incorporation of this intervention in the production process, effectiveness in microbial reduction and the effect of the treatment on the quality of meat.

Apart from the application of predictive models to assess microbial counts along with other indicators of meat quality for discrete values of a given variable (e.g. temperature or time of treatment), attempts have been made to translate empirical observations into theoretical predictions for a wider range of conditions.

One of the existing examples of such models relevant to *C. jejuni* is an empirical model built to predict the survival of *C. jejuni* as a function of temperature ranging from 4ºC to 30ºC (57). The authors used a simplified version of the Davey model (58) for the relationship between lag time, which is an initial exposure period at which the population size stays approximately constant, and temperature:

$$LT = A + \frac{B}{T} + \frac{C}{T^2} \tag{1.1}$$

Where *LT*=lag time, *T*=absolute temperature and *A*, *B*, and *C* are constants to be determined through fitting to experimental data.

For the relationship between the specific death rate of the organisms and temperature, the Boltzmann sigmoidal function was found to be a good fit to the obtained measurements:

$$SDR = SDR_{min} + \frac{SDR_{max} - SDR_{min}}{1 + exp[(T_{50} - T)/slope]} \tag{1.2}$$

Where SDR_{min} and SDR_{max} are minimum and maximum death rates, respectively, T_{50} is the temperature at which *SDR* is halfway between its minimum and maximum values, and *slope* is the rate of change of *SDR* as a function of temperature between its maximum values (57).

The results of the study suggested that the lag time decreases monotonically with temperature. Furthermore, specific death rate during the log phase was found to increase with temperature, and this increase occurred at a certain threshold (at the observed range the threshold appeared to occur between 16ºC and 20ºC). Finally, the study found that the maximum reduction of log CFU/ml of *C. jejuni* organisms on poultry patties or broth was not affected by temperature in the given range, i.e. 4-30°C (57).

While the above study focused on understanding the relationship between temperature and death rate, or temperature and lag time, other models may aim to quantify the relationship between microbial counts as a function of time, given a set of external conditions. With regards to these models, distinct survival curves as functions of time have been classified and assigned a suitable model distribution (59). One of the common distributions which has been found to provide a good fit for some of the types of survival curves is Weibull distribution (60, 61).

The Weibull distribution can be described as a probability distribution of individual failure events of individuals within a population (60). This distribution has many applications, from survival analysis of live organisms, to weather forecasting. It has been described in (60) as follows:

$$logN = logN_0 - \left(\frac{t}{\delta}\right)^p \quad (1.3)$$

Where N=number of live cells at time t, N_0= initial inoculum, p= parameter determining the shape of the distribution (concave or convex) and δ= time of the first 1-log reduction (simply because when t= δ, logN=logN_0-1).

A study which compared the goodness of fit of the Weibull and log-linear distributions as functions of *C. jejuni* survival and time in temperature range of -20ºC to 25ºC, also in poultry mince, found that the Weibull distribution provided a better fit for the observed data of log CFU/g across the whole range of temperatures considered, compared to the log-linear distribution, which matched the goodness of fit of the Weibull distribution only for non-zero temperatures (60).

A striking difference between the two studies described above (57, 60) was that although in both of these studies the same medium and temperatures were analysed, in the first one, a lag time of considerable length was observed (i.e. the population size remained approximately constant in the first days of incubation), while in the second study, the initial reduction in cell numbers was most abrupt and decreased as the time passed. What was consistent in both studies was the emergence of the subpopulation resistant to the conditions they were exposed to, which manifested through the levelling out of the death curve as the time progressed.

Reviewing survival curves obtained in other studies, in which different strains at various temperatures were used, revealed that whether or not a lag phase is observed in a given time frame may depend on both the strain tested and the incubation conditions (e.g. temperature, pH or atmosphere) (62–64). The two studies described above did perform their analysis on different strains of *C. jejuni*, which may explain the difference in the shapes of the survival curves against time (57, 60).

Apart from predictions of *C. jejuni* cell counts on food under various storage conditions, the data obtained from the analysis of survival of *C. jejuni* under various temperatures can be used in modelling the change in *C. jejuni* counts resulting from other food processing practices, e.g. scalding (31), which is a treatment of meat carcass with hot water or steam. A mathematical model, incorporating scalding process factors such as the volume of the scalding tank, average contamination of carcass prior to scalding, rate of carcasses entering the tank, or the rate of detachment of the bacteria from the carcass into the scalding water, plus the thermal inactivation data of *C. jejuni* strains subjected to scalding temperatures at varying pH values, could be a useful tool for food processing industries in the analysis of the effect of various factors on contamination of the final product (31). This particular model predicted that for a relevant range of model parameters, the level of *C. jejuni* contamination in the scalding water achieved a steady state in a short period of time, suggesting that the scalding process may be one of the sources of cross contamination in meat processing. The validation of the model against experimental data has been performed, confirming its predictive power (31).

To summarize, predictive models are a necessary tool in choosing the right set of decontamination methods at various food processing stages. Unfortunately, there seems to be no silver bullet solution which could lead to a complete eradication of *C. jejuni* contamination of food products. Rather, multifaceted approaches for pathogen control at every step of the production process (‘from farm to fork’), need to be further improved (65). It has been previously recommended that more standardized protocols should be developed for better comparability of results reporting on microbial reductions following a given intervention (53). Furthermore, it has been recommended based on current consumer trends and growing environmental concerns, that the assessment of natural disinfection methods (e.g. use of plant-based extracts) might be worthwhile, and some extracts from fruits and seeds have exhibited potential to reduce viability of *Campylobacter* on chicken samples without negatively affecting the sensory analysis of the meat (66). Such methods of chemical decontamination may aid physical decontamination methods with the additional benefit of being easier to accept by legislators and the consumers. Moreover, novel physical

disinfection methods, such as oscillating magnetic fields, use of enzymes, manothermosonication, pulsed electric fields, etc. may be of interest to consider (53).

1.2.3 Metabolic modelling and growth requirements

Genome-scale metabolic models (GSMs) aim to predict physiology and metabolism of organisms subjected to given environmental conditions. The development of a metabolic model typically follows four major steps, namely initial metabolic network reconstruction from gene annotation, refining the initial reconstruction with the use of other relevant data obtained from literature, conversion into a mathematical model, and validation of the reconstruction coupled with further refinement through comparison of the output of the model with reported phenomena (67). This type of model has been so far utilized more extensively for organisms such as *E. coli* (68) or *P. aeruginosa* (69). The first metabolic model of *C. jejuni* was proposed by Metris et al. (32). This model is based on genome sequence data obtained from the NCTC 11168 strain and relevant information found in literature on *C. jejuni*. Where information on *C. jejuni* was lacking, assumptions were made based on the data found for a closely related bacterium species, *Helicobacter pylori.* Information such as reactions for amino acids metabolism and nucleotide metabolism were drawn from the genome annotation. On the other hand, central metabolism reactions were mainly drawn from other literature sources (32). The model predicted, among other things, the predominance of essential genes associated with aromatic amino acid metabolism, tRNA metabolism and protein synthesis, TCA cycle, the cell envelope, and purine and pyrimidine metabolism (32).

The incorporation of metabolic reconstructions into mathematical models of bacterial populations have not been yet reported for *Campylobacter,* although it has been found to produce novel insights about colonies of other organisms such as *E. coli* (68). The lack of such models for *C. jejuni* may be due to this organism being understudied, compared to *E. coli*. To produce a model of such substantial detail (68) requires many organism-specific parameters to be derived from literature, for example the key metabolic requirements and products, with possible cross feeding mechanisms, or the rates of compound uptake and growth (68).

Similar to other chemoorganotrophic prokaryotes, *C. jejuni* uses organic compounds as energy sources. In particular, amino acids have been identified as primary substrate for *C. jejuni* (70). Although fucose, an abundant sugar in the mammalian gut (36), may sustain *C. jejuni* growth for some strains (71), *C. jejuni* is known to have limited capability of utilizing other carbohydrates as substrate (8, 29).

A recent genomic study of three strains of *C. jejuni* identified 486 genes which are essential for *C. jejuni* fitness (its survival and growth). Among these, genes responsible for the metabolism of lipids, coenzymes, carbohydrates, nucleotides and amino acids, were found (46). The appearance of nucleotides on this list may be particularly interesting when coupled with the findings presented in the previous sections, namely that eDNA forms a major component of *C. jejuni* biofilms (30). The presence of a nucleotide metabolism pathway suggests that it may be possible for *C. jejuni* to utilize the eDNA released by other cells as a nutrient source, and as there is an abundancy of eDNA in *C. jejuni* biofilms, this could potentially be an important factor in survival of *C. jejuni* populations in biofilms.

1.3 Individual host level models

1.3.1 Animal models of infection

Animal models can be used in order to identify virulence factors in *C. jejuni*, determine host responses to the presence of the pathogen, or test the viability of potential treatment methods (72). The use of non-human primate models, although on one hand desirable due to their closeness to humans, is limited, due to ethical considerations and the difficulty in keeping these animals, among other limitations (72).

Human volunteer studies have also been employed. In one such study directly related to *Campylobacter jejuni*, results suggested that the severity of acquired illness is strain dependent, the likelihood of exhibiting infection symptoms is dose dependent, and repeated exposure to a specific strain may increase immunity of the host (73). The latter finding agrees with the apparent decrease of colonization symptoms to *Campylobacter* exposure of people living in developing countries, compared to those living in industrialized countries (72). In a study using a ferret model, it was shown that NCTC 11168 has a low virulence compared with strain 81-176. Even at high doses, NCTC 11168 caused disease in only one out of nine animals, while all tested animals experienced infection symptoms after administering a high dose of 81-176. Furthermore, the study found a reduction in virulence of strains 81-176 with introduced mutations to their plasmid genes, suggesting that plasmids may be a significant factor in 81-176 virulence (74).

In order to produce an infection model which is relevant to human hosts, while maintaining ethical standards, antibiotic treatment of mice, used to eradicate their natural microflora, followed by introducing human microflora into their intestines, has been used (75). This microflora manipulation resulted in significant change in the outcome of *C. jejuni* colonization. Namely, mice with murine microflora were clear of the pathogen after 2 days of infection, while the mice with human microflora were found to be colonized for 6 weeks. The study concluded that specific gut microflora are essential in determining the outcome of

pathogen invasion, as the natural murine microflora exhibited resistance to *C. jejuni*, while the immune response of the mice with human microflora mimicked that of human campylobacteriosis (75). In another mice model study, which used antibiotic treatment prior to infection with *C. jejuni*, it was found that mice fed with zinc-deficient diet exhibited significantly more severe symptoms of campylobacteriosis than those on a standard or protein-deficient diet. Namely, the mice on the zinc-deficient diet suffered from bloody diahrrea and exhibited significantly increased weight loss due to the infection in comparison to mice on the other diets, for which only mild symptoms were observed (36).

In recent years, insect models, for example *Galleria mellonella* infection models, have been used to study various microorganism as an alternative to mammalian or avian models. Models of this type are desirable, due to, for example, reduced costs, improved commercial availability and lack of ethical approval required for use of these insects for research (76). Although insects lack an adaptive immune response, their innate immune response is very close to that of vertebrates (77). In contrast to mammalian or avian hosts, which are usually infected perorally (36, 73, 75), the insect larvae may easily be directly injected with a specific dose of the studied pathogen. As a result, more direct comparisons of virulence between strains may be derived (76). Typically, an intrahemocoelic injection method is used for inoculation of the larvae and it is recommended that 10-20 larvae are used for each tested condition for statistical significance (77). Markers of disease include melanisation, decrease in cocoon formation or motility, and death (77).

In one such study using a *G. mellonella* as a model organism for testing *C. jejuni* virulence, the effect of larvae infection with 67 *C. jejuni* isolates was tested (78). In congruence with common practice, a fixed inoculum size was directly injected into the haemocoel and the larvae incubated at 37ºC before assessment. One of the interesting observations was that *C. jejuni* cells recovered from infected larvae haemolymph were found to be in a coccoid, rather than the characteristic spiral shape. Furthermore, when infecting cultured mammalian and insect macrophages with *C. jejuni*, cell numbers were found to drop 100–1000-fold in comparison to the initial inoculum size in the first 4h post infection, and then remain constant or increase again when counted at the 24h mark (78). This finding suggests that *C. jejuni* cells experience stress at the initial stages of infection, but the population as a whole may be able to overcome it at later stages, provided that the initial inoculum is of sufficient size. Finally, from the comparison between larvae survival after a challenge with 6 different MLST types, it was suggested that the ST-21 group exhibits least virulence (with the mean survival rate at approx. 95%) and the highest virulence was observed for the group ST-257 (mean survival rate at approx. 76%).In contrast to the findings of the study outlined above, another *G. mellonella* study revealed a high virulence of a *C. jejuni* poultry isolate 13126,

which belongs to the ST-21 clonal complex (28). Although this particular isolate was not taken into account in the previously described study (78), this result calls for caution to be exercised before making inferences about differences between virulence of MLST groups.

Apart from generating particular data indicating the relative virulence of strains or properties of the host which may influence the severity of disease symptoms, important general theories have also been developed from this class of research. Data obtained from infection studies has led to the development of a Beta Poisson dose response model (79), with an aim to predict probability of infection or illness based on the administered dose. The Beta Poisson model has paved the way for future dose response models, and has found applications for a wide range of pathogens beyond *Campylobacter* species (80–83).

Although animal models have provided tremendous amount of information on many diseases, the variations between species has been reported to be a huge limitation, as the predictions of disease and effectiveness or side effects of tested treatments do not necessarily translate well from one species to another (84). *C. jejuni* is a good example of this, as it is believed that apart from some exceptions, *C. jejuni* does not generally cause illness in its other common hosts, while many cases of human disease caused by *C. jejuni* are reported each year (85). It has been suggested that modern technology may allow the shift from animal models to human-relevant data, by e.g. *in vitro* analysis of effect of disease on human tissues, or genomics approaches which may identify disease-specific genetic markers (84). It has been indicated that increased accessibility to human tissues of patients and healthy individuals for research purposes is essential to achieve statistically relevant results (84) In the case of *C. jejuni*, studies of human and poultry infection patterns may be of most interest.

1.4 Host population level models

1.4.1 Epidemiological models

Epidemiology is a branch of research dedicated to finding the causes, risk factors and transmission pathways associated with an illness, as well as predicting the impact of the disease on the populations, and developing suggestions for optimal control measures (86). Epidemiology studies rely on analysis of real-life data associated with a given disease (e.g. data collected from clinical records) (86). Epidemiology models are an important component of public health research, and as such, many such models have been developed for analysing data relevant to *C. jejuni* in an effort to minimise its burden on the populations worldwide.

Since epidemiology models rely heavily on data, statistical procedures are at the forefront of these types of studies, especially case control studies, which aim to identify and quantify risk factors associated with a disease. For example, case control studies may quantify the relationship between a dependent variable, such as disease incidence, and independent

variables, such as geographical location, age, gender, etc. Multivariate logistic regression models have been used in particular in studying these relationships (87, 88). It has been identified that contact with contaminated or undercooked retail chicken, international travel, eating in a restaurant, direct contact with animals and climate conditions are among significant risk factors of acquiring a *C. jejuni* infection (85). Furthermore, risk factors associated with susceptibility of individuals may be uncovered with case control studies, e.g. use of proton pump inhibitors has been associated with increased symptomatic *C. jejuni* infection rate (89). Other case control studies identified evidence for an increased risk of campylobacteriosis patients developing irritable bowel syndrome (IBS) (37), functional dyspepsia (FD) symptoms (90) or celiac disease (91). Extra gastrointestinal post infection complications associated with *C. jejuni* include Guillain-Barre syndrome, Miller-Fisher syndrome, bacteraemia, septicaemia, cardiovascular complications, meningitis, reactive arthritis and reproductive system failures (4).

Source attribution modelling also aids epidemiology studies by examining relative proportions of cases attributable to different sources (92). Multilocus sequence typing (MLST) allows to trace the phylogeny of isolates, which has contributed to finding that chicken is the most prominent source of human *C. jejuni* infections, followed by cattle and sheep (3). Furthermore, MLST has helped to classify *C. jejuni* isolates into distinct, highly diverse lineages, which aids in explaining the observed variation in *C. jejuni* phenotypes for different strains or strain variations. This categorisation of isolates in terms of their phylogeny is a key component in Genome Wide Association Studies (GWAS), where specific genetic factors associated with a given phenotype can be uncovered. For example, in 2017, a GWAS study found lineage specific genetic factors which may influence clinical incidence of *C. jejuni* (93). Interestingly, among the genes which were found to be associated with increased clinical incidence of *C. jejuni* in the ST-21 clonal complex, genes which have been identified to contribute to surface adhesion and biofilm formation (*kpsC* and *kpsD*) were identified in that study.

Lastly, quantitative risk assessment (QRA) methods, among other uses, may employ regression epidemiology models in order to identify acceptable thresholds for a value of a given risk factor (94). For *Campylobacter jejuni* for instance, quantitative risk assessment methods were applied to analyse the prevalence of this pathogen at various stages of food processing and thus pointed to specific areas in the process which may need improvement (95). With the use of QRA approach, it has also been recently suggested that a total eradication of *C. jejuni* on retail products may not be necessary, as only highly contaminated products pose a significant risk to consumers (85).

In 2008, the New Zealand Food Safety Authority (NZFS) produced a strategy aiming to reduce campylobacteriosis incidence due to poultry contamination, which incorporated improvements at all stages of the chicken's journey from farm to fork (i.e. primary production, processing, retail and consumer awareness) (92). An epidemiological study conducted in 2011, where, for example, source attribution techniques were employed, revealed that the incorporation of this multifaceted strategy resulted in an over 50% decline in campylobacteriosis cases in the country. The study considered that the decline could have been associated with other causes, such as changes in consumer behaviour or the climate. However, considering these factors, the authors suggested it was very unlikely that they could account for the whole extent of the decreased number of infections. For example, it was indeed found that poultry production was reduced by ~5% in the time period considered, and this could have partially contributed to the overall decline in *Campylobacter* infections, however, this decline was not sufficient to account for the full extent of the decline. Although it was stressed by the authors of the study that some contributing factors may have been overlooked, the success of the implemented strategy was concluded to be highly probable (92). This example illustrates how epidemiological studies may be used to evaluate efficiency of applied strategies in combating disease.

Although a number of transmission pathways of *C. jejuni* to humans have been described, it is believed that there are still some which are yet to be discovered. Apart from discovering all the ways humans may come into contact with *C. jejuni*, it has been suggested that focusing intervention strategies at the source (i.e. the farm) could subsequently lead to decrease in *C. jejuni* prevalence across both known and unknown pathways (85). One possible limitation of implementing this strategy is that due to *C. jejuni* being generally safe for livestock, farmers may lack the incentive to invest in measures designed to limit *C. jejuni* colonization in their flocks. Introduction of rewards (e.g. quality certifications) or policies may increase the incentive for farmers to implement more protective measures. A recent example of such an incentive is the ban of thinning procedures on RSPCA approved farms introduced in 2016, following a report released by the European Food Safety Authority which linked thinning to increased *C. jejuni* colonization among broiler chickens (56). Apart from introducing more control of *C. jejuni* on the farms, it has been suggested that finding a threshold for an acceptable level of meat contamination at the end of the processing stage and discarding or cooking highly contaminated samples may decrease the burden of *C. jejuni* on health of populations worldwide (85). Furthermore, close monitoring of new findings achieved by predictive models, which may indicate novel disinfection strategies, may also aid epidemiology studies by motivating assessment of these strategies on a larger scale,

which may in turn lead to policy changes and improvement in control of *C. jejuni* transmission.

1.5 Summary

We presented models which have been designed to understand and control *Campylobacter jejuni* and described recommendations for future research for all the types of models. Specifically, incorporation of mathematical modelling may aid understanding of *C. jejuni* biofilm formation both outside and inside the host. Predictive models may be improved by introduction of more standardized protocols for assessments of disinfection methods, and by assessment of novel physical disinfection strategies as well as assessment of the efficiency of plant extracts on *C. jejuni* eradication. Full description of metabolic pathways of *C. jejuni*, which is needed for successful application of metabolic models is yet to be achieved. A shift from animal models (except for those which are a source of human campylobacteriosis) to human-specific data may be made possible due to recent technology advancements, and this may lead to more accurate predictions of human infections. Epidemiology models may be aided by inclusion of clear instructions regarding the prescribed usage of statistical approaches in the documentation of generally used statistical software packages, as their misapplication has been reported to be of concern (96). Furthermore, monitoring advancements and potential strategies to motivate testing their efficiency on a larger scale through epidemiological studies may lead to improved control over *C. jejuni* globally.

In this literature review chapter, we tried to make it clear that a combination of different techniques and focus on various aspects, from a scale of the genome, through bacterial communities, up to affected host populations, are all important pieces of the health challenge puzzle posed by *C. jejuni.* Taken together, the proposed advancements could ultimately facilitate the reduction of *C. jejuni* burden on public health.

1.6 Research objectives and strategy

This study began with an extensive literature review, which was split into two topics –the current state of knowledge on *C. jejuni,* obtained by means of multiple modelling approaches (Chapter 1) and a review of mathematical modelling of biofilm formation (Chapter 2). The former was done to bring together what is known about *C. jejuni* and to identify knowledge gaps which are yet to be addressed. The latter was written to motivate the application of mathematical modelling techniques to *C. jejuni* biofilm formation research and to get familiarised with how mathematical modelling may be used to answer questions regarding biofilms.

In the rest of the study, we attempted to answer the following questions:

1) Does *C. jejuni* survive better in lower nutrient media compared to nutrient rich media?
2) How can *C. jejuni* form denser biofilms in aerobic conditions and harbour equivalent number of live cells in spite of being a microaerophile whose lysis is induced by aerobic stress?
3) How is it that DNAse have been shown to disrupt biofilm formation of some *C. jejuni* strains but there are many strains which possess genes encoding for DNAse in their genome and are good biofilm formers?
4) What are the genetic and environmental factors which influence biofilm formation ability of *C. jejuni?*

In order to answer these questions, three studies were designed and conducted. The first one attempted to answer some of the questions stated above with a novel mathematical model of biofilm formation, built on frameworks provided by previous mathematical models and adjusted to mimic certain patterns which have been observed for *C. jejuni* biofilms (Chapter 3). The second study aimed to analyse how environmental factors (i.e. oxygen conditions, level of nutrients and type of nutrients) influence biofilm formation of *C. jejuni* ATCC33291 in an experimental setup (Chapter 4). Finally, the objective of the third study was to uncover potential genetic markers which may play a role in the variation between *C. jejuni* strains regarding their biofilm forming ability (Chapter 5).

Chapter 2 - Challenges of Biofilm Control and Utilization: Lessons from Mathematical Modelling

The work presented in this chapter includes the following peer reviewed publication:

Dzianach, P.A., Dykes, G.A., Strachan, N.J., Forbes, K.J. and Pérez-Reche, F.J., 2019. Challenges of biofilm control and utilization: lessons from mathematical modelling. Journal of the Royal Society Interface, 16(155), p.20190042.

2.1 Introduction

It is estimated that bacteria and archaea constitute approximately half of all existing life on our planet (97). It should thereby not come as a surprise that microbes have such a profound impact on our environment and our day to day lives. It is evident that the control and utilization of these tiny, ubiquitous organisms can generate huge leaps to advance human society, be it through introducing improvements in environmental protection (98), general health and well-being (99) or in various industries, e.g. food (100), energy (101), water treatment (102), or mining (103). The immense complexity and diversity of the microbial world, and its sensitivity to environmental influences, physical or chemical alike, calls for a joining of forces between various science disciplines (for example biology, physics, mathematics, engineering, or chemistry), to fully equip the research field with the necessary tools for solving the associated challenges (104–106).

Bacteria may either exist in a "free-floating" planktonic state, or attached to a surface, forming biofilm communities (107). There are substantial differences between these two modes of bacterial existence, chemical gradients and stress responses being only the tip of the iceberg (108). In this chapter we will focus on the latter situation, i.e. bacteria growing in biofilms, although some comparisons to bacterial development in planktonic state will be included.

Biofilms can be defined as bacterial communities surrounded by polymeric matrices of extracellular matter and other associated products, most commonly attached to a surface or at an interface (109). The biofilm matrix itself can be an immensely complicated environment, ranging from one strain and all its associated products to multiple species (for example oral biofilms can contain more than 500 species of bacteria (110)). Generally, the associated products include eDNA, proteins, polysaccharides and lysed cell debris, but the matrix can also contain enzymes, RNA and abiotic materials (97, 111). Furthermore, biofilm communities typically grow in complex environments such as soil; a highly heterogeneous and geometrically intricate landscape (112, 113), which affects biological, ecological and physical processes in complicated ways.

Biofilm formation can be supported by virtually any nutrient sufficient environment, as is the case for general microbial growth (109). The biofilm phenomenon poses a significant challenge to industries and to human health, as bacteria within a mature biofilm structure are better protected against harsh environmental conditions and antimicrobial agents as compared to planktonic cultures (109). Indeed, such colonial growth can be seen as a strategy of unicellular organisms to gain the advantages that multi-cellular organisms have innately (114).

Biofilm control is of great importance to industries as their accumulation can cause significant economic losses, by causing, among other things, deterioration of equipment through inducing corrosion (115) or increasing fluid resistance (116). Furthermore, biofilm contamination may affect chemical processes involved in production, thus making them less effective. This is particularly important in the energy and chemical industries (117). Other note-worthy examples are the paper industry, where biofouling may have a detrimental effect on the quality of the final product, or the accumulation of biofilms below the waterline on the hulls of ships, which causes considerable losses for shipping industries by increasing drag, and what naturally follows, fuel consumption (117).

In contrast to generating losses, biofilm formation of some non-pathogenic bacteria can be utilized by industries, by e.g. inhibiting the growth of pathogens (118, 119), preventing fungi-related food spoilage (120), or engineering biofuels (121, 122). Microbes have also been recognized as useful in the treatment of wastewater (123, 124), cleaning up fuel spills (125), and even for their potential in generating electricity (101, 106, 126). The list of associations between biofilms and industries goes on and on and it is therefore no wonder that these bacterial communities are of great interest from an economical perspective.

Apart from generating significant interest directly from businesses, there are also great health concerns associated with biofilm formation (which are also connected with economic factors, albeit indirectly) (127). The problem is that there are innumerable species of human pathogens capable of forming biofilms, and many of these microbes, potentially dangerous to human health, are our constant co-habitants (128). Microbial contamination in the food, agricultural or medical sectors calls for, among other control measures, detailed exploration of possible disinfection methods, employed to prevent human disease outbreaks and to reduce the amount of food waste. The quest to gain control over microorganisms is extremely difficult, as these organisms have many tools at their disposal which aid their survival and growth. Developing resistance to antimicrobials (129) and cooperation with other microbial species (130), by e.g. quorum sensing (131), are a few examples of such survival tools.

It has been repeatedly shown that bacteria in a sessile growth phase are much harder to control than the bacteria grown in a free-floating state, and studies have been undertaken to understand what properties of biofilms give the bacteria embedded within a competitive edge against treatment (132). Mathematical models have significantly contributed to the field of biofilm formation in at least two important ways. First, mathematical models help to understand the key mechanisms involved in biofilm formation. These include quorum sensing (133–139), effects of multi-species interactions (140–142), antimicrobial resistance (143), or the mechanical properties of the extracellular matrix (144). Second, mathematical models are routinely used to inform strategies to prevent or promote biofilm formation in specific situations relevant to, e.g., food and water security (123, 145) or biofuel production (126, 146).

In this chapter, we give a concise summary of the current stage of application of mathematical models of biofilms, providing arguments for the continuation and further strengthening interdisciplinary collaboration within the field. We emphasise the applications of the models rather than their mathematical intricacies which are covered by other reviews (97, 147, 148). Section 2.2 describes results obtained from mathematical models used to understand key mechanisms for biofilm formation (see Table 2.1 for a summary of the reviewed models and Figure 2.1 for a schematic diagram of all sections discussed).The importance of mathematical modelling to address each of the selected topics is demonstrated by reviewing key findings based on state-of-the-art models that represent a substantial addition to the understanding gained through experimental approaches.

Table 2.1 *Summary of biofilm modelling work mentioned in this review.*

Author (Date)	**Model Description**	**Organism**	**Purpose**
O. Wanner, S. Gujer (1986)	1D, continuum, deterministic	Not specified	Study of the competition between autotrophs and heterotrophs in a multispecies biofilm (141).
W. Nichols et al. (1989)	1D, continuum, deterministic	*Pseudomonas aeruginosa*	Study of antibiotic penetration of biofilms of mucoid and non-mucoid strains (143).

E. Ben-Jacob et al.(1994)	2D, cellular automaton, stochastic	*Bacillus subtilis*	Exploration of patterns of bacterial growth in various nutrient conditions (149).
O. Wanner, P. Reichert (1995)	1D, continuum, deterministic	Not specified	Extension of previous work (141). General approach to modelling mixed species biofilms, exploring spatial profiles of chemical compounds and microbial organisms (150).
P. S. Stewart et al. (1996)	1D, continuum, deterministic	Not specified	Analysis of biocide action against biofilms (151).
C. Picioreanu et al. (2000)	2D, continuum, deterministic	Not specified	Study of the effect of biofilm surface roughness on the mass transport within the biofilm (152).
M. G. Dodds et al. (2000)	1D, continuum, deterministic	*Pseudomonas aeruginosa*	Analysis of antimicrobial resistance mechanisms of biofilms (153).
J. Dockery, J. Keener (2001)	1D, continuum, deterministic	*Pseudomonas aeruginosa*	General analysis of the quorum sensing mechanism in biofilms (133).
D. L. Chopp et al. (2002)	1D, continuum, deterministic	*Pseudomonas aeruginosa*	Prediction of acyl-HSL and oxygen

			concentration profiles within the biofilm and analysis of their effect on biofilm growth (154).
I.Chang et al. (2003)	3D, cellular automaton, stochastic	Not specified	Effect of transport limitation on microbial growth and biofilm structure (155).
K. Anguige et al. (2004)	1D, continuum	*Pseudomonas aeruginosa*	Analysis of effects of quorum sensing inhibitors and antibiotics on the quorum sensing mechanism of biofilms (134).
C. Picioreanu et al. (2004)	2D/3D, individual-based	Not specified	Analysis of the effect of multidimensional gradients on multispecies biofilm development (156).
J. Xavier et al. (2004)	3D, individual-based	Not specified	Comparison of CLSM data to spatial structures of multispecies biofilms generated by the model (157).
J. Xavier et al. (2005)	3D, individual-based	Not specified	Introduction of a general framework for IBM modelling (158) and evaluating the efficiency of biofilm treatment by

			detachment promoting agents (159).
K. Anguige et al. (2005)	1D, continuum	*Pseudomonas aeruginosa*	Quorum sensing inhibition (135); extension of (134).
S. M. Hunt et al. (2005)	3D, cellular automaton	Not specified	Analysis of antimicrobial action on biofilms, which focused on the scope of substrate limitation contribution on antimicrobial resistance (160).
J. D. Chanbless (2006)	3D, hybrid differential-discrete cellular automaton, stochastic	Not specified	Exploration of four hypothetical mechanisms of antimicrobial resistance, i.e. poor antimicrobial penetration, stress response mechanism, physiological heterogeneity within the biofilm and persister cells (161).
A. K. Marcus et al. (2007)	1D, conduction-based, deterministic	Not specified	Modelling the electrochemical processes in microbial fuel cells biofilms with focus on factors affecting electron flow (126).

J. Xavier K. Foster (2007)	2D, individual-based, deterministic	Not specified	Evolutionary outcomes of exopolymeric substances producers competing with non-producing individuals (142).
G. E. Kapellos (2007)	2D, hybrid differential-discrete cellular automaton, deterministic	Not specified	Analysis of biofilm growth dynamics in porous media. First modelling work to account for fluid flow through the biofilm (162).
F. Romero-Campero M. Pérez-Jiménez (2008)	P-system	*Vibrio fischeri*	Quorum sensing analysis using biochemical reaction networks (136).
J. Ward (2008)	1D, continuum, deterministic	Not specified	Investigation of anti-quorum sensing treatment of biofilms (135).
N. Jayasinghe R.Mahadevan, (2010)	1D, continuum model, combined with genome scale metabolism modelling	*Geobacter sulfurreducens*	Analysis of the effect of maintenance energy requirements on maximum current production and thickness of biofilms in microbial fuel cells (106).
M. Frederick et al. (2011)	2D, continuum, stochastic	Not specified	Analysis of how quorum sensing controlled EPS production affects

			biofilm formation (138).
Z. Wang et al. (2011)	2D, cellular automaton, deterministic	*Caldicellulosiruptfor obsidiansis, Clostridium thermocellum*	Study of cellulose degradation by biofilms in biofuel production (146, 163).
L. Lardon et al. (2011)	2D, individual-based	Not specified	Introduction of a biofilm modelling platform for non-programmers; iDynoMiCS (164).
D. Rodriguez et al. (2012)	2D/3D, cellular automaton, stochastic	Not specified	Studying effects of surface roughness patterns on biofilm formation in the presence of flow (165).
M. Asally et al. (2012)	2D, hybrid differential-discrete cellular automaton, deterministic	*Bacillus subtilis*	Theoretical analysis of mechanical forces behind emergent pattern formation of biofilms (166).
F. Pérez-Reche(2012)	3D, network, stochastic	Not specified	Analysis of network representation of soil samples with regards to potential microbial invasions (113).
R. Ferrier et al. (2013)	2D, individual-based, stochastic	*Listeria monocytogenes*	Estimating counts of food spoilage organisms on the surface of cheese (167).
A. Ehret, M. Böl (2013)	3D, continuum, deterministic	*Pseudomonas aeruginosa*	Study of mechanical role of EPS matrix

			on biofilms, representing the EPS matrix as a worm-like chain network (144).
S. Bottero et al. (2013)	2D, cellular automaton, stochastic	Not specified	Examination of factors influencing the development of flow paths in a biofilm formed in porous media (168).
W. Harcombe (2014)	2D, differential-discrete model, combined with genome scale metabolism modelling	*Escherichia coli* *Salmonella enterica* *Methylobacterium extorquens*	Proposed a modelling framework for incorporating genomic scale information on the scale of microbial communities with the aim to predict the behaviour of multispecies consortia (169).
N. Jayasinghe et al. (2014)	1D, continuum model, combined with genome scale metabolism modelling	*Geobacter sulfurreducens*	Metabolic modelling of spatial heterogeneity of biofilms in microbial fuel cells (170).
J. Cole et al. (2015)	3D, continuum model, combined with genome scale metabolism modelling	*Escherichia coli*	Analysis of the effect of metabolic interactions within densely packed biofilm colonies, i.e. the relation between a cell's position

			within a colony and its metabolism (68).
B. Emerenini et al. (2015)	2D/3D, continuum, deterministic	Not specified	Analysis of biofilm detachment regulated by quorum sensing mechanism (139).
R. Bennett et al. (2016)	Hydrodynamic, deterministic	*Pseudomonas aeruginosa* et al.	Analysis of individual cells flagellar spinning movements on the surface in early biofilm development (171).
P. Phalak et el. (2016)	1D differential-discrete model combined with genome scale metabolism modelling	*Pseudomonas aeruginosa, Staphylococcus aureus*	Role of metabolic factors on the spatial distribution of cells in a two species biofilm. The species were chosen for their common occurrence in chronic wound infections (69) .
M. Azari et al. (2017)	Activated Sludge Model	*Candidatus brocadia* et al.	Wastewater treatment reactor study (123).
B. Né Dicte Martin et al. (2017)	2D, cellular automaton, stochastic	*Streptococcus gordonii, Porphyromonas gingivalis*	Assessment of mixed species interactions in oral biofilms (140)
I.Tack et al. (2017)	2D, individual based, stochastic	*Escherichia coli*	Analysis of the effect of various environmental factors on the

			biofilm morphology (172).
K. Coyte (2017)	2D, hydrodynamic, game theory	*Escherichia coli*	Analysis of the relative success of microbial strategies in porous media for various flow conditions (173).
S. Stump et al. (2018)	2D, cellular automaton, stochastic	Not specified	Study of the competition between co-operators and cheaters within a microbial community (174).

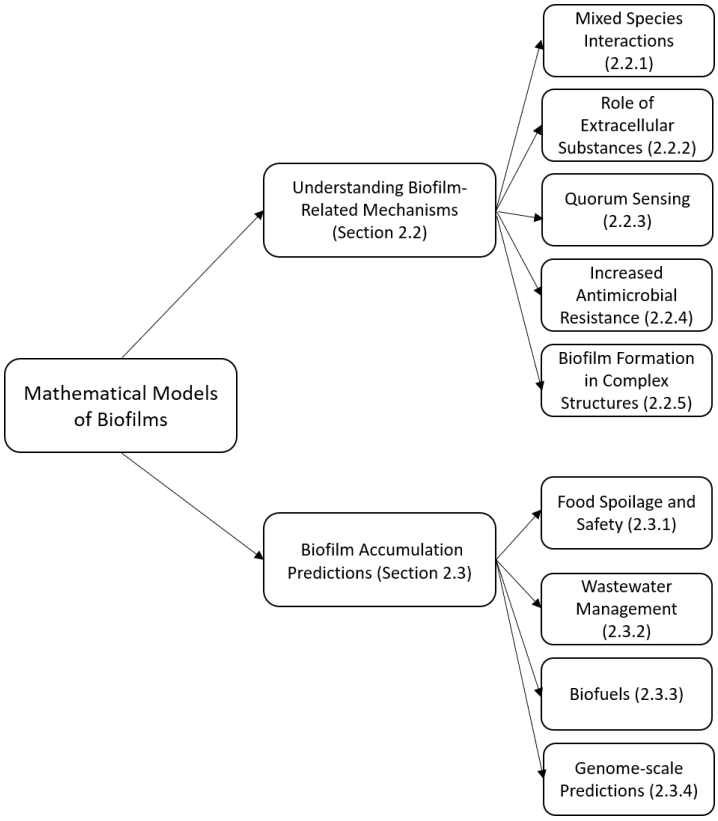

Figure 2.1 *Schematic diagram of the review. The biofilm models are categorised according to their purpose. Firstly, models which aimed to understand various biofilm formation mechanisms are discussed. We give examples of how mathematical modelling explained some observed phenomena arising from mixed species interactions, extracellular substances, quorum sensing mechanism, apparent antimicrobial resistance of biofilms and biofilm formation in complex structures. Secondly, attention is turned to second type of biofilm model, which aim to predict levels of biofilm accumulation. These models are generally specific to a given area of interest. We give examples of applications of these predictive models in the food industry, wastewater management and in engineering biofuels.*

2.2 Understanding biofilm-related mechanisms with mathematical models

Ever since the 1980s, efforts have been made to use mathematical descriptions to supplement experimental observations of biofilm communities. Many biofilm models have appeared since the initial efforts which considered one-dimensional, mono-species descriptions (175). These have been extended to add more spatial dimensions, more bacterial species, or by analysing the effects of varying environmental properties such as temperature, pH, fluid flow or spatial constraints from rough surfaces or porous media. The biofilm

models are either stochastic (145, 164, 176, 177), taking into account a certain degree of randomness of biological processes, or deterministic (178–180), if the stochasticity analysis is not needed to answer a particular question. They can be individual-based (156–158, 164, 167, 172), where each bacterial cell is considered as an entity, or mesoscopic (181–183), where an entity of interest is a whole colony or a microcolony of cells, and a single event may be for example population doubling. The models developed can focus on describing the biofilm at the scale of the whole population, or at the level of the individual cells, taking into account the details of cell structure and how it affects its behaviour (171). The fact that different models have been developed to focus on different spatial and temporal scales reflects the inherent multi scalar nature of the processes involved in biofilm formation (184, 185).

Although biofilm models may significantly differ from each other, they also have many things in common. Fundamental processes such as attachment, microbial growth, nutrient uptake, cellular death, extracellular products generation, detachment, and some chemical processes are usually introduced in some manner, albeit the methods used vary. For example, microbial growth in an individual-based model is introduced by a division of a cell with a set of rules governing the structural changes in the matrix following the introduction of a daughter cell. On the other hand, in models in which biomass is treated as a continuum, growth may be portrayed in terms of continuous biomass expansion and movement (97). Furthermore, diffusion of chemical compounds is generally introduced by solving Fick's law, convection is often governed by Navier-Stokes equations for fluid flow or their approximations, and nutrient uptake and biomass growth implementation usually includes a form of Monod equation (147, 148).

The following section presents examples in which mathematical modelling has proven instrumental to understand complex factors in biofilm growth whose elucidation using experimental methods remains a challenge. We will discuss the role of extracellular matrix and quorum sensing, the emergent antimicrobial resistance of biofilms and models which test viability of treatment methods, biofilm formation in complex structures and in mixed species biofilms. The list of topics presented here is by no means exhaustive. Due to the complexity of the field, we were forced to leave out many aspects, for example, the effect of motility of cells or factors influencing attachment (see, e.g. (186–188) for mathematical models incorporating some of these factors). We believe however, that the aspects we present give a taste of how mathematical modelling has been employed in biofilm research to this date.

2.2.1 Role of extracellular substances

The general role of the biofilm extracellular matrix (ECM) is to hold the biofilm together and fix it in place, but it has also been reported to be utilized by cells as a nutrient source (97, 111). By keeping the cells closer together, accumulation of quorum sensing signalling molecules is more likely to occur, making communication mechanisms more effective (111). Furthermore, the immobilizing properties of the ECM have the effect of keeping extracellular enzymes close to the cells and thus the ECM may act as an external digestive system (189). Other fundamental roles include facilitating gene transfer (190) or inducing formation of complex, self-organised structures (166). The ECM has also been reported to protect the biofilm cells from desiccation, biocides, antibiotics, heavy metals, UV light, host immune responses, and protozoan grazers (189).

In IbM models, individual agents such as bacteria cells or ECM material are treated as discrete entities, with specific properties assigned to them, such as their biomass, size and interactions with the environment. These agents are typically placed in continuous space, which is what puts IbM models apart from Cellular Automaton (CA) models, in which space is discretised in the form of a lattice (156).

A study using an individual based model (IbM) in 3 dimensions has been conducted to assess the potential of enzymic disruption of the ECM as a biofilm control strategy (158, 159). Prior to the theoretical study, the ability of NaOH to break down *Staphylococcus epidermidis* biofilms was confirmed experimentally, resulting in the need to identify factors affecting the efficiency of the treatment which could potentially be applicable for other bacterial species (191). The simulations had two stages. In the first stage, a biofilm was developed without the presence of disruptive enzymes. Subsequently, after a simulation time of 60 days, the biofilm was treated with a chemical compromising the ECM matrix, along with activating flow in order to trigger the detachment effect of the weakened biofilm structure. The modelling study found that 99% of biofilm removal resulting from the treatment occurred quickly, i.e. within a couple of hours. However, it took much longer for the remaining biofilm to be removed, i.e. 94 % of the total treatment time. Another interesting result obtained by the study was that the efficiency of the treatment in the simulations depended strongly on the ratios between the decay rate of the treatment substance in the biofilm, the rate at which the substance was able to compromise the ECM produced by the bacteria in question, and the rate at which the bacteria produced ECM. In some cases, the production of ECM was sufficient to counteract the effects of the treatment, resulting in persistence of the biofilm. The results of the study thus underlined the role of ECM material in biofilm prevalence, as well as provided possible reasoning behind differences in the relative success of biofilm treatment targeted at various bacterial strains.

The results of mathematical analysis of the role of ECM in protecting cells from antimicrobials will be discussed in later sections on antimicrobial resistance of biofilms. Now we introduce another modelling example, which analysed the influence of the ECM on the interactions between different species within the biofilm community (142). This individual based modelling study of mixed species biofilms has challenged the common perception of exopolymeric substances (EPS) production within the ECM matrix as a purely cooperative behaviour. Computational analysis identified the potential evolutionary advantage of EPS production in terms of aiding the individual's genes propagation. The study considered two species, in all other aspects equal, except that one produced EPS and the other did not. The non-EPS producer grew faster, as it had more resources available to allocate for reproduction compared to the other species. Simulations of the competition between two species have shown that the outcome was strongly dependent on the ratio of EPS produced per biomass formed and the ratio between the density of the EPS to the biomass. In some cases, the non-producing species indeed had an advantage over the EPS producers. It is interesting, however, that the EPS producers were favoured when the density of the EPS was lower than the density of biomass, for a wide range of EPS production rates and diffusion coefficients of the growth-limiting compound. This extended to being able to "suffocate" its rival with its generated product, while displacing the individuals of its species towards the top of the biofilm, where nutrients were more abundant. The authors of this study argued that considering EPS-producing behaviour solely as a group-benefiting sacrifice may be wrong, as this behaviour may be capable of causing a detrimental effect towards the neighbours of the producers.

2.2.2 Role of quorum sensing on biofilm formation

Quorum sensing (QS) is a means of cell-cell communication using signal molecules (autoinducers), allowing bacteria to sense the changes in their environment and react appropriately by activating or inhibiting gene expression (192). This phenomenon is thought to have a greater impact on bacterial communities in biofilms, as opposed to the planktonic phase, due to closer clustering of cells, which increases the number of signalling molecules in the external environment of the cells and may thus be a cause of increased QS associated gene expression (132). The QS mechanism has been reported to greatly affect biofilm formation. It has been suggested to play a significant role in attachment of cells or their detachment. For example, disrupting the QS mechanism in *P. aeruginosa* biofilms has been observed to result in thinner biofilms (193). The effect of quorum sensing on *P. aeruginosa* biofilms may well be a consequence of the fact that approximately 6% of all *P. aeruginosa* genes seem to be regulated by this communication mechanism (194).

Synthetic engineering of Quorum Sensing Inhibitors (QSI) has been suggested as a possible solution to aid eradication of unwanted biofilms. It has been observed experimentally that supplementing tobramycin as an antibiotic treatment of *P. aeruginosa* biofilms with a garlic extract, a natural QSI, was successful in killing all biofilm cells, a result that was not obtained when using either one of the compounds alone. Interestingly, disrupting the growth of cells within biofilms through manipulating their quorum sensing mechanism is not solely a man-made concept. For example, it has been observed that inhibition of quorum sensing can be imposed on one bacterial species by another within a mixed species biofilm (195).

Several mathematical models have been developed over the years to describe the role of QS on biofilm communities (133, 136, 137, 139, 154, 181, 196, 197). For instance, the study in Ref. (196) predicted diminished role of the QS mechanism in a biofilm exposed to high flow rates, in agreement with experimental observations.

The factors that may influence the effectiveness of *P. aeruginosa* biofilm treatment by disrupting cell-cell communication were analysed in a theoretical study (135). A critical biofilm depth was predicted, above which the treatment with QSI inhibitors would not be successful. This is thought to be partly due to a predicted exponential increase of the successful concentrations of QSI, or for that matter, any kind of antimicrobial compound, with biofilm depth (135). In contrast, in the case of planktonic cultures, the concentration of antimicrobials needed to eliminate the population of cells has been predicted by a previous theoretical study to increase linearly with the amount of treated biomass (134), which may be one of the direct causes of the difference in antimicrobial sensitivity between these two modes of bacterial growth.

In another application, a two-dimensional, deterministic model designed to study the quorum sensing mechanism has been proposed by Frederick et al. (138). Specifically, it aimed to investigate whether the QS regulation of EPS production by cells may be beneficial compared to a non-regulated, steady extracellular excretion process. Cases when EPS could serve as a nutrient source and when it could not, were investigated separately under high and low nutrient conditions. It was found that upregulated EPS production does not provide an advantage in terms of achieving higher population numbers, when compared to steady, low EPS production. It may, however, increase the optical density of the biofilm and thus protect the cells from environmental stresses or trap nutrients and thus lead to out-competition of the low-EPS producing rivals in nutrient rich conditions, even though the EPS production comes at a cost of slower growth (138).

2.2.3 Increased antimicrobial resistance

The structure and chemical composition of a mature biofilm provides a barrier which in many cases protects embedded cells from antimicrobials. This causes significant concern in the medical sector, among other industries (198) . Biofilm-caused infections often result in the development of chronic illnesses in patients, with available treatments inadequate in completely eradicating the bacteria within the biofilm. These can include foreign-body infections, e.g. biofilm formation on surgically inserted medical implants, or infections of regular tissue, e.g. lung tissue (199). Chronic patients must often maintain a constant, life-long treatment with antibiotics in order to keep the biofilms at a manageable level. However, this solution, among other things, disrupts the normal gut flora which may cause further deterioration of the overall health of the patient and may as a consequence cause the emergence of bacterial infections resistant to all types of available antibiotics. This in turn renders further treatment even more challenging and ultimate eradication of the infection difficult (132). Increased antimicrobial resistance of cells in biofilms is believed to be caused by many factors including, for example, increased level of mutation in biofilms in comparison to their planktonic counterparts. This phenomenon in turn is believed to emerge due to increased cell-cell communication in the biofilm community, where cells are naturally bundled closer together than in the case of bacteria floating in a free planktonic state (132).

The increase in mutations can cause upregulation of genes responsible for production of enzymes which degrade antimicrobial agents, or increased activity of efflux pumps, which expel the antimicrobial agent out of the cell membrane, making the bacteria more tolerant to antibiotic exposure.

In addition to increase in mutations and its effects in increasing antimicrobial resistance, development of chemical gradients in the biofilm layers is also believed to contribute to the persistence of treated biofilms. The chemical gradients of nutrients and other substances within the biofilm structure cause the emergence of dormant cells in the layers of the biofilm where nutrients become limited, while the dividing cells occupy the outer layers, closer to the biofilm surface. Some commonly used antibiotics exclusively target either dormant or active cells which is why using only one type may not prove sufficient to kill all cells within the biofilm. However, applying both of those antibiotics at the same time seems to be able to overcome this particular problem. For example, synergistic treatment with ciprofloxacin and colistin have been observed to be successful in clinical trials on patients in the early stages of cystic fibrosis (132).

Another advantage gained by the cells from the structural properties of the biofilm matrix is that diffusion of antimicrobials through the matrix may be significantly delayed, or even inhibited due to the chemical composition of the matrix, by breaking down or trapping

the antimicrobial compound before it reaches the cells within biofilm depths. Pre-treatment of the biofilm with enzymes degrading the biofilm matrix has been demonstrated to be a successful strategy by rendering the biofilm more susceptible to application of antimicrobials in a study involving *P. aeruginosa* biofilms (132).

Numerous modelling efforts have been employed in order to address the challenge of biofilm treatment with antimicrobials (158, 200–205), for example, a hybrid differential-discrete approach which tested four biofilm survival mechanisms separately (i.e. slow penetration, stress response, altered microenvironment and emergence of persisters). It was found by the study that the survival behaviours predicted by the simulations for each of the mechanisms were clearly distinct from each other. This result can be useful for determining the most dominant protection mechanism in an observed scenario and thus could prove informative in terms of choosing prospective disinfection strategies (201).

In another example, a continuous, diffusion-reaction, one-dimensional model, has been employed in order to predict antibiotic penetration into *P. aeruginosa* biofilms, in order to test the viability of antibiotic treatment for cystic fibrosis patients (143). Tobramycin and cefsulodin were chosen as antimicrobial compounds, and a mucoid and non-mucoid version of the *P. aeruginosa* biofilm were modelled in the calculations, in order to assess how the physical barrier of mucus affects the resistance of the biofilm embedded bacteria to chemical treatment. Interestingly, the results pointed to the conclusion that even though the diffusion of the antibiotic was substantially delayed in the mucoid phenotype when compared to the non-mucoid phenotype, the penetration time difference was not significant enough to account for the reported antimicrobial resistance. That is, the time it took for the antibiotic concentrations to reach high levels at the base of a 100 μm thick biofilm was still well within the common treatment time of cystic fibrosis patients. Furthermore, even when accounting for adsorption of the antibiotic to the exopolysaccharide, the concentration of the antibiotic at the base of the biofilm was eventually able to reach the concentration at the substratum. In the light of these calculations, it was concluded that the exopolysaccharide itself should not be considered as a significant physical protection barrier for *P. aeruginosa* biofilms against antibiotics.

Another hypothesis tested in (202) was whether the effect of bacterial production of enzymes is sufficient to effectively break down the antimicrobial compound. Assuming the enzymatic breakdown of an antibiotic in the model led to a phenomenon in which the concentration of antibiotic at the base of the biofilm could not rise above a certain threshold, as the diffusing substance would be continuously removed by the cell-produced enzymes. Simultaneously, it was observed that bacterial cells exposed to cefsulodin grew very slowly,

and thus it was hypothesized that slow growth may be another likely reason for increased tolerance of the bacteria. There may be many reasons for this phenomenon, for example, bacteria in a state of low metabolic activity may naturally allow less uptake of substances into the cells, therefore decreasing uptake of the toxin. Furthermore, low metabolic activity may be caused by upregulated production of toxin-degrading enzymes or upregulated activity of toxin-expelling efflux pumps. Results of experimental studies support the hypothesis that the concentration of biocides required for successful disinfection is much greater when applied to biofilms compared to planktonic cultures (206).

In another theoretical study, the efficiency of a biocide, benzalkonium chloride and peracetic acid, against *P. aeruginosa* biofilm was analysed (207). When comparing the susceptibility of different strains of *P. aeruginosa* to benzalkonium chloride treatment, considerable differences have been found between the resistance of strains grown in biofilms (in contrast with planktonic cultures where no significant difference was found). In particular, the difference in the time it took for the antimicrobial activity to reach the depths of the biofilm cluster, and the resulting changes in the total inactivation rate of the bacterial cells, all seemed to confirm the crucial role of ECM in determining disinfection efficiency. Moreover, it has been found that, in agreement with the modelling study, most cells within the biofilm have been deactivated during a short treatment time of 25 min, with few live cells remaining.

At present, biofilm treatment with enzymes is applied in industrial (208) and marine applications, and research is being undertaken to apply this strategy in the hospital setting with regards to development of antibacterial coatings for implants (132, 209).

2.2.4 Biofilm formation in complex structures

Experiments and models often describe biofilm communities growing on relatively simple substrates (e.g. flat surfaces). However, extremely flat surfaces on, e.g., the micrometre scale are an exception only found in some artificial settings (165) and most natural biofilms grow on rugose surfaces or porous media. Indeed, most bacteria on the planet inhabit structurally complex environments such as oceans or soils (112, 210).

The opacity of natural porous media makes it very challenging to study biofilm formation using only experiments. This fact has been recognised in e.g. predicting biofilm growth inside the cheese matrix, among other complex food structures (211) or questions regarding bacterial invasions of the gut (212). Applications of mathematical modelling to understand microbial growth in porous media is still limited but we believe that mathematical models can significantly help understanding this phenomenon. A theoretical framework for generic

biological invasions in porous media found that the shape, size and connectivity between pores within the medium plays a fundamental role in determining the extent of a potential microbial invasion (213). In this study, the structural heterogeneity of the soil pore space was captured through a network description with edges and nodes representing channels and bifurcation points in the pore space, respectively. Biological invasions were numerically simulated as a stochastic epidemic spreading on the pore space network. Based on the topology of the networks of the porous medium, the authors argued that structural heterogeneity typically favours biological invasions. The growth of biofilms in porous media has been recently studied experimentally (214) and theoretically (162, 168, 173, 215) but understanding is still limited due to the complexity of the problem. The difficulty of considering microbial accumulation in porous media is amplified by the fact that this network of flow channels is generally not static, i.e. various events, including microbial activities, lead to repeated clogging and unclogging of channels, formation of new channels, etc. (168). An approach combining fluid dynamics with game theory and experimental techniques revealed that in porous media, relatively strong and weak flow conditions favour fast and slow growing microorganisms, respectively (173).

Mathematical models have also been applied to study the effect of heterogeneity of abiotic surfaces on biofilm formation (165, 216–219). Some of these studies use computer fluid dynamics (CFD) modelling which may be combined with reconstruction of specific surface topography by Surface Element Integration (SEI) techniques, to assess the combined effect of flow and roughness patterns on biofilm accumulation (216, 218). These are highly advanced models, which can provide a detailed analysis of biofilm formation in a specific scenario. However, we discuss below in more detail results of a study which addressed the effect of surface roughness on biofilm formation with a cellular automaton, which we believe give a more general view of the problem (165). In cellular automata, space is discretised into equally sized patches, forming a lattice. Each patch may contain several objects (e.g. cells, extracellular material, oxygen or nutrients in (165)) and rules are introduced as to how objects interact with each other and with their environment. Properties of both objects and the environment may be defined as required. The authors in Ref. (165) argued that surface roughness may aid or inhibit biofilm formation when the flow of liquid above the biofilm is of considerable force, depending on the topography of the surface (165). The study focused on roughness on the length scale of a bacterial cell, i.e. at around one micron. The motivation for studying surface roughness of such magnitude was to address biofilm growth on mechanically milled surfaces, as the effect of roughness patterns of these surfaces may be an important factor for industrial applications. The modelling study found that in the case when flow is an important factor, biofilms growing on flat surfaces are easily

washed out. However, for otherwise identical environmental conditions, if blocks of size comparable to a single bacterium are fixed on the surface, the bacteria at the cracks between these blocks may become sheltered from the erosion effects of the flow, and are thus allowed to colonize, expand, and spread to downstream regions of the surface. This study found that one of the key factors determining whether roughness was beneficial to the development of the biofilm or not, was the spacing between the roughness blocks. If the spacing was too small, the resulting biofilms were flat, with less cells, as space for development was scarce; if the spacing was too large, the sheltering effect was insufficient to prevent flow-induced detachment. Furthermore, increasing the height of the blocks was also predicted to present a problem for the bacteria, as at sufficiently low niches nutrients could become limited, inhibiting biofilm development at the sheltered locations.

The results of the study discussed above provide a better understanding of how exactly some surface roughness patterns affect biofilm formation. In comparison, through experimental observations, it has been reported that when mimicking the conditions of a drinking water system, with flow adjusted to 10 cm s^{-1} , matt stainless steel accumulated a significantly greater number of microorganisms than electro-polished or bright annealed stainless steel (220). A separate experimental study on 316L stainless steel confirmed that bacteria may exhibit higher colonization levels at the cavities present on the unpolished metal surface (221). Interestingly, although many experimental studies simply conclude that increased surface roughness seems to promote biofilm accumulation (220–223), when investigated more closely, the surface topography, i.e. the depth and size of the cavities on the surface, has been found to be of more importance (224–227). The latter conclusions are supported by the modelling study of Rodriguez et al. (177).

It is worth noting, that nowadays the engineering of surface coatings with topographies designed to reduce biofouling are extensively studied, as technological advances allow for creating topographies of exquisite detail (226–228). In addition to the topography, other fundamental factors have to be considered in such designs. These include, but are not limited to, the surface free energy, wettability, elasticity, and antimicrobial properties of the surface (227).

2.2.5 Mixed species interactions

A single species biofilm is in most cases a laboratory construct, as the natural environment is full of microbial life and growth of single species seldom occurs in isolation. It is therefore mixed-species biofilms that are mostly apparent *in situ*, and thus the study of inter-species interactions within a biofilm is of great importance in addressing the challenges

associated with biofilm control. Studying the role of mixed species interactions on biofilm growth is experimentally challenging (140) and mathematical models can be of great help (140, 229, 230). In particular, we describe two recent applications of mathematical models which reveal key mechanisms in biofilm communities involving multiple species.

Recently, a new 2D cellular automata (discrete space and time) model has been developed to study biofilm formation of two species of bacteria, *Streptococcus gordonii* and *Porphyromonas gingivalis* (140). These two species have been identified as the leading causes of periodontitis, commonly referred to as gum infection, which can lead to tooth loss around the infected area. The study was performed to address the gaps in knowledge on the initial development of this two species biofilm, which follows after adhesion to periodontal tissues. Experiments informed by the model were performed to verify simulation outputs against observation. The model was designed to test whether the relationship between *S. gordonii* and *P. gingivalis* in the initial stages of biofilm development was independent, competitive or detrimental. The results of the simulations agreed with experimental observations only for the detrimental case, i.e. when it was assumed in the model that *S. gordonii* produces a compound which slows down the growth of *P. gingivalis*. This finding is in line with the fact that *S. gordonii* is known to be able to produce hydrogen peroxide, while *P. gingivalis* is known to be sensitive to this compound. Furthermore, it has been suggested by array analysis and reverse transcription PCR that oxidative stress response may be triggered in *P. gingivalis* in the presence of *S. gordonii* (231). In summary, the model has been able to provide evidence for a detrimental effect of *S. gordonii* on the growth of *P. gingivalis* in a two-species biofilm, following adhesion.

In another recent example, a stochastic two-dimensional cellular automaton model was applied to study mutualism versus exploitation in a microbial context (174). In particular, the study analysed potential mechanisms which could promote the success of bacteria producing nutrients for other organisms, over "cheating" bacteria which did not produce any nutrients. The results of the contest between the two species exhibiting these distinct behaviours were mapped against the distance between the microbes and the distance at which the produced resources could reach other microbes. It was shown that, consistently, for high cell dispersal and high reach of the shared resource, cheaters had a competitive advantage, and after reaching a certain threshold for these parameters, extinction of the co-operators was predicted. It was reasoned that for these conditions, the cells were forced to interact with many random neighbours, thus making co-operators open to exploitation. In contrast, the case when both cell dispersal and reach of the resources were low, provided an opportunity for groups of co-operators to persist against the invasion of the cheaters. Interestingly, for intermediate conditions, i.e. high cell dispersal and low reach of the resource, or low cell

dispersal and high reach of the resource, the co-operators also were found to persist. In the former case, it was found that the uncertainty of the interactions between neighbours harmed the exploiter, as it led to uncertainty of resources. In the latter case, the community exhibited self-organised pattern formation, in which co-operators organised themselves into stripes or spots. The conditions within these organised groups were such that they limited the growth of cheaters. It is noteworthy that such patterns are reminiscent of similar phenomena observed in biofilms.

2.3 Applications of mathematical models in predicting biofilm formation

Biofilm models have proliferated due to a need to answer particular questions stemming from areas where biofilm formation is a significant concern. Today, modern theoretical biofilm models are recognized for their ability to, among other things, analyse spatial interactions between organisms within a biofilm on an individual scale (232). Other models may focus their analysis on predictions of biofilm formation in specific environments [10,26,27,32]. In the previous section, we have discussed the former, i.e. models developed in order to understand the role of various factors on biofilm formation. In this section, we will focus on the models which aim to predict accumulation of biofilms. For example, the output of such models may be a prediction of bacterial counts on a given surface (145), or a detailed biofilm composition in the studied environment (123).

2.3.1 Food spoilage and safety

It is recognized that food spoilage depends on factors such as storage conditions, initial unwanted microbial counts in the food and their properties, and finally, the properties of the food involved, such as its pH or moisture. Estimating the shelf life of food products has been aided by means of mathematical models developed as early as the 19th century (235, 236), and the value of these microbial count models for the food industry is now widely appreciated at the product development stage (237).

Empirical models build on data obtained from storage trials are common among models employed to predict shelf life (238–240). These models are characterised by a systematic experimental approach, in which the effect of a specific variable (e.g. temperature) on microbial growth is assessed. Data collection is followed by fitting experimental data with a theoretical curve in order to analyse the correlations between considered factors, formulate general hypotheses, and subsequently allow for making better

predictions. One of the notable examples in this area is the work by Ratkowsky et al. (240), in which the authors proposed a general law governing the relationship between the temperature and growth rates of bacteria. The results of the Ratkowsky et al. study were found to fit experimental data better than what was predicted by Arrhenius Law (235, 241) (this is a classical law describing the relation between chemical reaction rates and temperature). Furthermore, a slight modification of the Ratkowsky et al. model (240) was found to fit empirical data for a temperature dependency study of *Lactobacillus plantarum* growth (239). Apart from temperature, other factors affecting growth have been empirically modelled, e.g. the effect of carbon dioxide on growth of *Photobacterium phosphoreum* and *Shewanella putrefaciens* (238).

More recently, predictive modelling has been employed to estimate bacterial growth in seafood, dairy, bakery, vegetable, meat products, and other products, e,g, infant formula or acidified sauces (211). For example, one of the recent approaches used an individual-based stochastic model, able to accurately predict *Listeria monocytogenes* counts on soft cheese (167). The individual based approach, so far uncommon in the area of predicting the microbial shelf life of food products, was introduced in order to account for variability in the microenvironment of individual cells.

The area of predictive modelling for food safety is so vast that it is beyond the scope of this chapter to go into the amount of detail it deserves. For an extensive, recent evaluation of this particular topic, the reader is encouraged to turn to the book by Mahony and Seman (211).

It is noteworthy that apart from predicting growth of microorganisms during food storage, empirical mathematical modelling has also been applied to address other food safety concerns. For example, a relationship describing cross contamination of *Escherichia coli* and *Listeria monocytogenes* from slicer to deli meat has been proposed based on experimental data (242, 243).

2.3.2 Wastewater management

The use of bacteria in the Activated Sludge (AS) process, designed to treat water systems, dates back over a hundred years and it is safe to say that this invention revolutionised wastewater management (244). Computational modelling of microbial communities can contribute to engineering safe water treatment reactors by, for example, testing for mathematically plausible causes for the occurrence of some observed phenomenon. This may include testing the nature of interactions between microorganisms present in the reactor (123). Such models aim to simulate a typical environment of a

wastewater system, in order to predict the distribution and relative concentrations of various microorganisms and their effectiveness in water treatment.

Activated Sludge Models (ASM) is the name given to the specific type of a biofilm model designed to optimize the AS process. ASM models describe processes such as oxygen consumption, sludge production, nitrification and denitrification in the activated sludge designed to treat water systems (245). ASM models serve as a good example for specialised models which can be widely adopted in the field they are designed for (246). These models can aid the daily operations of plants, as well as the development of plans for introducing modifications. A careful design and continuous improvement are fundamental in using ASM models as tools for the wastewater industry, as significant decisions with financial and environmental implications may be based on their predictions. With the incorporation of computational models into water treatment industry comes the necessity to develop stringent procedures for accurate software usage and interpretation of the model's outputs, a task which has been taken on by the International Water Association (247). It was estimated, that in 2009, the number of ASM users worldwide was between 3000 and 5000 and included university and public researchers, as well as private company employees (247).

The ASM1 model describes the water purification system by a series of processes which take place in the reactor. The processes are governed by substrate-dependent rates and by stoichiometry of the occurring reactions in each process (246). The rates of all processes are described by various equations; for example, growth of biomass is unsurprisingly modelled by use of Monod relationships (248). The other processes modelled by ASM1 are the decay of biomass, ammonification of organic nitrogen and hydrolysis (246).

A very recent example of a biofilm model designed for wastewater management purposes was presented by Azari et al. (123). The model had been developed with the aim of identifying the most important parameters affecting biofilm formation in an anammox reactor; a reactor engineered to remove ammonium from wastewater. The framework of the study was based on Activated Sludge Model no. 1 (ASM1). It has been found by the model that biofilm formation and ammonium removal was most affected by the maximum specific growth rate of organisms and heterotrophic biomass yield. The levels of nitrogen compounds and biofilm composition predicted by the model were in good agreement with experimental findings, suggesting that the results obtained by the simulations were reliable (123).

2.3.3 Biofuels

With advancements in technology, energy consumption has been rapidly rising. The need to move from non-renewable energy sources such as fossil fuels, to sustainable

solutions which rely on renewable energy sources, is apparent. Most people are aware of such solutions being applied in the form of harnessing solar, wind, geothermal or tidal energy. Surprisingly, it does not seem to be commonly known that microbes are also being utilized by the energy industry, for instance in engineering biofuels such as e.g. bioethanol, biodiesel or biohydrogen (249). However, biofuels have been claimed to have the biggest potential for reducing CO_2 release into the atmosphere (122). This is largely due to the fact that the demand for fuels makes up a majority of the overall demand for energy (250).

Biofuels can be produced by thermochemical means or by microbial fermentation (122). In the latter case, degradation of biomass (e.g. cellulose) by microbes (e.g. yeast, bacteria or mould) is a key process in biofuel production (251) Although there is already an established procedure for engineering biofuels, research is being undertaken to make this process more efficient (121, 146).The area of biofuels is a multifaceted one, as for instance complex chemical and biological reactions, as well as engineering solutions have to be designed and perfected for process optimisation. Advanced technologies, e.g. genomics, have been identified to be fundamental for maximizing the efficiency of biofuel production methods (121). Furthermore, given the undeniably immense global scale impact of the energy industry, the efforts for engineering biofuels should be done in close cooperation with environmental scientist (252). One review on microalgal biofuels listed fundamental biology, systems biology, metabolic modelling, strain development, bioprocess engineering, integrated production chain and the whole system design, as areas which need to be included in the biofuel research portfolio. The biggest share of mathematical modelling in aiding biofuel production process engineering probably lies in metabolic modelling, which is a key part of the systems biology approach to metabolic engineering (253). However, as such techniques are performed on the scale of genomes, rather than bacterial populations, these models are beyond the scope of this review. Although we have not found in the literature the link of population scale metabolic modelling to biofuel production, it should be noted that some recently published studies combined genome scale metabolic reconstructions with differential equations for the diffusion of metabolites, thus creating genome scale resolution models of biofilm populations (69).

There are not many papers available which explicitly link biofuels to biofilm formation, and this may be due to the fact that smaller scale modelling integrated in the system biology approach has been found more applicable for this field. We will presently discuss results of a modelling study which did focus on population scale degradation of cellulose.

A cellular automaton model has been developed which is able to mimic experimentally observed structure of biofilms formed by *Caldicellulosiruptor obsidiansis*

(234), and in a separate study, those formed by *Caldicellulosiruptfor obsidiansis* and *Clostridium thermocellum* on cellulose substrate (146). In the latter study, the observed thickness of the biofilm was achieved in the simulation by incorporating a detachment mechanism, which was activated once the biofilm thickness approached an observed threshold. It is quite plausible that a colony which feeds on the substrate to which it adheres will exhibit such behaviour, as this allows detached cells to float towards areas where nutrients are unexploited, i.e. to the non-colonized areas of the substrate.

Analysis of both experimental and computational results obtained from the study published in (146) seemed to point to the conclusion that cellulose degradation was synchronous to biofilm formation of the particular species. Moreover, only cellulose areas to which bacterial cells were attached exhibited degradation and increasing number of planktonic cells in the culture did not produce a significant effect. In the light of obtained results, the authors concluded that the process of cellulose degradation could theoretically be sped up by covering the cellulose substrate with a highly concentrated inoculum of cellulose-degrading cells (146).

2.3.4 Application of genome-scale reconstructions in biofilm modelling

With recent advancements in genomics, proteomics and metabolomics, there has been a rise in biofilm models which incorporate genome-scale data for obtaining more sophisticated predictions for microbial communities (68, 69, 106, 169, 170). The aim of incorporation of genome scale data in biofilm modelling is to improve the quantitative understanding of spatial and temporal variation of the microenvironment of cells embedded within a biofilm, which is believed to have a critical impact on biofilm development (69). A table of available genome-scale metabolic reconstructions which have been validated by experimental data can be accessed through Systems Biology Research Group web page (254). These reconstructions can be used to feed more information into biofilm models, e.g. the metabolic by-products, compound uptake fluxes, or the secretion of toxins and growth inhibitors of the documented strains. It has been suggested that the accuracy of predictions related to spatial partitioning of species within a mixed-species biofilm is enhanced by inclusion of the effect of metabolic factors (69, 169).

The studies which explicitly coupled genomic scale data and biofilm modelling have targeted e.g. illness related biofilms (69) or microbial fuel cells biofilms (106, 170). In another study of this kind which focused on *E. coli* biofilms, it was suggested that a similar methodology may also be useful for models of tissues or tumours (68). In essence, these studies incorporate differential equations for the diffusion of metabolites in population scale

models, and they do seem promising in terms of improving prediction power of mathematical models of biofilms. For example, in a modelling study of *E. coli* colonies grown on glucose minimal agar, incorporation of data from *E. coli* metabolic reconstruction led to the discovery of a feature of *E. coli* colonies which has not been recognised previously. The study found that glucose and oxygen gradients within the colony gave rise to four distinctly spaced metabolic phenotypes, namely, rapidly growing cells at the bottom edge of the colony, where both glucose and oxygen concentrations were high, nearly dormant cells in the interior, where both glucose and oxygen levels were low, and two other subpopulations between which acetate cross-feeding was found to take place. The first subpopulation, located at the base of the agar, exhibited high glucose consumption and acetate production due to high glucose concentrations. The second subpopulation, located at the regime of high oxygen concentrations and low glucose concentrations, exhibited a phenotype which favoured acetate consumption. In terms of the predictive power of this modelling study, the height to width ratios of simulated colonies were in agreement with those of colonies grown experimentally (68).

2.4 Summary

Mathematics can be used to understand and exploit the world around us. Examples of mathematical models of biofilm formation presented in this chapter only scrape the surface of the vast number of models which have been developed, from their earliest descriptions until the present. We presented some examples of biofilm models which significantly advanced our understanding of biofilm communities and generated results applicable, for example, to medicine, the food industry, dentistry, water management and for engineering more environmentally friendly energy.

Although computational models have been found useful over the years in providing practical answers about microbial communities, they do all have considerable limitations. The fact that a model is necessarily a significant simplification of reality is both a handicap and a strength, depending on the point of view and application. Just as the biofilm field is complex, so is the branch of biofilm modelling. This creates obstacles between model development and applications, because if the model is to be trusted, it must be verifiable in a specific setup for which it has been created. Furthermore, the wide use of any given model is difficult to achieve, as any model would have to go through modifications to become usable for another research problem. This requires understanding of the language in which the model source code was written, and a thorough grasp of the implemented processes. Luckily, when building a model to address a specific problem, one may build on the general rules

adapted by existing models and choose suitable methods of implementation for the question which needs to be answered. For instance, empirical models give an idea of the relations between specific factors affecting biofilm formation, e.g. the relationship between temperature and growth rates. Although these are built on specific experimental results, as evidence of their reliability builds up, they become widely adapted, as has been the case with Monod growth equations, for example. Empirical modelling has been particularly favoured when estimating bacterial counts is the priority of the study, as is the case in e.g. developing food spoilage prevention methods. On the other hand, in studying the interactions between biofilm components on the scale of bacteria cells, the mechanisms of biofilm organisation and structuring, or when considering structurally complex environments such as rough surfaces and porous media, spatial, individual based or cellular automaton models seem to be a suitable choice, as does the game theory approach. Furthermore, treating the biomass as a continuous, viscoelastic substance, may allow for applying mechanics laws in studying the material properties and behaviour of the biomass. Finally, for analysis of e.g. antimicrobial penetration of a biofilm, a one-dimensional model treating biomass as a continuum may be fitting for its purpose.

In their current form, mathematical models of biofilms can play a key role in addressing many important questions. For example, a proper combination of experimental and theoretical approaches will help understanding the behaviour of biofilm communities in some habitats that can be reasonably complex (e.g. through structural or chemical heterogeneity). Other questions will require holistic approaches accounting for biofilm formation at multiple scales, interactions between species and other factors. For instance, biofilms are likely to promote survival and persistence of pathogens in food-related environments (26). In this context, biofilms can be regarded as just one element of a larger multifaceted problem involving domains ranging from the natural environment to food production factories and consumers. Integrating the key factors in a single framework to address biofilms associated problems (e.g. risk assessment of food contamination), is a challenge that will necessarily involve mathematical modelling and data analysis combined with experimental approaches.

It seems that although great improvement has been seen over the years with regards to computational models of biofilm formation, with substantial useful information gathered from computational analysis, much work is yet to be done to bridge the gap between theoretical and practical aspects, in order to synergistically build a general set of principles by means of which microbial development can be understood. Although not an easy endeavour, it is a necessary next step to fully realize the potential of biofilm models in addressing new challenges associated with biofilm control and utilization. A relatively

recent, however fast developing field of systems biology promises to provide such an integrated framework (255). Systems biology has already been successful in engineering new solutions for e.g. biofuel or pharmaceutical industry (256). The idea behind this research field is to develop fine-detailed models of ecosystems which take advantage of the new advances in genome sequencing data collection (257). Among a plethora of potential applications of this technology, when paired with advances in computing, it can lead to development of highly sophisticated biofilm models. The high resolution methodology of systems biology has already been to some extent applied at the scale of whole populations of bacteria cells, for example by combining genome-scale metabolic modelling techniques with partial differential equations to model the spatial distribution of metabolites within the biofilm (69). The systems biology approach requires a high level of cooperation between various disciplines. In building such fine-resolution models, apart from biology, expertise in fields such as chemistry, physics, engineering, and informatics may be necessary, depending on the research question. It is likely we will see more field-specialised biofilm models develop, as is the case with ASM models for wastewater management or shelf life prediction models. Before incorporating solutions to challenges of microbial control and utilization on a large scale, potential environmental concerns should be addressed, thus further widening the desirable network of collaboration in the biofilm research field. This sentiment has already been expressed by researchers in the biofuel field (252), however, it should extend to all areas capable of producing a large-scale impact on the environment.

Chapter 3 - A Novel Method to Approach Biofilm Modelling Through a Continuous Time Stochastic Cellular Automaton Biofilm Formation Model – Case Study of *Campylobacter jejuni* Biofilms

3.1 Introduction

Biofilms may be described as close packed bacterial communities surrounded by extracellular material. It is important to analyse biofilm communities separately from planktonic cultures, as behaviours of cells in these two states differ substantially (108). Moreover, biofilms are considered important mechanisms for survival of microorganisms in hostile environments. For instance, they have been shown to enhance resistance of bacteria to antimicrobial treatments (258). Increasing resistance to antibiotics is of particular significance to public health, as it contributes to persistence of chronic infections, among other concerns (69).

Mathematical modelling of biofilms has been incorporated in biofilm research for many years. It has proven to yield useful insights in many sectors in which biofilm formation is a fundamental issue, by for example helping to understand key mechanisms of biofilm formation or by making predictions on biofilm accumulation (52). Most models focus on estimates of the morphology of biofilms without estimating the chances of biofilm formation. Even though stochasticity has been included to various extents in biofilm modelling (158, 161, 164, 165, 172, 259), both modelling and experimental approaches generally focus on biofilms in the invasive mode, i.e. when the probability that the biofilm forms is high. Although in the biofilm modelling literature it is usually acknowledged that biological systems are often governed by stochastic dynamics, the probability of a successful surface invasion has not so far received much attention. Addressing this issue may be of particular importance for an organism such as *Campylobacter jejuni*, which on the one hand is known to be fastidious in terms of its survival and on the other, has been shown to be able to cause disease from a relatively small number of ingested organisms, i.e. approximately 500 (8).

C. jejuni is a microaerophilic pathogen, commonly found in a variety of environments including farms, food, and the intestinal tract of various animals. In particular, *C. jejuni* is frequently found in the gastrointestinal tract of poultry, causing subsequent contamination of poultry products. *C. jejuni* has been recognised to be one of the leading causes of human cases of gastroenteritis globally. In 2018, the World Health Organisation (WHO) reported *C. jejuni* to be the most common bacterial cause of gastroenteritis in the world (18). In 2017, it

was estimated that *Campylobacter* spp. was responsible for around 96 million foodborne infections around the world (260). Despite being fastidious in terms of its growth requirements, *C. jejuni* is quite a successful pathogen. A question which has been puzzling researchers is how *C. jejuni* manages to survive in the food production chain where conditions would be expected to be suboptimal for this microorganism (261). It is now believed that biofilm formation or attachment to existing biofilms of other species are both fundamental factors in mediating survival of *C. jejuni* outside of its host (262, 263). Although various models for *C. jejuni* have been proposed (30–32, 36, 46, 57, 60, 73–75), up until now, no attempt has been made to study *C. jejuni* biofilm formation through mathematical modelling.

In this paper, we propose a mathematical model that allows a number of observations regarding *C. jejuni* biofilm formation to be understood in a unified manner. Firstly, *C. jejuni* has been reported to exhibit higher biofilm formation in low nutrient media (MHB) compared to nutrient rich media (Brucella broth, Bolton broth) (40, 49). Furthermore, in spite of being a very sensitive organism to atmospheric oxygen concentrations, it was found that biofilms formed in food-chain-relevant aerobic conditions may be in some cases denser than those achieved in microaerobic conditions, while being equivalent reservoirs of live cells (47). Another study (42) also reported interesting facts that are difficult to reconcile with these results. Namely, aerobic conditions were found to induce bacterial lysis of *C. jejuni*. The same study found that starvation conditions significantly inhibit lysis.

Apart from the aforementioned observations, there are also questions which remain to be answered regarding the potential effect of extracellular matrix (ECM) disruption on biofilm formation of *C. jejuni*. Since extracellular DNA (eDNA) is a major constituent of *C. jejuni* biofilms (30, 42), it is puzzling that many *C. jejuni* strains (almost half, according to a study of 2791 strains) contain at least one gene in their genome which encodes for an eDNA disruption enzyme (DNAse) (264). Furthermore, although there are some strains with the DNAse gene which have been classified as poor biofilm formers, there are also many such strains which have been classified as relatively good biofilm formers, and there seems to be no clear association between presence of any of the three DNAse encoding genes in the genome and biofilm forming ability (265). For this reason, the mechanism of ECM disruption has been introduced in our biofilm model, to assess how it affects the simulated biofilms.

The biofilm formation model reported here has been based on the existing cellular automata (CA) biofilm models (149, 165, 266–268). Our model accounts for fundamental processes

such as the rearrangement of cells within the biofilm, diffusion of chemical compounds outside and inside of the biofilm, growth, lysis, deactivation due to insufficient nutrient resources, and an optional mechanism of ECM disruption. We consider the biofilm formation process as a biological invasion, which is inherently stochastic (113, 259, 269) and can occur with a certain probability, depending on the conditions the cells are subjected to. This description is reminiscent of models for spread of infections and biological invasions (113, 270, 271). It is important to note that there are many complex mechanisms involved in biofilm formation (some of them described in detail in Section 2.2), which were not included in this model, e.g., the quorum sensing, mixed species interactions, or attachment of cells to an existing biofilm from the planktonic state.

We notice that despite the specific application to *C. jejuni*, our model can be widely applied to biofilm formation of any bacterial species. Our main motivation to propose this model, however, was to understand the biofilm formation mechanisms of *C jejuni* - an important food-borne pathogen for which biofilm formation is still puzzling and poorly understood.

3.2 Model Description

Biofilm formation is simulated in a 2-dimensional environment with $L \times H$ patches of linear length δx arranged on a square lattice (Figure 3.1). The sites at the bottom of the lattice represent the solid surface on which the biofilm growth can initiate. The remaining sites represent the space above the solid surface where the biofilm can grow. Each patch in this region represents an area of space that can be occupied by fluid (*F*), a cell (*C*) or ECM (*E*). The state of the *i*-th patch at time t is given by a variable $\sigma(i,t)$ that can take values *F*, *C* or *E*. The nearest four patches to the *i*-th patch define its neighbourhood, $\mathcal{N}(i)$. The top and bottom boundaries of the system are closed, while periodic boundary conditions are set on the vertical sides.

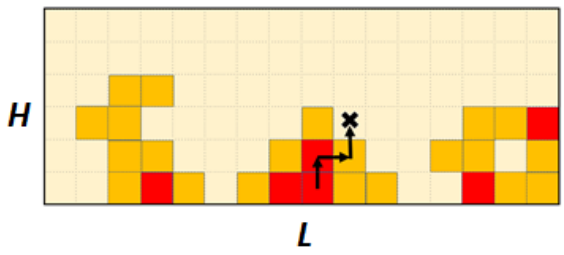

Figure 3.1 *Illustration of a directed random path. End of the path is marked by a cross. The start of the path is where cell division has occurred, thus calling for redistribution of the biofilm material to accommodate the new cell. In this example, no empty sites are encountered on the path, thus, the shoving algorithm will terminate at the end of the path. Red coloured patches represent live cells and orange patches represent the ECM material.*

Two types of chemical compounds are considered: a carbon source and oxygen with concentration fields $S_c(i,t)$ and $S_o(i,t)$, respectively. In the case of *C. jejuni*, the carbon

source represents growth limiting compounds such as amino acids (70). Since *C. jejuni* is a microaerophilic organism, it requires oxygen to grow but small concentrations are optimal for cell survival. Indeed, it was observed that high oxygen concentrations promote bacterial lysis (42).

The system dynamics involve cell vital transitions (duplication, lysis, or deactivation) and substrate rearrangements associated with cell nutrient consumption and diffusion. We assume that concentration changes of the substrate are fast compared to vital dynamics events. Under this assumption, the system dynamics can be implemented as a sequential algorithm in which vital dynamics events are followed by fast substrate rearrangement, as shown in Figure 3.2. The next sections describe the implementation of cell vital dynamics and substrate dynamics.

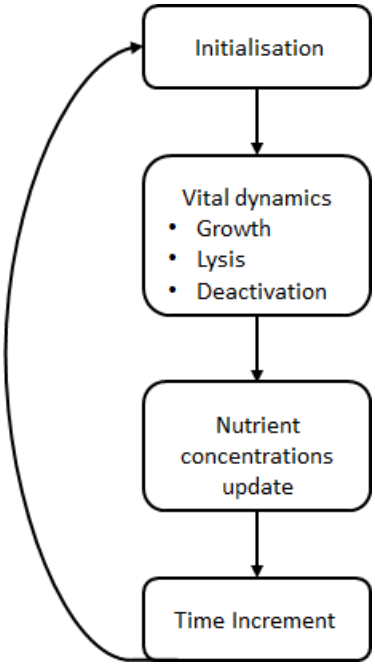

Figure 3.2 *Schematic diagram of the biofilm model simulation algorithm.*

The model assumes that duplication (i.e., growth), lysis or deactivation of a cell at patch i can occur randomly with transition rates $\lambda_{\mathrm{g}}(i,t)$, $\lambda_{\mathrm{l}}(i,t)$ and $\lambda_{\mathrm{d}}(i,t)$, respectively. Such rates depend on the substrate concentration fields $S_{\mathrm{c}}(i,t)$ and $S_{\mathrm{o}}(i,t)$ as described below. Transitions of the state of cells are simulated as random events in continuous time using a kinetic Monte Carlo algorithm (also known as n-fold way or Gillespie algorithm) (272), as described in Appendix A.1. This algorithm allows the time increment Δt between consecutive events to be efficiently obtained. We now describe the implementation of the three transitions of cells.

3.2.1 Growth

Growth occurs through duplication of cells. If a cell in patch *i* duplicates, the offspring is placed at a randomly chosen site in the neighbourhood $\mathcal{N}(i)$ of patch *i*:

$$C(i,t) \xrightarrow{\lambda_g(i,t)} C(i,t+\Delta t) + C(j,t+\Delta t), \qquad j \in \mathcal{N}(i) \tag{3.1}$$

Only for cases when one of the sites in $\mathcal{N}(i)$ are in the fluid state, the offspring is randomly placed in one of these sites. In contrast, if all sites in $\mathcal{N}(i)$ are occupied by cells or ECM, a random directed path is applied to redistribute the biofilm mass (defined as the set of patches occupied by cells or ECM). The path is defined by first selecting an endpoint at random at the interface between the colony holding site *i* and the liquid medium (e.g., see the patch indicated by '×' in Figure 3.1). The distances in vertical and horizontal directions from the site at which duplication occurred to the endpoint are recorded. Then, randomly ordered, directed steps in the vertical and horizontal directions are applied, during which the surplus mass is pushed to the neighbouring site on the path, displacing what is already there to the next step of the path. The path in our algorithm is "directed", as the only movement allowed is that which shortens the distance to the endpoint. In other words, if the endpoint lies on the upper left side to the location where growth occurred, the only steps taken will be to the left and upwards, in randomized order. This shortens the algorithm substantially compared to a regular random path, especially for growth within large colonies. The algorithm ends once a site occupied by fluid is reached, either through reaching the endpoint of the path or by reaching a fluid site along the way, for example a pore or a channel within the biofilm matrix. This mass redistribution algorithm is reminiscent of the one proposed in (266), albeit the latter chose the end point of the path by calculating the shortest distance from the growth location to the biofilm surface, rather than randomly. In our model, as we allowed for the algorithm ending when encountering an empty or partially filled space (the latter only possible when disruption of ECM is allowed), we left the choice of the biofilm-liquid interface endpoint as random. This was done, as in our model, choosing the shortest distance to the biofilm interface would not necessarily mean choosing the path of less resistance, as there may be pores along the way whose filling would offer less resistance.

In reality, cells would reorient themselves before moving further, provided that there is enough space to allow it. However, previously reported biofilm simulations, which took into account reorienting of cells, have revealed that cells at the periphery may also be effectively pushed as a result of cell replications inside the biofilm (273), which motivated our simplified version of the biomass distribution.

We assume that growth is promoted by the carbon source concentration $S_c(i,t)$ and optionally by ECM disruption. The growth rate $\lambda_g(i,t)$ of a cell at site *i* and time *t* is given by:

$$\lambda_g(i,t) = r_g \frac{S_c(i,t)}{S_c(i,t) + K_g} + \sum_{j \in \mathcal{N}(i)} r_E \cdot \begin{cases} 1 & \text{if } \sigma(j,t) = E \\ 0 & \text{otherwise} \end{cases} \quad (3.2)$$

The first element of the above equation represents a standard Monod growth curve depending on the local growth limiting nutrient concentration $S_c(i,t)$ with maximum specific growth rate r_g and Monod growth half saturation coefficient K_g. The second contribution to $\lambda_g(i,t)$ accounts for the effect of ECM disruption on cell growth. For r_E =0, no growth benefit is achieved from the ECM disruption and this may simulate the case when cells produce an enzyme which breaks down the biofilm matrix. For r_E >0, however, there is a growth rate increase from ECM disruption, and this may simulate utilization of extracellular material as a nutrient source. In particular, we assume that the growth rate is incremented by a bonus amount r_E [h^{-1}] whenever a cell finds itself in the neighbourhood of at least one ECM particle. The growth rate bonus is assumed to be proportional to the number of neighbouring ECM particles, as we allow for simultaneous depletion of ECM material in $\mathcal{N}(i)$.

For gradual depletion of ECM as a result of its disruption, we introduced a parameter N_E, which represents the number of times the growth rate of a cell neighbouring a patch containing ECM may be increased. Each time a cell causes depletion of part of the ECM material in a neighbouring patch occupied by ECM, the amount of ECM in that patch decreases by $\frac{1}{N_E}$ of its original value. In our simulations, N_E =4.

The blue dashed curve in Figure 3.3 illustrates how the growth rate increases with increasing nutrient concentration and asymptotically approaches the maximum value, r_g. Here, the parameters r_g and K_g take the values given in Table 3.1. If ECM uptake is active, the rate $\lambda_g(i,t)$ will increase irrespective of the value of $S_c(i,t)$ (i.e., the curve for $\lambda_g(i,t)$ in Figure 3.3 would be vertically shifted).

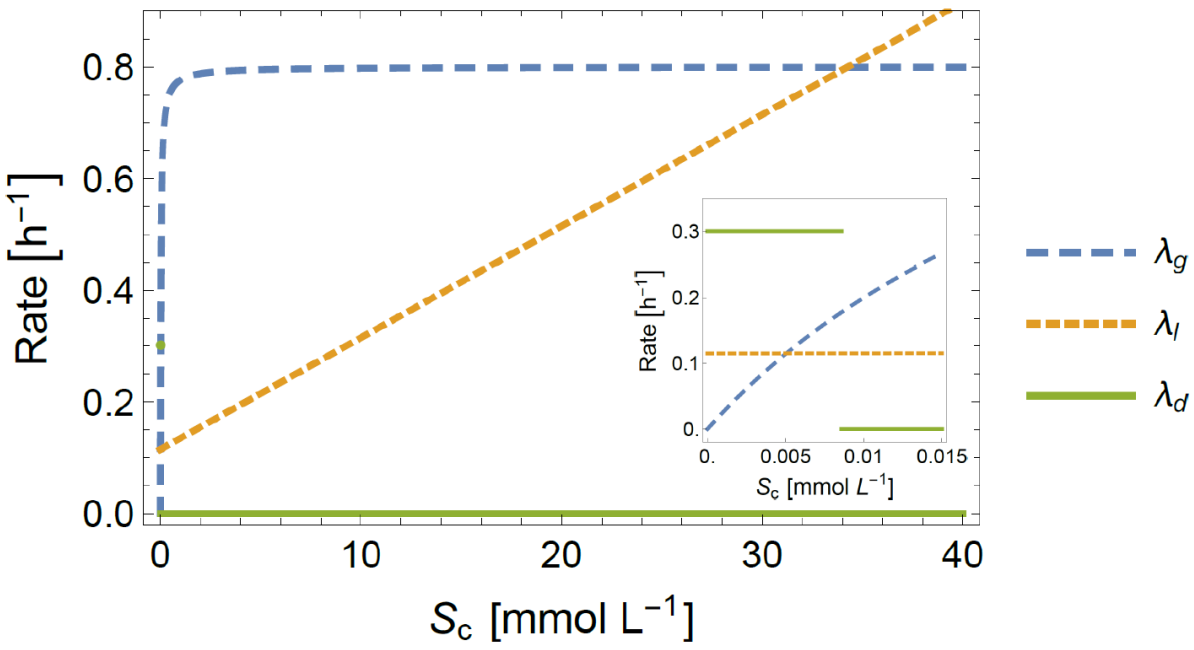

Figure 3.3 *Effect of concentration of growth limiting compound Sc on the rates of growth, λ_g, lysis, λ_l, and deactivation, λ_d, for $S_o = 0.05\ mmol\ L^{-1}$ and parameters given in Table 3.1. The inset shows the dependence of the rates on the nutrient concentration for the model parameters in Table 3.1.*

3.2.2 Lysis

Whenever lysis occurs at site *i*, the living cell at this site is removed from the system and replaced by ECM material:

$$C(i,t) \xrightarrow{\lambda_l(i,t)} E(i,t+\Delta t) \tag{3.3}$$

This simulates the bursting of the cell membrane and release of the cell material into the biofilm matrix. Since both starvation conditions and enhanced oxygen conditions were found experimentally to significantly impact the measured lysis capacity of *C. jejuni* (42), we take both compounds into account in the lysis rate of the model. In particular, we assume a simple linear relation between the substrate concentrations and the lysis rate:

$$\lambda_l(i,t) = a_c S_c(i,t) + a_o S_o(i,t) \tag{3.4}$$

Here, $a_c, a_o \geq 0$ quantify the effect of concentration of chemical compounds on the lysis rate. Figure 3.3 illustrates the increase of the lysis rate of a cell as the nutrient concentration increases at the patch occupied by the cell.

3.2.3 Deactivation associated with starvation

Deactivation occurs when the uptake rate of the growth-limiting compound falls below a certain threshold value. In a deactivation event, a live cell is removed from the system in a similar way to a lysis event. Unlike a lysis event however, no extracellular material is released through deactivation. This simulates the cell simply "shutting down" or entering a coccoid, viable but non-culturable (VBNC) state of minimal metabolic activity. Although theoretically cells may be resuscitated from the VBNC state, this process is poorly

understood and the likelihood is very low – hence, in this model, the deactivation process is irreversible. The deactivation rate is implemented as follows:

$$\lambda_{\mathrm{d}}(i,t) = r_d \Theta\big(u_{\min} - u_{\mathrm{c}}(i,t)\big), \tag{3.5}$$

where $\Theta(x)$ is the Heaviside step function, whose value is zero for $x < 0$ and one for $x > 0$. The quantity $u_c(i,t)$ is the uptake rate of compound S_{c} , implemented in the model by a Monod equation with an intrinsic maximum uptake rate, $U_{\mathrm{c}}^{\max}$, and a half saturation constant K_{c} (Eq. 3.10). According to Eq. (3.5), deactivation can only occur when the uptake rate is below a minimum value $u_{\min}$. Accordingly, deactivation is restricted to values of the nutrient concentration below a threshold, i.e., for $S_c < S_{\mathrm{c}}^{\mathrm{d}} = K_{\mathrm{c}} u_{\min} (U_c^{\max} - u_{\min})^{-1}$. Figure 3.3 illustrates the dependence of $\lambda_{\mathrm{d}}(i,t)$ on $S_{\mathrm{c}}(i,t)$ for the values of the parameters $u_{\min}$, $U_c^{\max}$ and K_{c} given in Table 3.1. In this example, the deactivation rate is positive for very low values of the nutrient concentration (for $S_{\mathrm{c}}(i,t) < S_{\mathrm{c}}^{\mathrm{d}} = 0.0086$).

3.2.4 Substrate dynamics

Changes in the concentrations of chemical compounds (i.e., oxygen and the nutrient source) occur in our model through their uptake by live cells and diffusion. The diffusion process in our model is governed by a modified Fick's second law (Eq. (3.9)), with changes introduced to include the influence of nutrient uptake by bacteria on chemical concentrations and the differences in the diffusion coefficient at different stages of biofilm development (266). The latter modification has been introduced to account for the heterogeneity of the biofilm structure, which resembles a porous material (274, 275). Diffusion in porous media is a well-known example of anomalous (non-Fickian) diffusion, in which the quadratic displacement of molecules is not proportional to time (275). Instead, anomalous diffusion corresponds to a time-dependent diffusion coefficient, $D(t)$ (275). Introducing a density dependent diffusion coefficient has been previously applied to mathematical models of biofilms (274). In our model, the diffusion coefficient at time *t* is given by the following equation:

$$D(t) = \rho_{\mathrm{B}}(t) D_{\mathrm{B}} + \big(1 - \rho_{\mathrm{B}}(t)\big) D_{\mathrm{F}} \tag{3.6}$$

In Eq. (3.6), D_{B} represents the diffusion constant through biofilm mass (i.e., when patch *i* is occupied by a cell or ECM) and D_{F} represents the diffusion constant in the fluid. The quantity $\rho_{\mathrm{B}}(t)$ is the density of the biofilm and it is calculated as follows:

$$\rho_{\mathrm{B}}(t) = \frac{\text{Area occupied by biomass}}{\text{Area occupied by biomass} + \text{No. pores}} \tag{3.7}$$

$$\text{Area occupied by biomass} = \sum_{i \in L*H} \begin{cases} 1, & \sigma(i,t) = C \\ 1 - \dfrac{n_{\mathrm{E}}(i,t)}{N_E}, & \sigma(i,t) = E \\ 0 & \sigma(i,t) = F \end{cases} \tag{3.8}$$

When patch *i* is occupied by ECM and the ECM uptake mechanism is active, $n_{\mathrm{E}}(i,t)$ gives the number of times part of the ECM has been consumed at site *i* by cells in $\mathcal{N}(i)$. The pores are defined as sites for which $\sigma(i,t) = F$, which are contained within the biofilm, i.e., no path can be drawn from the site to the biofilm boundary layer without crossing a site which is occupied by biomass. Note that if there is no biomass present, $D(t) = D_{\mathrm{F}}$, and in contrast, if the biofilm is fully compact, $D(t) = D_{\mathrm{B}}$.

The change of compound *k* (here, nutrient or oxygen) concentration at site *i* with time is given by:

$$\frac{\partial S_k(i,t)}{\partial t} = D(t)\nabla^2 S_k(i,t) - u_k(i,t), \qquad k = \mathrm{c}, \mathrm{o} \tag{3.9}$$

Here,

$$u_k(i,t) = \begin{cases} U_k^{\max} \dfrac{S_k(i,t)}{S_k(i,t) + K_k}, & \text{if } \sigma(i,t) = C \\ 0, & \text{otherwise} \end{cases}, \tag{3.10}$$

is the uptake of the *k*-th chemical compound at patch *i* if it is occupied by a cell at time *t*. In common with previous models for biofilm formation (266, 268), we assume that diffusion and nutrient uptake are fast compared to vital cell transitions. In practice, we assume that immediately after a change of the biofilm geometry or composition due to cell division, cell deactivation, or lysis , the chemical compounds reach a quasi-steady state with $\frac{\partial S_{\mathrm{c}}(i,t)}{\partial t} = \frac{\partial S_{\mathrm{o}}(i,t)}{\partial t} = 0$. After that, it is assumed that the substrate concentrations remain constant until the next cell transition event when the concentrations reach a new quasi-steady state (See Appendix A.2 for a detailed description of the implementation of the algorithm for the substrate dynamics).

Table 3.1 *Parameter values used in our simulations and corresponding references motivating the choice of the parameter values.*

Parameter	Description	Value(s)	Unit	Reference
D_B	Diffusion of the chemical compounds in	9×10^{-11} and 9×10^{-10}	m^2h^{-1}	Calibrated through numerous simulations to

	the biofilm matrix			produce conditions in which biofilm development would affect chemical concentrations for given parameters.
D_F	Diffusion of the chemical compounds in liquid	9 x 10^{-10}	m^2h^{-1}	As above.
U_o^{max}	Maximum uptake rate of oxygen	4.5 x 10^{-10}	mmol h^{-1}	Order of magnitude obtained from oxidation rates of *C. jejuni* (276).
K_o	Monod coefficient for oxygen uptake rate	3 x 10^{-3}	mmol L^{-1}	Based on saturation constant estimates for *C. jejuni* in regards to changes in dissolved oxygen concentrations (276).
U_c^{max}	Maximum uptake rate of the carbon source	4.5 x 10^{-9}	mmol h^{-1}	Assumed larger than maximum oxygen uptake, due to higher need to utilize

				the carbon source.
K_c	Monod coefficient for growth limiting compound uptake	0.03	mmol L^{-1}	Assumed larger than K_o due to higher need to utilize the carbon source.
r_g	Maximum division rate of *C. jejuni*	0.8	h^{-1}	(277) Maximum growth rate of *C. jejuni* 104 in 42ºC, in BHI medium.
K_g	Monod coefficient for growth	0.03	mmol L^{-1}	Assumed to be equal to K_c, due to the growth limiting nature of S_c and previously found direct proportionality of growth rate and substrate uptake for *E. coli* (278).
r_d	Deactivation rate	0.3	h^{-1}	(279) Approximate death rate of *C. jejuni* in water at 37ºC in stationary conditions.
u_{min}	Minimum uptake rate of S_c	1 x 10^{-9}	mmol h^{-1}	Assumption based on U_c^{max}.
a_c	Lysis coefficient for carbon	0.02	L $mmol^{-1}$ h^{-1}	Approximated to produce a slightly higher

	source concentration			lysis rate value in aerobic conditions than the maximum growth rate, and a slightly lower lysis rate value in microaerobic conditions.
a_{o}	Lysis coefficient for oxygen concentration	Oxygen sensitive: 2.3	L mmol^{-1} h^{-1}	As above.

3.3 Results

3.3.1 Behaviour of a single cell for given chemical concentrations

Before studying the predicted behaviour of a population of cells forming a biofilm, it is instructive to study the behaviour of a single cell depending on the chemical concentrations $S_c(i,t)$ and $S_o(i,t)$ at the patch i occupied by the cell. In particular, we consider the probability $P_g(i,t)$ that growth occurs before lysis or deactivation. This can be expressed in terms of the rates for growth, lysis and deactivation as follows:

$$P_g(i,t) = \frac{\lambda_g(i,t)}{\lambda_g(i,t) + \lambda_l(i,t) + \lambda_d(i,t)} \tag{3.11}$$

The dependence of P_g on the chemical concentrations at the patch can be obtained by using the dependence of the rates on the chemical concentrations given by Eqs. (3.2), (3.4) and (3.5). The colour map in Figure 3.4 shows the probability P_g for a cell as a function of chemical concentrations S_c and S_o in the patch occupied by the cell for ECM uptake set to zero, $r_E = 0$, and the model parameters given in Table 3.1. The chemical concentrations space (S_c, S_o) can be naturally split into two regions separated by a threshold line where $P_g = \frac{1}{2}$. We shall refer to this line as the growth likelihood threshold (in a probabilistic sense). Below the line, the growth probability is $P_g > \frac{1}{2}$ so that growth is more likely than deactivation and lysis. In contrast, $P_g < \frac{1}{2}$ above the line and growth is less likely than deactivation or lysis. One can show that $P_g = \frac{1}{2}$ implies that the effective rate of growth

$\lambda_g - \lambda_l - \lambda_d$ becomes zero, i.e., the rate of growth is exactly balanced by the rates of no growth at the threshold line.

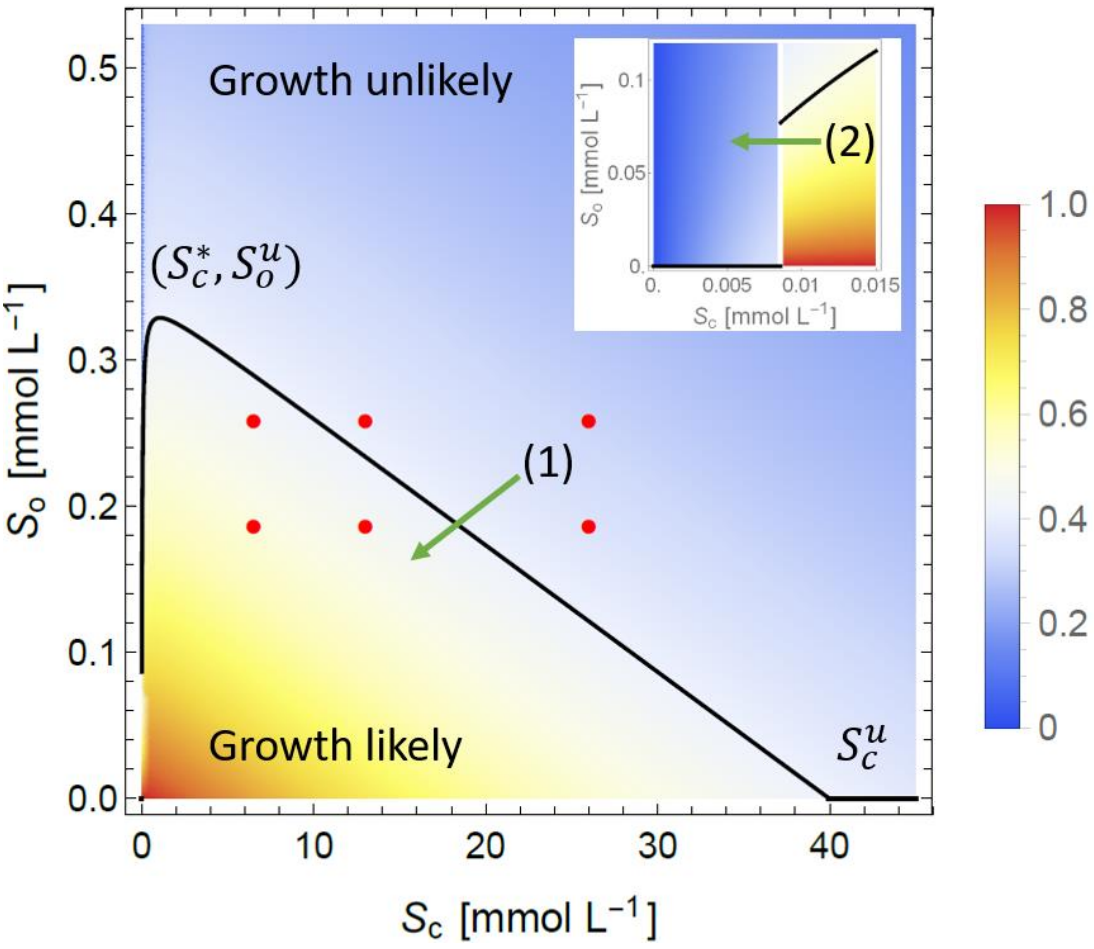

Figure 3.4 *Colour map for the probability of growth of a cell (P_g) as a function of the nutrient and oxygen concentrations at the patch occupied by the cell. The black curve corresponding to $P_g = \frac{1}{2}$ splits the space of chemical concentrations into a region where growth is more likely than lysis or deactivation ($P_g > \frac{1}{2}$, below the curve) and a region where lysis or deactivation are more likely than growth ($P_g < \frac{1}{2}$, above the curve). The value S_c^u indicates the value of the chemical nutrient concentration above which growth is unlikely for any oxygen concentration. At the point (S_c^*, S_o^u), the nutrient concentration S_o^u is optimal in the sense that growth is favoured for the widest possible range of oxygen concentrations, i.e., for $S_o \in (0, S_o^*)$. (1) indicates a cell in a patch where growth is initially unlikely but could become favourable if biofilm develops around the patch, thus leading to a local reduction of S_c and S_o as marked by the arrow. The inset shows a magnification of the behaviour for small values of the chemical concentrations. (2) indicates a cell in a patch where growth is initially favoured, but would become unlikely, as S_c and S_o reduce locally if biofilm forms around the patch. The red circles show the values of supplied chemical compounds used in the numerical simulations of Section 3.3.2.*

In the absence of ECM consumption, the oxygen concentration at the growth threshold (i.e., when $P_g(i,t) = \frac{1}{2}$) is a function of the nutrient concentration:

$$S_o = f_{th}(S_c) = \max\left[0, -\frac{1}{a_o}\left(a_c S_c - \frac{r_g S_c}{K_g + S_c} + r_d \Theta\left(S_c^d - S_c\right)\right)\right]. \qquad (3.12)$$

The equation arises from setting $P_g(i,t) = \frac{1}{2}$ in Equation (3.11) and solving for S_o. Here, the max function ensures that the oxygen concentration is non-negative at the growth threshold. For the parameter values in Table 3.1, $f_\mathrm{th}(S_\mathrm{c})$ is zero for $S_\mathrm{c} < S_\mathrm{c}^\mathrm{d}$ (see the inset of Figure 3.4). This is due to the dominance of the rate of lysis in this range of concentrations (see the inset in Figure 3.3) which makes growth essentially impossible. Increasing S_c leads to an increase of $f_\mathrm{th}(S_\mathrm{c})$ until a maximum is reached with value:

$$S_\mathrm{o}^\mathrm{u} = \frac{r_\mathrm{g}}{a_\mathrm{o}}\left(1 - \sqrt{\frac{K_\mathrm{g}a_\mathrm{c}}{r_\mathrm{g}}}\right)^2 \tag{3.13}$$

for a nutrient concentration

$$S_\mathrm{c}^* = K_\mathrm{g}\left(\sqrt{\frac{r_\mathrm{g}}{K_\mathrm{g}a_\mathrm{c}}} - 1\right). \tag{3.14}$$

This expression is valid when $S_\mathrm{c}^\mathrm{d} < S_\mathrm{c}^*$ which is indeed the case for the values $S_\mathrm{c}^* = 1.065$ and $S_\mathrm{c}^\mathrm{d} = 0.0086$ mmol L^{-1} obtained for the model parameters in Table 3.1. The prediction of a peak in $f_\mathrm{th}(S_c)$ implies that the tolerance of cell growth to oxygen is maximal for intermediate values of the nutrient concentration. For $S_\mathrm{c} < S_\mathrm{c}^*$, growth becomes possible for higher concentration of oxygen when increasing the nutrient concentration. Increasing the nutrient concentration beyond S_c^*, however, leads to a decreasing tolerance to oxygen due to the monotonic increase of the lysis rate with S_c compared to the growth rate that is limited in our model by the cell uptake capability (compare the curves for λ_g and λ_l in Figure 3.3). In fact, growth becomes unlikely (i.e., $P_\mathrm{g} < \frac{1}{2}$) for any concentration of oxygen if the nutrient concentration is above the value:

$$S_\mathrm{c}^\mathrm{u} = \frac{r_\mathrm{g}}{a_\mathrm{c}} - K_\mathrm{g}\,. \tag{3.15}$$

In particular, $S_\mathrm{c}^\mathrm{u} = 39.97$ mmol L^{-1} for the parameters given in Table 3.1 (see Figure 3.4).

To summarise, the analysis of the behaviour of one cell predicts the existence of an optimal nutrient concentration for which the effective growth rate is maximised and thus increases the survival potential of the population in higher oxygen conditions. As a consequence, our analysis suggests that lower nutrient media may increase tolerance of *C. jejuni* to higher oxygen conditions. This result may serve as a logical argument for a possible connection between *C. jejuni* being observed to grow better in lower nutrient media and its oxygen susceptibility (40, 48).

The analysis of the behaviour of a single cell also reveals the expected dependence of the probability of growth on the parameters of the vital dynamics of the model. The value of the nutrient concentration S_c^* for which the tolerance to oxygen is maximal increases with the rate of growth, r_g, and decreases with the growth half saturation coefficient, K_g, and nutrient-induced lysis rate, a_c. The maximally tolerated oxygen concentration, S_o^u, and maximal nutrient concentration for which growth is likely to occur, S_c^u, exhibit similar trends with r_g, $K_g$and a_c. In addition, S_o^u decreases with the oxygen-related lysis rate, a_o. As intuitively expected, these trends suggest that the overall chances of cell growth increase with the rate of growth and decrease with the rates of lysis. The decrease with the half-saturation K_g is perhaps less intuitive but also makes sense since larger values of K_g induce a slower increase of the total growth rate, λ_g, with the nutrient concentration (see Eq. (3.2)).

3.3.2 Collective behaviour – biofilm formation

Here, we present results on the collective behaviour of cells by means of numerical simulations performed to study the effect of substrate concentrations and diffusion of chemicals on biofilm formation. In order to clearly identify the role played by these factors, in the following sections we consider biofilms formed by cells that do not disrupt ECM. The influence of ECM disruption on biofilm formation is studied separately in the final section.

Numerical simulations are run on a rectangular environment with $L \times H$ patches, where $L = 20$ and $H = 50$. The linear size of each patch is set to $\delta x = 5$ μm, the simulation space is a rectangle of area 100×250 μm^2. All numerical simulations begin with a single triggering cell located at a randomly chosen patch, i^* at the bottom of the system; the rest of the patches are occupied by fluid. The oxygen and carbon source concentrations are kept constant throughout the entire simulation at the upper horizontal boundary of the system. Such constant concentrations are denoted as S_o^H and S_c^H to denote the fact that they correspond to patches at a height H. We assume that the concentration of chemical compounds is initially homogeneous in the system, i.e., concentrations are set to the values S_o^H and S_c^H at all the patches in the system. Following this, we shall refer to S_o^H and S_c^H as initial chemical concentrations. Values for the model parameters have been taken from the literature if available or, if they could not be directly inferred from the literature, those parameters were assigned representative values (see Table 3.1).

Effect of supplied oxygen and nutrient concentrations

Numerical simulations presented in this section consider six different conditions in terms of the supplied chemical compounds concentration, S_c^H and S_o^H, that cover the regimes of likely and unlikely growth predicted for a single cell in Section 3.3.1 (see the circles in Figure 3.4). More explicitly, simulations were run for two concentrations of oxygen: $S_{\mathrm{o}}^H = 0.188$ and 0.26 mmol L^{-1} corresponding to microaerobic and aerobic conditions in Ref. (48) respectively (approximate values from (48) were converted from ppm to mmol L^{-1} units). For each value of S_{o}^H, we explored the effect of three values of the supplied nutrient concentration, $S_c^H = 6.5, 13$ and 26 mmol L^{-1}. In particular, $S_c^H = 13$ mmol L^{-1} corresponds to the sum of measured concentrations of aspartate, glutamate, serine and proline in MHB in the experiments of Ref. (263). In order to gain statistical insight on biofilm formation, we ran 100 stochastic realisations for each pair (S_c^H, S_o^H), starting from a randomly placed cell at the bottom of the system. Simulations presented in this section are ran until the biofilm reaches the upper horizontal boundary of the simulation rectangle, the cell population is extinct, or the simulation time reaches the maximum time, set to 72h. The diffusion constant in patches occupied by the biofilm (i.e. either in *C* or *E*-state) is set to D_{B} =9x10^{-11} m^2/s which is 10 times smaller than the diffusion in the liquid (see Table 3.1 and Eq. (3.7)). Figure 3.5 shows growth curves (number of live cells, $n_{\mathrm{c}}(t)$, vs. time) for several values of the supplied chemical concentrations, S_c^H and S_o^H. We observed three possible outcomes for a given stochastic realisation: (i) The biofilm may not start spreading (for instance, the realization indicated by the blue point in Figure 3.5 (c)). (ii) The biofilm may start spreading but the population of live cells becomes extinct before they invade the system in the vertical direction (see the realisation indicated by the red curve in Figure 3.5 (a)). (iii) The biofilm may invade the system by reaching the upper boundary of the simulation rectangle at height *H* (see, for instance, the curves in Figure 3.5 (b) which stop when the system reaches the boundary).

Biofilms do not start spreading (outcome (i)) if the triggering cell undergoes lysis or deactivation before duplicating. This occurs with probability $1 - P_g(i^*, t)$, where $P_g(i^*, t)$ is the probability that the triggering cell duplicates (see Eq. (3.11)), i.e., that the biofilm starts growing. Figure 3.6 shows analytical (continuous line) and numerical estimates (squares) of the probability that the biofilm starts growing, as a function of supplied nutrient concentration, S_c^H, for oxygen concentration $S_o^H = 0.3$ mmol L^{-1}. The peak of the probability suggests that biofilm formation is more likely for intermediate values of S_c^H. Indeed, at very low values of $S_{\mathrm{c}}^{\mathrm{H}}$, deactivation can occur (for $S_{\mathrm{c}}^{\mathrm{H}} < S_{\mathrm{c}}^{\mathrm{d}}$) and the rate of growth is small (see Figure 3.3). Growth is also hindered for high enough values of S_c^H due to a relative dominance of lysis.

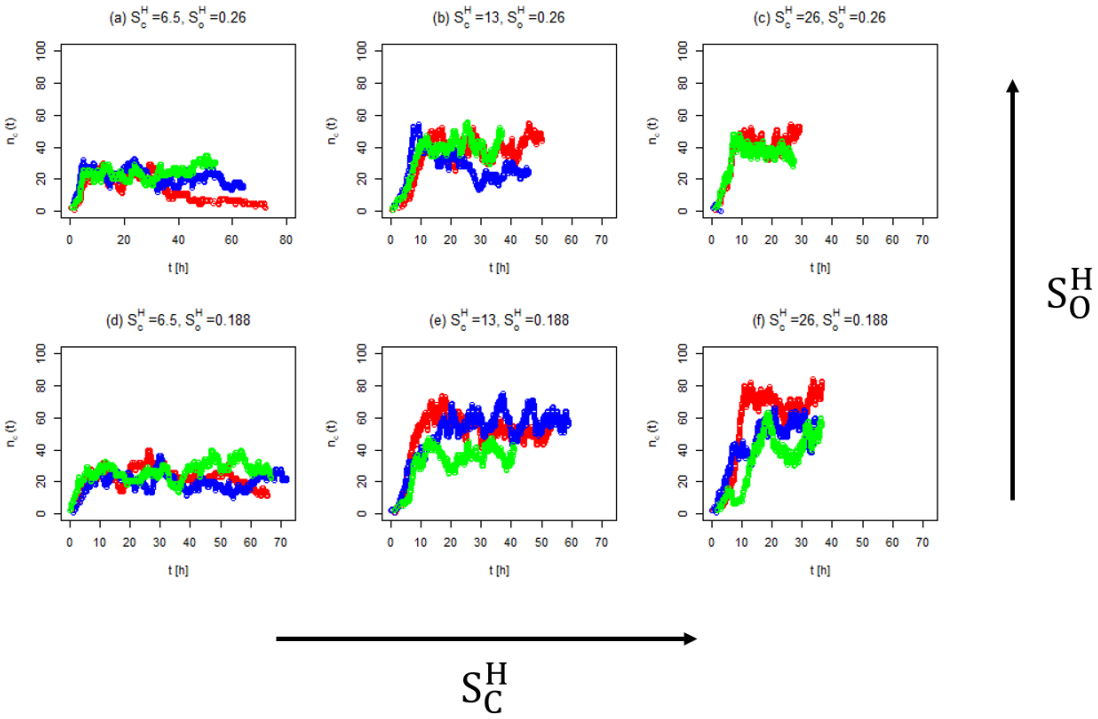

Figure 3.5 *Effect of varying the initial chemical concentrations, S_c^H and S_o^H, on the number of live cells as a function of time for tmax=72h. Three stochastic realisations are shown for each combination of chemical concentrations. Some curves end abruptly when the biofilm reaches the upper boundary of the system. The chemical concentration units are [mmol L^{-1}].*

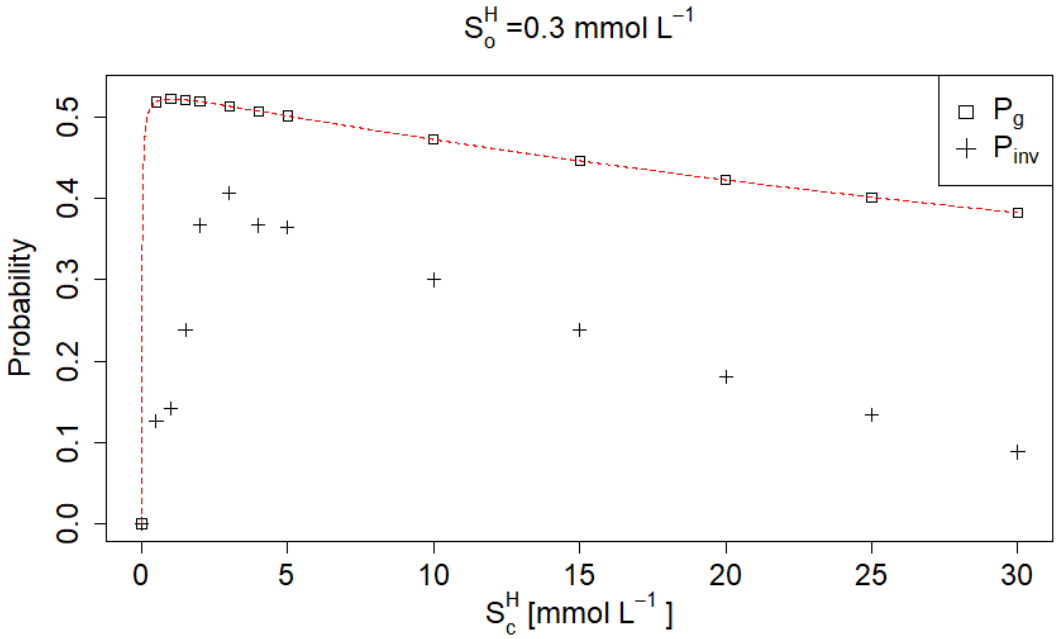

Figure 3.6 *Probability that a triggering cell in high oxygen conditions $S_o^H = 0.3\ mmol\ L^{-1}$ leads to an invasive biofilm (crosses) compared to probability of growth of a single invading cell, $P_g(i,t)$ (squares). Maximum probability is observed for relatively small nutrient concentration, $S_c^H = 3\ mmol\ L^{-1}$. Crosses indicate the survival probability of a triggering cell colony within 5h simulations, n=1000. The red dashed line is a function of $P_g(i,t)$ against S_c^H for $S_o^H = 0.3\ mmol\ L^{-1}$. The squares have been added to make the direct comparison between analytical and numerical results easier. It can be seen that although both curves have a similar shape, $P_g(i,t)$ overestimates the survival probability of the colony. This is most likely due to the fact that $P_g(i,t)$ calculates the probability that the triggering cell will duplicate before dying. As such, one of the ways in which $P_g(i,t)$ may overestimate the chances of success of the triggering cell to establish a biofilm colony, comes from the fact that for the values of S_c^H we have considered, the calculation of $P_g(i,t)$ cannot consider the effect of nutrient depletion.*

When biofilms form, they can survive and invade the system to reach the upper boundary with probability P_{inv}. This probability is reminiscent of the probability of an epidemic in a lattice which was mapped to percolation (280, 281). By definition, $P_{\mathrm{inv}} \leq P_{\mathrm{g}}$ since invasion requires duplication of the triggering cell and sustained growth of the population of live cells afterwards. For given $S_{\mathrm{o}}^{\mathrm{H}}$, however, P_{inv} depends on $S_{\mathrm{c}}^{\mathrm{H}}$ in a similar way as P_{g} in the sense that it displays a peak at intermediate values of $S_{\mathrm{c}}^{\mathrm{H}}$ (see the crosses in Figure 3.6). The maximum of P_{inv}, however, is located at a higher value of $S_{\mathrm{c}}^{\mathrm{H}}$ than that of P_{g}. In other words, the optimal nutrient concentration for biofilm invasion is higher than the nutrient concentration that optimises the effective growth rate of individual cells. This is qualitatively expected since the local nutrient concentration within a biofilm is smaller than $S_{\mathrm{c}}^{\mathrm{H}}$ due to the lower value of the diffusion constant in the biofilm compared to that in the fluid and nutrient consumed by the cells.

The diminished concentration of chemical compounds in patches occupied by cells and ECM has important consequences on the collective behaviour of invading biofilms. In particular, this may lead to large biofilm invasions in conditions that are considered unfavourable for

growth of individual cells. In order to quantify these effects, we consider the mean number of live cells at a given time, defined as $\overline{n_c} = \sum_t n_c(t)\ /T$ for each invading biofilm. Here, T is the number of time steps (i.e. vital dynamics events) in a realisation. Figure 3.7 presents relative frequency histograms for $\overline{n_c}$ for the same supplied chemical concentrations as in Figure 3.9. The histograms are bimodal, indicating a clear distinction between the situations when the colony does not invade (peak at low values of $\overline{n_c}$) and those in which the biofilm grows and spans the system. For any of the two oxygen levels considered in Figure 3.7, the peak of histograms at low $\overline{n_c}$ increases with S_c^H, meaning that invasion becomes more unlikely as S_c^H increases (i.e., P_{inv} decreases with S_c^H). On the other hand, the location of the peak on the right of the histograms moves to higher values of $\overline{n_c}$ as S_c^H increases. This implies that biofilm invasion becomes increasingly unlikely at higher nutrient concentrations but those biofilms that manage to span the system at high nutrient concentrations contain (on average) more live cells at a given time than those spanning at lower nutrient concentrations. This can be qualitatively explained in terms of the phase diagram in Figure 3.4 as a shelter effect induced by the biofilm. At high nutrient concentrations, the growth of individual cells will be unlikely (e.g., they are at the point (1) in Figure 3.4). In spite of that, if cells start duplicating and form a large enough biofilm, the chemical concentrations will decrease for some cells and their chances to duplicate increase (see the green arrow in Figure 3.4 which indicates the transition of point (1) to the region where growth is likely).

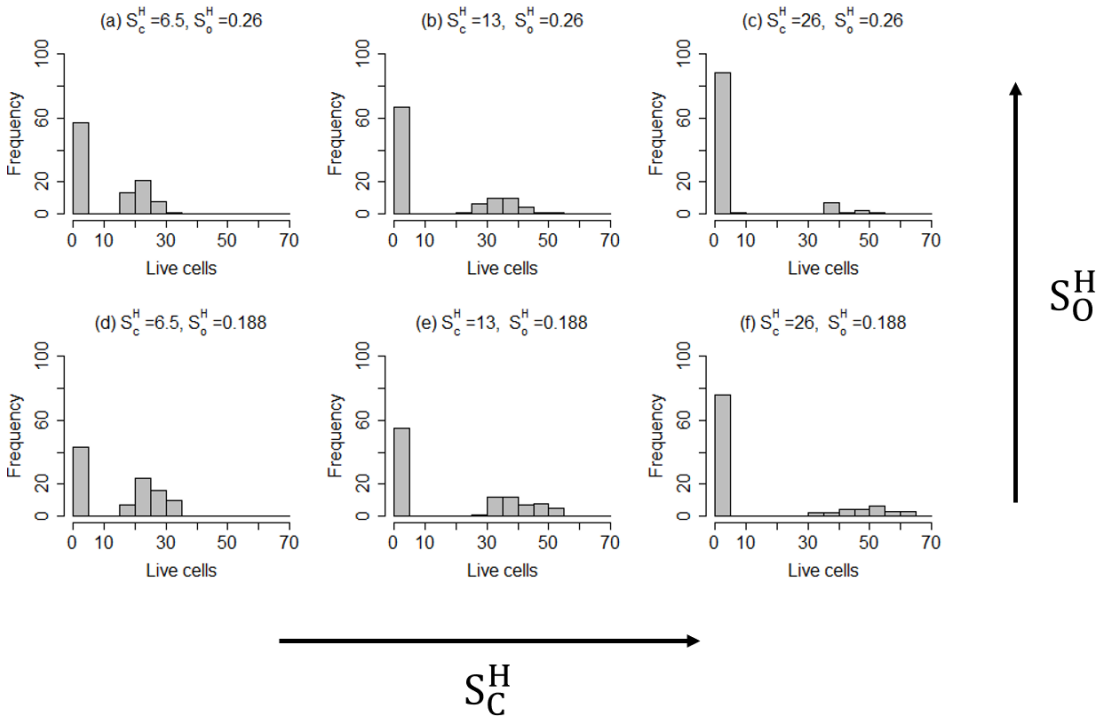

Figure 3.7 *The bar plots show the mean live cell counts from each of the 100 realisations, tmax=72h. The chemical concentration units are [mmol L^{-1}].*

The relative frequency of growth events in invasive biofilms takes values that are marginally above 1/2 for all values of supplied nutrient and oxygen concentrations considered (see Figure 3.8). For all considered nutrient concentrations, increasing the level of oxygen significantly increased the frequency of lysis and decreased the frequency of deactivation in such a way that the frequency of growth remained essentially unaffected. As a result, nearly equivalent numbers of live cells were observed to reside within the colonies stemming from a single invading cell for both low and high oxygen case, and higher density of the biofilms was obtained in the aerobic conditions (due to increase in release of extracellular material in the lytic process).

It should be recalled here that our simulations only focused on microcolonies – i.e., single cells invading the surface. In reality, there will be many cells invading the surface at the same time. Our results predict that the number of the cells in each of the invasive microcolonies could be equivalent in aerobic and microaerobic conditions, but at the same time, the invasion probability for each cell colonising the surface is lower in higher oxygen conditions. This leads to prediction of overall lower biofilm formation in aerobic conditions for the parameters specified in the model. This result is supported by experimental observations I made in Chapter 4, where less biofilm formation was observed for *C. jejuni* cultivated under aerobic conditions in MHB compared to microaerobic conditions. It is also in line with results obtained by Amy Teh et al., where it was observed that for 8 different strains of *C. jejuni*, biofilm formation in MHB was equivalent or lower in aerobic conditions compared to microaerobic conditions (48). On the other hand, the same study reported equivalent biofilm formation in Brucella broth under aerobic and microaerobic cultivation. It was suggested in the study that this could be due to lower dissolved oxygen content in

Brucella broth exposed to aerobic conditions when compared to MHB.

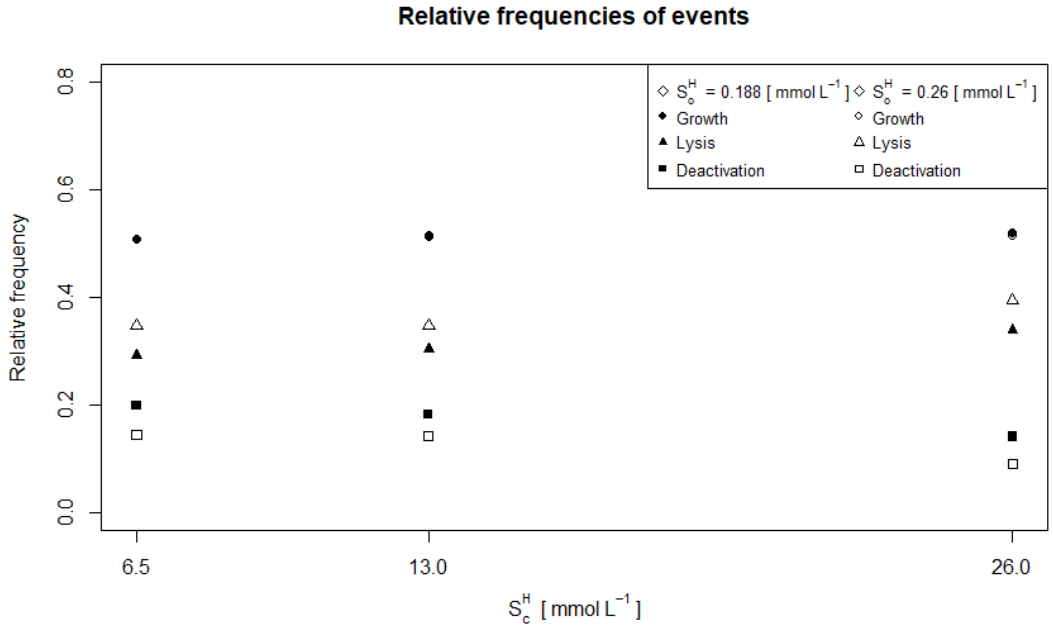

Figure 3.8 *Relative frequencies of (●,○) growth, (▲,△) lysis and (■,□) deactivation events for invasive biofilms with different initial chemical concentrations S_c^H and S_o^H. The filled shapes represent lower oxygen conditions, where S_o^H =0.188 mmol L^{-1}, and empty shapes represent cases where S_o^H =0.26 mmol L^{-1}. Not that the frequency of growth in aerobic conditions is not visible, as it was found to be nearly equivalent to the frequency of growth in microaerobic conditions for the set of parameters specified by Table 3.1.*

One can regard the relative frequency of growth events as a numerical estimate for the growth probability P_{g} introduced above. The fact that the relative frequency of growth events is slightly larger than 0.5 for invasive biofilms agrees with the requirement $P_{\mathrm{g}} > \frac{1}{2}$ for cell growth to be likely.

At the level of individual cells, the considered values of $S_{\mathrm{c}}^{\mathrm{H}}$ and $S_{\mathrm{o}}^{\mathrm{H}}$ cover conditions of likely and unlikely growth (see circles in Figure 3.4). In spite of that, invasive biofilms are not very sensitive to the specific values of $S_{\mathrm{c}}^{\mathrm{H}}$ and $S_{\mathrm{o}}^{\mathrm{H}}$. This suggests that biofilm formation leads to a self-organized state in which live cells are effectively near the $P_{\mathrm{g}} = \frac{1}{2}$ threshold line in Figure 3.4 that separates the regimes of likely and unlikely growth.

The morphology of invasive biofilms depends on the supplied chemical concentrations. For lower nutrient concentrations, invasive biofilms describe ramified patterns with lower density (compare the panels in Figure 3.9 (a) and (b) with Figure 3.9 (c)). Increasing the supply of nutrients increases the frequency of lysis events and decreases the frequency of deactivation events. As a result, the amount of ECM increases, and biofilms become denser (Figure 3.9 (c)). At low fixed values of $S_{\mathrm{c}}^{\mathrm{H}}$, increasing the supply of oxygen, $S_{\mathrm{o}}^{\mathrm{H}}$ has a

similar effect as that observed when increasing S_c^H at constant S_o^H (compare the left and right panel in Figure 3.9 (a)).

Each pattern in Figure 3.9 corresponds to the realisation in which the largest biofilm size out of 100 trials was obtained for the given set of parameters. Figure 3.10 (a) shows that on average, for those biofilms that managed to invade the system, the biofilm size (including ECM and cells) was found to increase by 5-9% in aerobic conditions compared to microaerobic conditions for lower nutrient concentrations (i.e. S_c^H=6.5 and S_c^H =13 mmol L^{-1}), and was equivalent for the higher nutrient concentration (S_c^H=26 mmol L^{-1}). Interestingly, while a similar pattern was observed for biofilm density (Figure 3.10 (b)), this was not true for all realisations, as the panel in Figure 3.9 (b) shows (note a higher number of pores within the biofilm on the right panel in Figure 3.9 (b), compared to the left panel).

The specific concentrations of supplied chemicals do not only influence the morphology of biofilms but also affect the location of live cells. For relatively low values of S_c^H and S_o^H, live cells are mostly observed at the top of the biofilm where nutrients are more abundant (see Figure 3.9 (a) and (b)). In contrast, for high levels of oxygen and nutrients, live cells are observed in deeper regions of the biofilm where the local concentrations are low enough for cell growth to be likely (Figure 3.9 (c)). In such cases, the shelter provided by the biofilm is crucial for a population of live cells to exist. Figure 3.9 suggests, however, that for the conditions considered, most cell divisions occurred at the top of the biofilm, and this is most likely due to nutrient limitations in the lower parts of the biofilm.

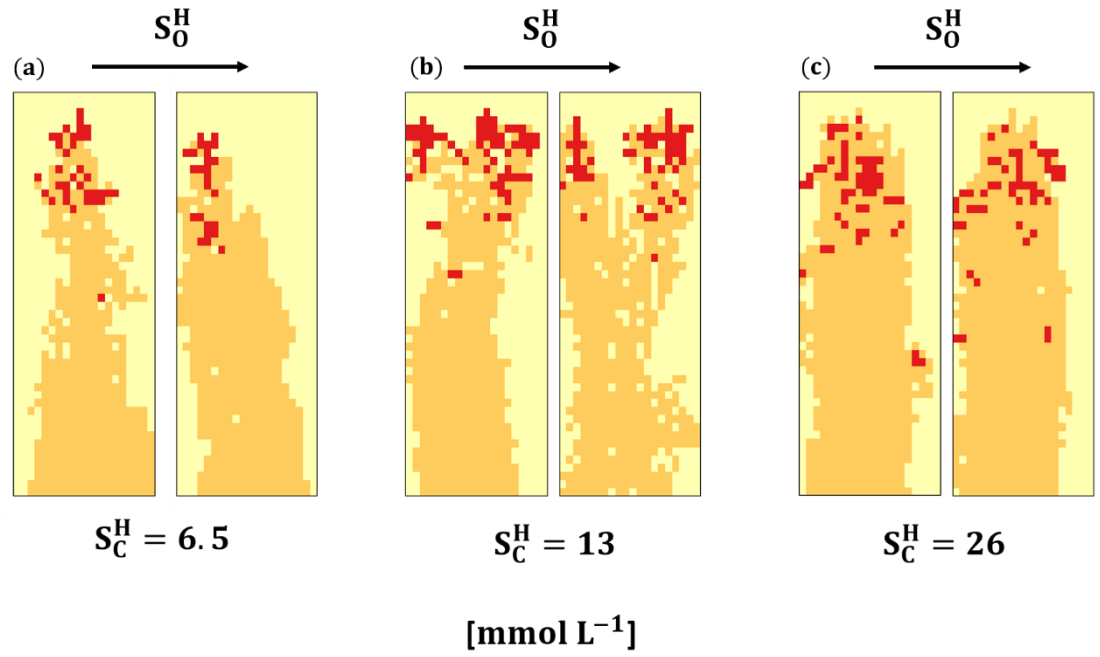

Figure 3.9 *Structure of invasive biofilms for nutrient concentrations (a) $S_c^H = 6.5\ mmol\ L^{-1}$, (b) $S_c^H = 13\ mmol\ L^{-1}$ and (c) $S_c^H = 26\ mmol\ L^{-1}$. The left and right panels for each value of S_c^H concentration correspond to low and high oxygen conditions, $S_o^H = 0.188\ mmol\ L^{-1}$ and $S_o^H = 0.26\ mmol\ L^{-1}$, respectively. Red patches represent live cells and orange patches are occupied by ECM material. All patterns correspond to the time at which biofilms span the system vertically. Furthermore, each pattern corresponds to the realisation in which the largest biofilm size out of 100 trials was obtained for the given set of parameters.*

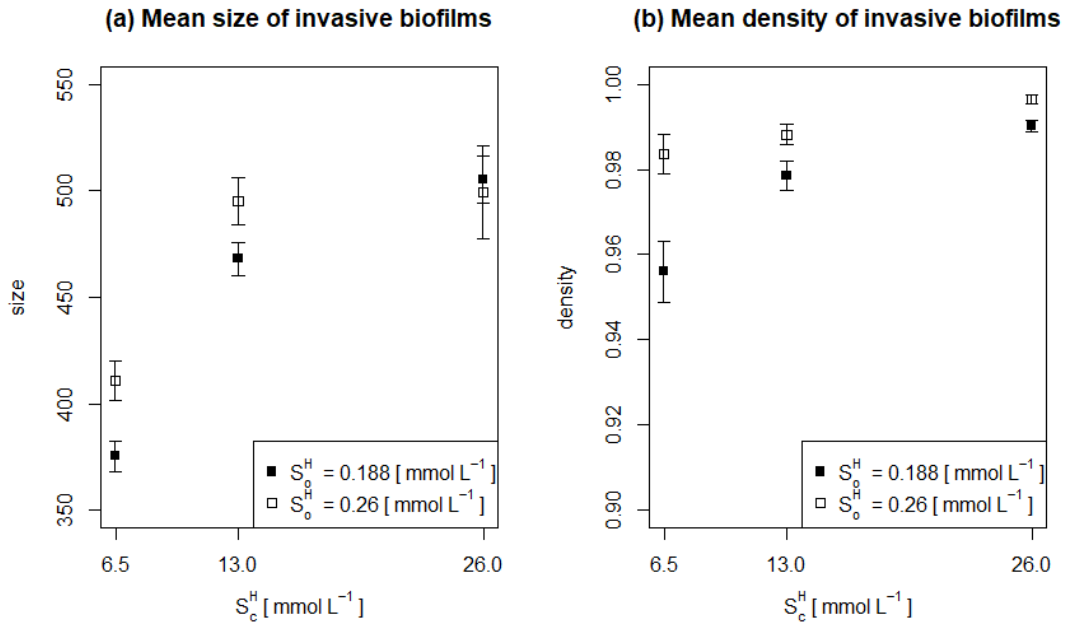

Figure 3.10 *Average size (a) and density (b) of invasive biofilms. The error bars represent the standard error of the mean.*

Effect of the diffusion coefficient in the biofilm

In order to investigate the influence of diffusion of chemical compounds in the biofilm, we ran simulations for a diffusion constant D_B =9x10^{-10} m^2/s and compared with the results obtained in the previous section, where we set D_B =9x10^{-11} m^2/s.

In the case when chemical concentrations in the media are low enough, the increase in the diffusion coefficient was found to decrease the probability of invasion, P_{inv} (from 0.57 to 0.39, for n=100), in agreement with predictions of Figure 3.4. However, for successful invasions, the mean number of live cells at a given time, $\overline{n_c}$ was found to increase by more than 5-fold (from 24 to 123). In particular, more cells were observed at deeper regions of the biofilm for high diffusion which allows the concentrations of nutrient within the biofilm to be higher compared to cases with lower diffusion (compare the patterns in panels (a) and (b) of Figure 3.11).

The effects of enhanced diffusion can be opposite in the case when the supply of nutrients and oxygen is high. Indeed, one observes high levels of nutrient and oxygen deeper inside the biofilm and this has a detrimental effect on growth due to high lysis rates. This is illustrated in Figure 3.11 where biofilms can invade if diffusion is low (panel (c)) but the invasion is unlikely for higher diffusion (see a non-invasive biofilm consisting of just ECM in panel (d)).

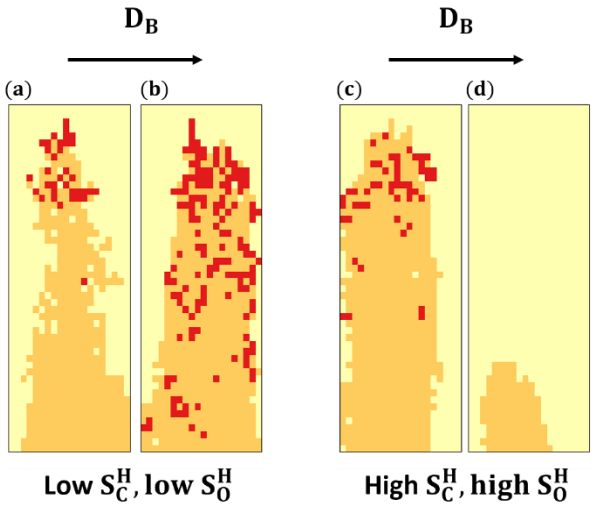

Figure 3.11 *Effect of increasing the diffusion constant D_B of nutrient and oxygen within biofilms. Panels (a) and (b) correspond to D_B =9x10^{-11} m^2/h and D_B =9x10^{-10} m^2/h, respectively, for low supply of chemical compounds ($S_c^H = 6.5\ mmol\ L^{-1}$ and $S_o^H = 0.188\ mmol\ L^{-1}$). Panels (c) and (d) show a similar arrangement for high supply of chemical compounds ($S_c^H = 26\ mmol\ L^{-1}$ and $S_o^H = 0.26\ mmol\ L^{-1}$). In all panels the realisations for which the maximum biofilm size was achieved were chosen.*

Effect of ECM disruption

The motivation for the analysis of the effect of ECM disruption was on the one hand the importance of extracellular DNA (eDNA) as a major component of *C. jejuni* extracellular biofilm matrix, and on the other, no systematic difference in biofilm formation of *C. jejuni* strains which possess genes encoding for DNAse – an enzyme known to cause eDNA disruption. Two cases were considered – disruption only with no direct benefit to the cell ($r_{\mathrm{E}} = 0$), and a disruption followed by an increase of the growth rate for the cell ($r_{\mathrm{E}} = 0.5\ h^{-1}$), to simulate potential utilization of ECM as a nutrient source.

Two distinct values for the Monod coefficient for growth, K_{g}, and varying values of the lysis coefficient for oxygen concentration, a_o, were considered. This was done to assess how our biofilm system is affected by incorporating an ECM disruption mechanism and how this may be influenced by oxygen sensitivity (measured by a_o) or by growth rate sensitivity to changing nutrient concentrations (measured by K_{g}). All other parameters, including initial nutrient and oxygen concentration values, were fixed ($S_{\mathrm{C}}^{\mathrm{H}} = 6.5$ mmol L^{-1} and $S_{\mathrm{O}}^{\mathrm{H}} = 0.26$ mmol L^{-1}). As in the previous sections, we ran 100 stochastic realisations for each value of the triple ($r_E,\ K_{\mathrm{g}},\ a_o$). In these simulations, maximum biofilm formation time was set to $t_{max} = 12$h. Figure 3.12 illustrates P_{inv} as a function of a_o in three different scenarios. The points indicated by asterisks represent cases with no ECM disruption. The black triangles represent disruption only with no growth benefit (i.e., r_E = 0), and the white triangles represent ECM disruption with growth benefit. The statistics on the invasion probability were obtained by a non-parametric bootstrap method on the vector containing information whether or not each realisation was invasive, taking on values 0 and 1, with 0's indicating unsuccessful invasions and 1's indicating successful invasions, obtained from the simulation data. 10,000 samples of size 50 were randomly chosen from the vector and the mean and standard deviation of the means were used to generate Figure 3.12 (the error bars represent twice the standard deviation). No difference was observed in invasion probability between cells exhibiting the disruption mechanism only, without the growth benefit, and the cells which did not exhibit the disruption mechanism (compare asterisks and black triangles in Figure 3.12). On the other hand, ECM uptake with growth benefit ($r_{\mathrm{E}} = 0.5\ h^{-1}$) resulted in a significantly higher invasion probability in most scenarios. The effect of K_{g} on invasion probability was found to be quite subtle – when $r_{\mathrm{E}} = 0.5\ h^{-1}$, no difference was observed between invasion probability for $K_{\mathrm{g}} = 0.03\ mmol\ L^{-1}$ and $K_{\mathrm{g}} = 6.5\ mmol\ L^{-1}$. For $r_{\mathrm{E}} = 0$ or ECM disruption mechanism inactive, however, it can be seen that for high values of a_o, the invasion probability is lower for the higher value of K_{g}. This is due to the fact that increasing K_{g} while keeping other parameters constant decreases the initial growth rate. This

in turn lowers the probability of growth, $P_g(i,t)$. As a_o increases, so does the lysis rate, which causes a further decrease in $P_g(i,t)$.

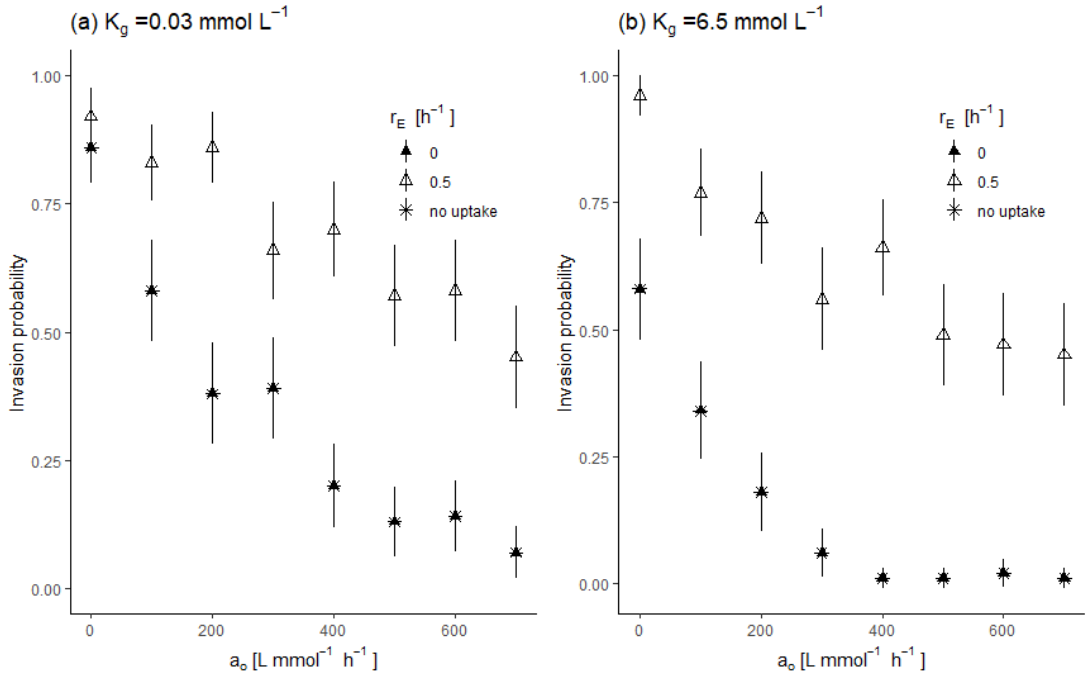

Figure 3.12 *Effect of ECM uptake on the probability that biofilms growing for 12h invade the system in n=100 realisations for (a) $K_g = 0.03\ mmol\ L^{-1}$ and (b) $K_g = 6.5\ mmol\ L^{-1}$. Standard deviation and means generated by bootstrap method using 10,000 samples of size 50 from the set of 100 realisations.*

Comparison between the number of cells in the successful invasions again did not reveal any difference between the numbers in the case when the extracellular material was not disrupted and the case when it was disrupted without growth benefit (Figure 3.13). Moreover, for the lower value of the Monod growth coefficient ($K_g = 0.03$ mmol L^{-1}), only slight, if any, increase in cell numbers was achieved when ECM disruption benefited growth (Figure 3.13 (a)). On the other hand, for higher values of the growth Monod coefficient (Figure 3.13 (b)), adding a growth benefit increased cell numbers as well as made invasion possible for higher values of a_o (i.e. $a_o \geq 400$ L mmol^{-1} h^{-1}). This result may have arisen from the fact that in our model we have assumed a direct proportionality between the Monod coefficient for growth and the Monod coefficient for limiting compound uptake, based on evidence found in literature (278). In particular, we have set $K_g = K_c$ (Table 3.1). Thus, although increasing K_g has a detrimental effect on the initial growth rates, increasing K_c in turn decreases nutrient uptake rates of bacteria, which may benefit the growth of the population, as we have in fact observed in Figure 3.13.

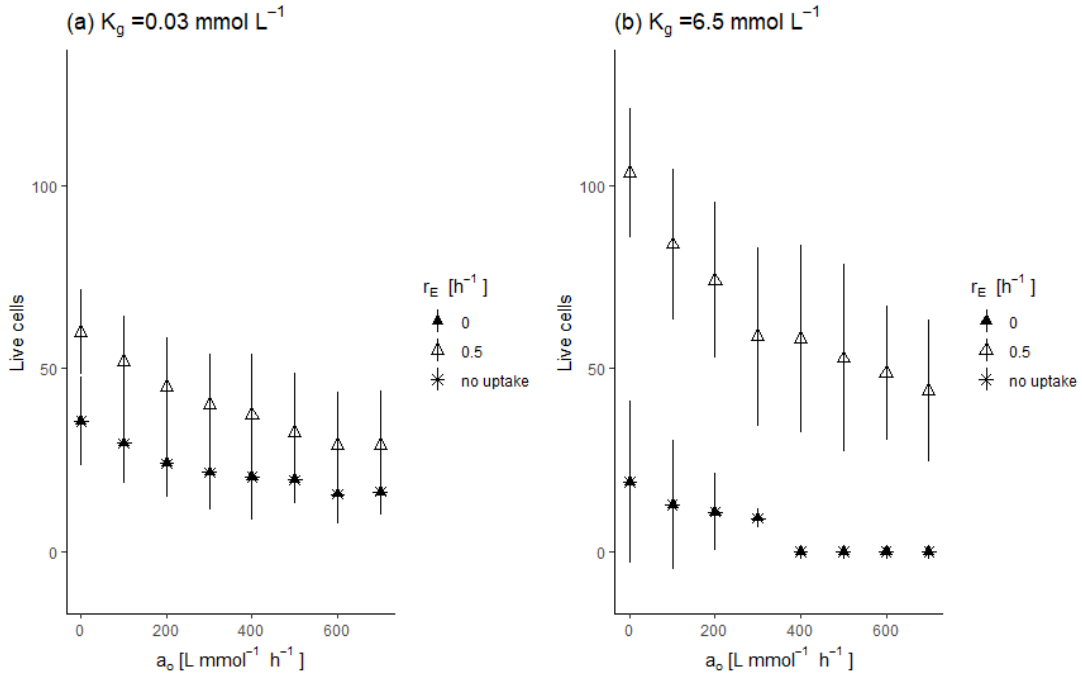

Figure 3.13 *Effect of ECM uptake on biofilm live cell counts after simulations with tmax=12h and n=100. Standard deviation and means generated by bootstrap method (10,000 samples of size 50 from the set of 100 realisations).*

Overall, studying the effect of the extracellular disruption mechanism showed that it is possible for a wide range of parameters to obtain equivalent biofilms even under oxygen stress with and without extracellular matrix disruption (i.e., for cases when disruption was active with $r_E = 0$ and when disruption was inactive). As was mentioned in the introduction, one of the puzzling properties of *C. jejuni* is that a large proportion of *C. jejuni* strains possesses one of the three genes which encodes for a biofilm matrix disruption enzyme, DNAse (264), and there seems to be no association between biofilm forming ability and the possession of these genes across *C. jejuni* phylogeny (265). Through our result, the observed lack of difference in the biofilm forming ability of strains which encode for a biofilm matrix disruption enzyme and those which do not may be explained.

It should be noted that DNAse has been shown to disrupt biofilms of some strains (42, 51, 264), and mutation in a DNAse encoding gene has been shown to restore biofilm forming ability of a poor biofilm former, RM1221 strain (264). It therefore seems that biofilm formation of some *C. jejuni* strains may in fact be inhibited by the extracellular matrix disruption. A question which remains to be answered is what factors determine whether ECM disruption is beneficial or detrimental to biofilm formation of *C. jejuni*.

Finally, as it has been observed that for some strains, ECM disruption does inhibit biofilm formation, it should be possible to uncover such regimes in this model. Perhaps a change in the diffusion coefficient or more aggressive disruption would change the invasion outcomes.

3.4 Summary

The model presented here is an individual-based, stochastic cellular automaton, in which events such as cell division and death are governed by a Poisson process with mean local event rates affected by changes in concentrations of carbon sources and oxygen. The results obtained from our analysis can be summarized as follows:

- The analysis of the behaviour of one cell predicts existence of the optimal nutrient concentration for which the effective growth rate is maximised and thus it increases the survival potential of the population in higher oxygen conditions. This in turn suggests that lower nutrient media may increase tolerance of *C. jejuni* to higher oxygen conditions.
- Development of a biofilm, through decreasing local compound concentrations within its boundaries, may push the population towards the regime in which the effective growth rate is positive.
- The model predicts that the survival probability of individual cells placed on a surface decreases when nutrient or oxygen levels are raised. On the other hand, however, the means of live cell numbers in situations when the cell is successful in establishing a colony tend to be lower in the lower nutrient conditions, i.e., there appears to be a trade-off, where on the one hand, increasing nutrient conditions may be detrimental to individual cells, but at the same time it may bring benefit to the whole population.
- Our results suggest that lower biofilm formation observed in aerobic conditions in Mueller Hinton Broth compared to cultivation in microaerobic conditions may be due to lower surface invasion probability of individual cells, and that those cells which manage to invade may generate microcolonies of equivalent size in aerobic and microaerobic conditions.
- Studying the effect of the extracellular disruption mechanism showed that it is possible for a wide range of parameters to obtain equivalent biofilms even under oxygen stress with and without extracellular matrix disruption.

Our analysis focused on comparison of the patterns found in experimental observations to those generated by the model. It should be noted, however, that the model presented here is very versatile and the analysis conducted so far does not fully address the relative contributions of all possible variables implemented in the model which govern the system. It is possible that additional interesting predictions may be obtained with time.
The model could be improved by increasing the number of dimensions to 3 or increasing the system space in order to study more mature biofilms. Furthermore, incorporation of a

metabolic network reconstruction has been successfully incorporated in a biofilm model of *Eschierichia coli* (172). Once enough information about *C. jejuni* metabolism is revealed, additional mechanisms such as release of metabolic products could be incorporated into this model and the effect of these mechanisms on the biofilm population could be assessed.

Chapter 4 - The Influence of Media and Atmospheric Conditions on *Campylobacter jejuni* Biofilm Formation

4.1 Introduction

Campylobacter jejuni is a microaerophilic bacterial species and its survival under atmospheric conditions has been frequently attributed to protective properties of biofilms which may harbour them (27, 38, 51). It has been suggested that the survival of *C. jejuni* in the environment may be largely due to its attachment to biofilms formed by other organisms, including other pathogens associated with the poultry industry (263). The environment is naturally filled with microbial life and it is not a surprise that biofilms encountered in natural settings contain mixed species in most cases. Furthermore, relatively fastidious growth requirements of *C. jejuni*, such as its inability to grow below 30ºC (26), indicate that *C. jejuni* may in many cases survive solely by passive attachment to surfaces or biofilms of other species, which is not followed by growth until suitable conditions arise (263). Literature reveals that *C. jejuni* is quite successful in surviving such 'bad times', for example, there is evidence of *C. jejuni* surviving in 4ºC distilled water in a culturable state for as long as 20 days (282).

Understanding growth requirements, either as biofilms or in planktonic cultures, of *C. jejuni*, is crucial in establishing control over this pathogen. On one hand, it may aid culturing this pathogen in laboratories, as this has been reported to be a fastidious task (50, 262), and on the other, it may help gain more understanding about *C. jejuni* prevalence both inside and outside the host. Inside the host, there is some evidence on the impact of *C. jejuni* metabolism on its virulence and ability to colonise the host, and it has been suggested that further investigations of *C. jejuni* metabolism may aid the development of drugs against *C. jejuni* (29). Outside the host, *C. jejuni* has been reported to prevail by attachment to surfaces in forms of biofilms, for example in watering supplies and plumbing systems of farms and processing plants, and nutritional properties of the media have been reported to influence the ability of *C. jejuni* to form biofilms on abiotic surfaces (40).

Although a consensus on the absence of growth of *C. jejuni* at or below ambient temperatures seem to have been reached, the survival and growth through biofilm formation of *C. jejuni* under aerobic stress varies among different reports. In some cases, it has been reported that exposing *C. jejuni* to aerobic conditions (albeit under optimal temperatures) may enhance biofilm formation (42, 47, 48), while in others, it has been shown to have an inhibitory effect (40, 48). Whether incubation in aerobic conditions enhances or inhibits biofilm formation has been suggested to depend on the particular strain and on the media. For example, levels of dissolved oxygen in various media have been suggested to play a role

in the observed results, as strains which exhibited lower or equivalent biofilm formation in microaerobic conditions when cultivated in Mueller Hinton Broth, formed higher or equivalent biofilms in aerobic conditions when cultivated in Bolton broth, which was also shown to have a significantly lower level of dissolved oxygen (48).

There have been various reports on the effect of nutrient levels on biofilm formation of *C. jejuni*. Feng at al. (2018) reported that starvation conditions inhibit biofilm formation (i.e. incubation in PBS), while Reeser et al. and Teh et al. reported higher biofilm formation in lower nutrient MHB compared to higher nutrient NB2 or Bolton and Brucella both (40, 49). On the other hand, analysis of *C. jejuni* proliferation in planktonic state revealed higher cell concentrations in highly nutritious Triptone Soya broth with 0.6% yeast extract (TSBYE) compared to NB2 and MHB (50).

As the composition of the different types of media varies substantially, other factors than the level of nutrient may be responsible for the observed differences in biofilm formation. To systematically test how nutrient levels may affect biofilm formation, *C. jejuni* ATCC 33291 cultures were incubated in dilutions of three broths (MHB, NB2 and TSBYE) with PBS, at 0.2, 0.4, 0.6, 0.8, regular concentrations and double concentrations. In separate experiments, the effect of glucose and yeast extract on biofilm formation was determined in order to test how some of the broth components may affect biofilm formation. The experiments were performed under both aerobic and microaerobic conditions.

4.2 Materials and methods

Bacterial strains and growth conditions

Campylobacter jejuni ATCC 33291 (Oxoid, VIC, Australia) was used in all experiments performed in this study. The strain was maintained at -80°C in Mueller Hinton Broth (Oxoid, VIC, Australia) and 15% glycerol. Cells were resuscitated by incubation on *Campylobacter* blood-free selective agar base (Oxoid, VIC, Australia) for 48h at 37ºC in microaerobic conditions (5% O_2, 10% CO_2, 85% N_2) generated by Campygen gas packs (Oxoid, VIC, Australia) in air-tight containers. For biofilm formation assays, cells harvested from the plates with a sterile loop were suspended in 5ml of PBS, which was used as inoculum for biofilm formation assay. Viable cell counts determined the size of the inoculums to be approx. 10^7-10^8 CFU/ml for all experiments.

4.2.1 Effect of nutrient levels on biofilm formation

Crystal violet biofilm assays

A 180µl of each medium (MHB, TSBYE, NB2 – Oxoid, VIC, Australia) at dilutions 0 (PBS), 20%, 40%, 60%, 80% and at regular concentration was transferred to a microtiter 96-

well plate. In a separate plate, 180µl of regularly concentrated TSB and double concentrated TSB, MHB and NB2 were transferred to individual wells. 20µl of the inoculum was inserted in each well containing the media. The plates were incubated for 2 days at 37ºC under either aerobic or microaerobic conditions, without shaking. Uninoculated media dilutions were used as negative control. The optical density of the biofilm samples was measured using the commonly used crystal violet assay with detailed method described in (48), at 550nm. The optical density measurements were performed in triplicate, and three independent experiments were conducted.

4.2.2 Effect of media components on biofilm formation

Crystal violet biofilm assays

MHB and NB2 basal media were supplemented with either 0.6% yeast extract (to produce MHBYE and NB2YE) or 0.25% glucose. TSB was also prepared, this time without yeast extract. The absorbance measurements were obtained analogously to the method presented above.

Viable cells in biofilm supernatants

Two-day biofilms were cultivated in MHB, NB2 or TSBYE as described in the crystal violet biofilm assays method. Following incubation, viable cell counting was performed by sampling 10 µl from the biofilm supernatants and serial dilution, followed by spread plating on Campylobacter blood-free selective agar (CCDA – Oxoid, VIC, Australia) and incubation for two days at 37ºC in microaerobic conditions.

Statistical analysis

All experiments were performed in triplicate, with three independent experiments, giving nine data points in total for each condition tested. A three-way analysis of variance (ANOVA) with Tukey's post-hoc test was performed to determine the effect of different type of media, their dilutions, and atmospheric conditions on biofilm formation. A two tailed Student's t-test was performed to compare the biofilm formation between two sets of conditions. The statistical significance was determined at a 95% confidence level using RStudio.

4.3 Results and discussion

4.3.1 Biofilm formation under different conditions

The composition of media used in this assay, as per manufacturer documentation, is summarised in Table 4.1. Results obtained from ANOVA on all data showed that in general, oxygen, media, and concentration of the media all significantly affect biofilm formation of

C. jejuni ATCC33291 ($p<0.01$), with the type of media and oxygen conditions being the highest contributing factors. Results obtained from Tukey's post hoc test showed that in general, biofilm formation was significantly higher ($p < 0.01$) in TSBYE broth as compared to MHB and NB2 broths. Furthermore, it also showed that biofilm formation was generally higher in Mueller Hinton Broth ($p<0.01$) than in NB2.

Table 4.1 Composition of media used in the biofilm assays [g/l].

	Mueller Hinton Broth (MHB)	Triptone Soya Broth (TSB)	Nutrient Broth no. 2 (NB2)
Beef dehydrated infusion	17.5	-	-
Casein hydrolysate	17.5	-	-
Starch	1.5	-	-
Enzymatic digest of soya bean	-	3.0	-
Pancreatic digest of casein	-	17.0	-
Sodium chloride	-	5.0	5.0
Dipotassium hydrogen phosphate	-	2.5	-
Glucose	0 or 2.5 for MHBG	2.5	0 or 2.5 for MHBG
Yeast extract	0 or 6.0 for MHBYE	0 or 6.0 for TSBYE	0 or 6.0 for NB2YE

Lab-Lemco powder (meat extract)	-	-	10
Peptone	-	-	10

Both TSBYE and NB2 are high nutrient media, while MHB is characterised as a low nutrient media. In the assays, TSBYE was the best performing medium, achieving highest levels of biofilm formation ($p<0.01$) and highest numbers of viable cells in biofilm supernatants in microaerobic conditions ($p<0.01$, Figure 4.1). Biofilm formation was second highest in MHB and then NB2, although equivalent numbers of cells were measured in the supernatants of these biofilms (Figure 4.1). This is similar to findings of Moore et al. (50), where TSBYE was also found to produce the highest proliferation of *C. jejuni* cells relative to MHB and NB2, although in that assay NB2 resulted in higher proliferation of cells than MHB. The authors suggested that higher proliferation of cells in TSBYE compared to other assessed broths may be due to this media having a higher concentration of peptones, proteins and meat infusion (50). On the other hand, greater biofilm formation in MHB compared to NB2 is consistent with findings obtained by Teh et al. (49). The fact that highest biofilm formation was obtained in a highly nutritious medium (TSBYE) confounds previous postulates of low nutrient media enhancing biofilm formation (40, 48, 49). Instead, it seems rather that the extent of biofilm formation is affected by the type of nutrient more than by the level of nutrients, as will be further discussed in the following section.

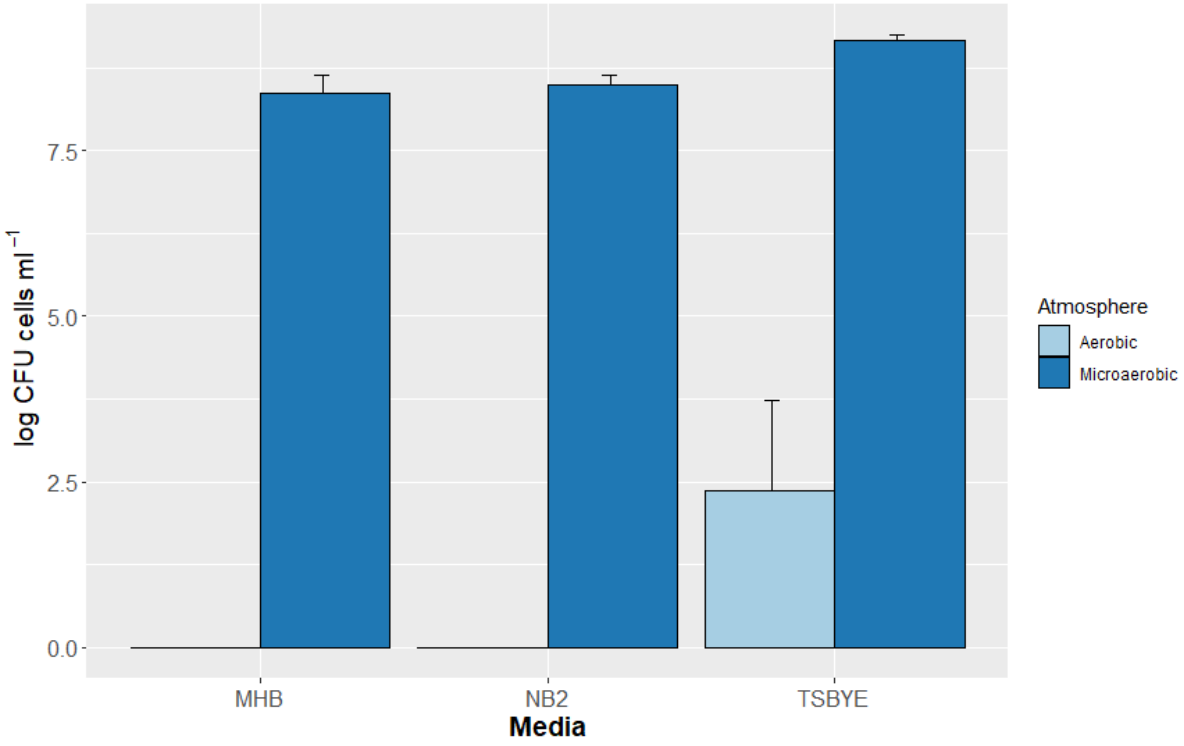

Figure 4.1 *Viable cell counts of biofilm supernatants. The means were obtained from triplicate measurements of three independent experiments and the error bars represent two-sided 95% Confidence Intervals obtained from the t-distribution. The dark blue bars represent biofilm cultivated in microaerobic conditions and light blue bars represent biofilm cultivated in aerobic conditions.*

Regarding atmospheric conditions, Tukey's post hoc test revealed significantly higher biofilm formation under microaerobic conditions compared to aerobic conditions ($p<0.01$). In fact, for cultures cultivated aerobically in MHB and NB2, the 95% CI interval of the mean contains 0, which indicates there may have been no significant biofilm formation under these conditions. Consistent with the absorbance measurements, no viable cells were detected in the biofilm supernatants of aerobic cultures in MHB and NB2 media, while a small number of viable cells (mean of approximately 10^3 CFU/ml) were found in aerobically incubated TSBYE cultures (Figure 4.1). Even the relatively low concentrations of 10^3 CFU/ml can be deemed as considerable, as ingestion of as little as 500 *C. jejuni* organisms has been shown to cause disease (8). The fact that TSBYE was the only broth for which viable *C. jejuni* cells were recovered after aerobic cultivation is likely to partially result from higher growth observed in these conditions overall, which is demonstrated by higher concentration of cells in microaerobic conditions compared to other broths (Figure 4.1).

4.3.2 Effect of nutrient levels on biofilm formation

The effect of changing concentration of specific media on biofilm formation is illustrated in Figure 4.2 and Figure 4.3. Although the broth dilutions were recognised as significant factor ($p<0.01$) in the ANOVA analysis, the type of media used, and atmospheric conditions were

indicated to be of more importance. Tukey's post hoc test revealed that in general, the only significant differences ($p<0.05$) could be detected between higher dilutions (from 60% to double concentration) and the lowest dilution (20%). Namely, for a wide range of broth dilutions, equivalent absorbance measurements were obtained. More prominent differences were observed between different media used. These results suggest that it is unlikely that low level of nutrients supports biofilm formation. Rather, it seems that specific broth properties, for example, the type of nutrients they have, may play a bigger role. That being said, increasing nutrient levels to higher than recommended (i.e., double concentration) was found to decrease biofilm formation in microaerobic conditions for double concentration of MHB (MHBx2, $p<0.01$), with p-values determined with a two-sided Welch t-test. On the other hand, there was no significant difference observed between the means of results obtained from double concentrated TSB and NB2 compared to their regular concentrations (Figure 4.3).

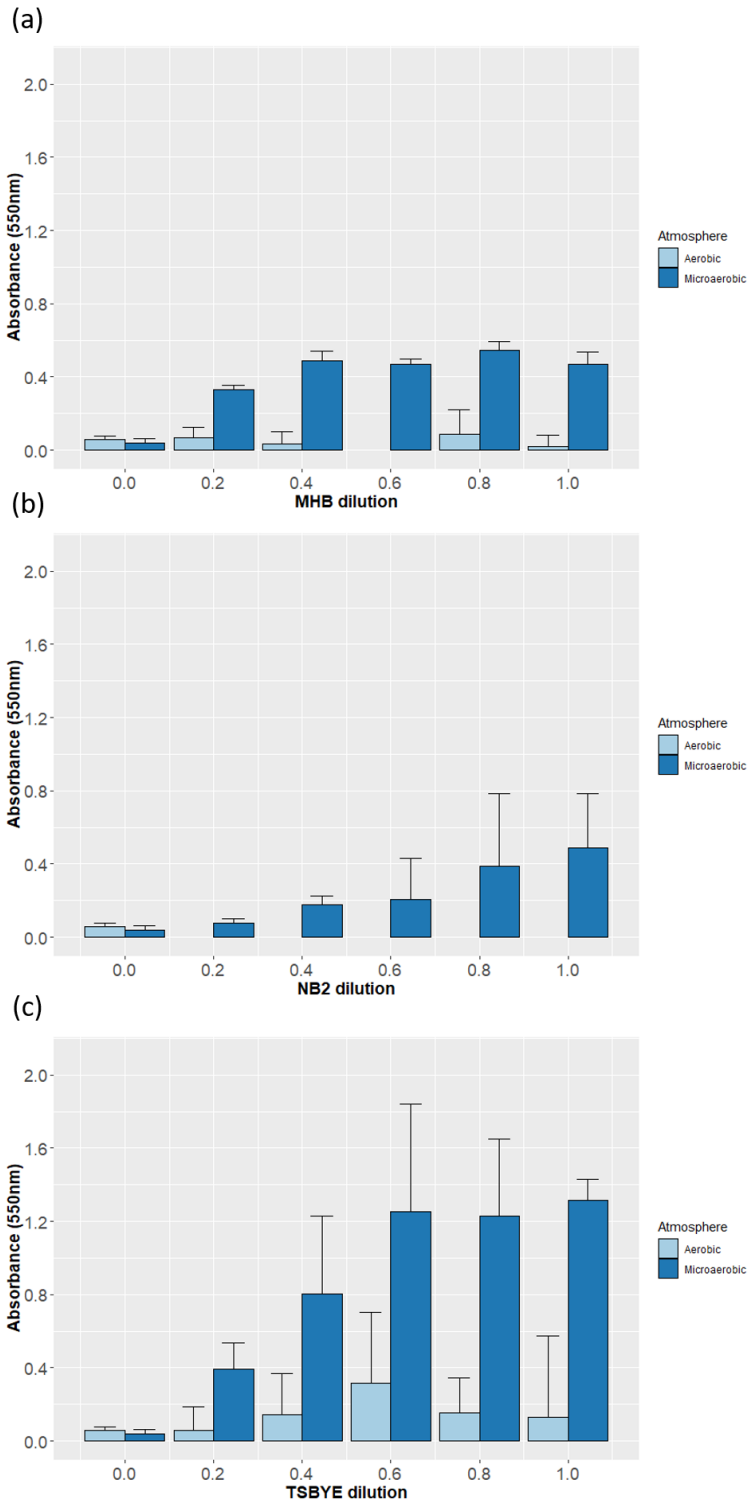

Figure 4.2 *Absorbance of biofilms measured by crystal violet assay for (A) diluted media of MHB, (B) diluted media of NB2, (C) diluted media of TSBYE. The absorbance was calculated by the difference between absorbance measured in biofilm cultures and the absorbance measured for negative controls. The means were obtained from triplicate measurements of three independent experiments and the error bars represent two-sided 95% Confidence Intervals obtained from the t-distribution. The dark blue bars represent biofilm cultivated in microaerobic conditions and light blue bars represent biofilm cultivated in aerobic conditions.*

Since the relatively low-nutrient medium, MHB, was the only one which exhibited inhibited biofilm formation at a double concentration, this suggests that MHB may contain a compound which inhibits *C. jejuni* ATCC33291 growth in higher concentrations. MHB is a relatively simple broth, containing only three ingredients, namely beef infusion, casein hydrolysate and starch (Table 4.1). Beef infusion and casein hydrolysate are unlikely to inhibit *C. jejuni* proliferation (50). On the other hand, starch is already used as an ingredient of edible coatings in the food industry, partly due to its antimicrobial properties (283). Furthermore, although inclusion of a small amount of soluble starch has been shown in another study to promote growth of *C. jejuni*, the same study reported that higher concentrations of starch inhibited its growth (284). These considerations suggest that most likely starch may be the reason for less biofilm formation observed in MHBx2.

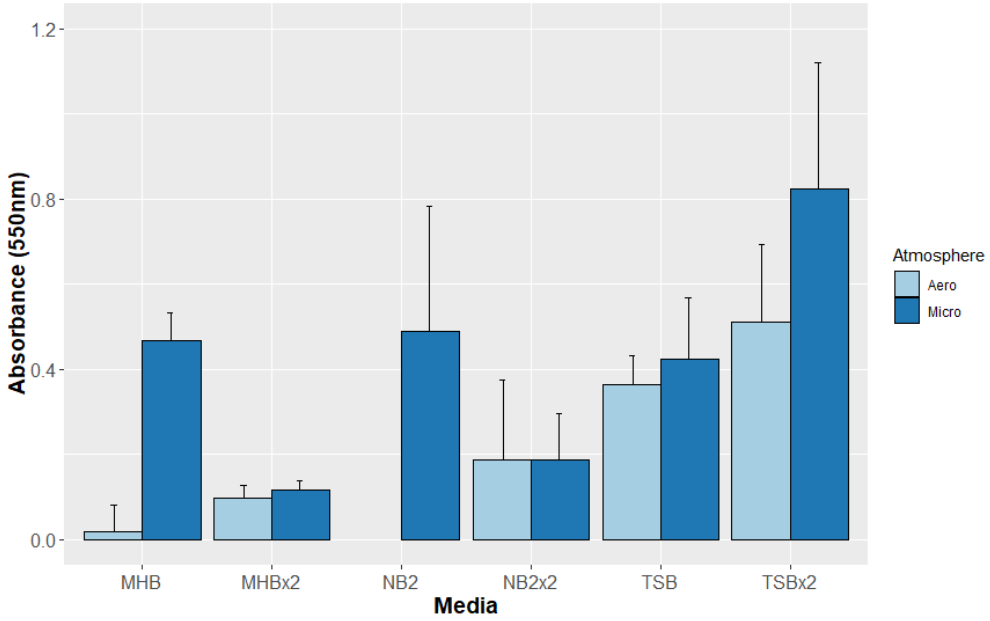

Figure 4.3 *Absorbance of biofilms measured by crystal violet assay for regularly concentrated (MHB, NB2, TSB) and double concentrated (MHBx2, NB2x2, TSBx2) broths. The absorbance was calculated by the difference between absorbance measured in biofilm cultures and the absorbance measured for negative controls. The means were obtained from triplicate measurements of three independent experiments and the error bars represent two-sided 95% Confidence Intervals obtained from the t-distribution. The dark blue bars represent biofilm cultivated in microaerobic conditions and light blue bars represent biofilm cultivated in aerobic conditions.*

4.3.3 Effect of media components on biofilm formation

The addition of 0.6% yeast extract to MHB or NB2 resulted in inhibited biofilm formation in microaerobic conditions ($p<0.01$ for MHB and $p<0.05$ for NB2). On the other hand, TSB supplementation with yeast extract (TSBYE) was found to enhance biofilm formation compared to those grown in the regular TSB medium ($p<0.01$). The results are illustrated in Figure 4.4.

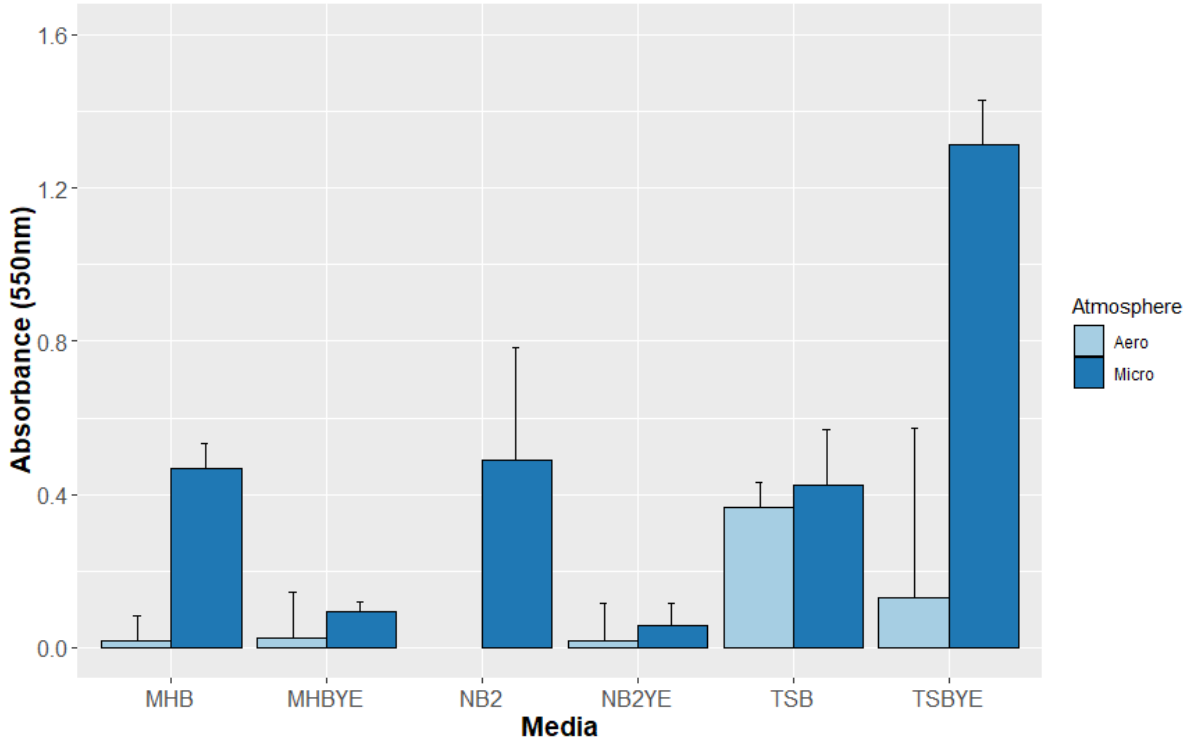

Figure 4.4 *Absorbance of biofilms measured by crystal violet assay for regular (MHB, NB2, TSB) and yeast extract supplemented (MHBYE, NB2YE, TSBYE) broths. The absorbance was calculated by the difference between absorbance measured in biofilm cultures and the absorbance measured for negative controls. The means were obtained from triplicate measurements of three independent experiments and the error bars represent two-sided 95% Confidence Intervals obtained from the t-distribution. The dark blue bars represent biofilm cultivated in microaerobic conditions and light blue bars represent biofilm cultivated in aerobic conditions.*

Yeast extract is a source of nitrogen, amino acids, carbon, and vitamins which may potentially serve as nutrient sources for *C. jejuni*. However, it has been previously shown in an assay of 13 strains, that *C. jejuni* possibly does not grow in yeast extract alone. On the other hand, when yeast extract was combined with casamino acids, growth was observed, although no growth was detected after cultivation in casamino acids alone (284). This suggests that these two ingredients complement each other in terms of growth requirements of *C. jejuni.* On the other hand, in the same study, addition of beef extract to the casamino acids and yeast extract mix inhibited growth of *C. jejuni*, while the cells grew well in beef extract with casamino acids (284). Therefore, it seems that the interaction between ingredients of the growth medium is of significant importance in determining the proliferation of *C. jejuni.* It may be that in TSBYE, addition of yeast extract complements the nutrient mixture, while in MHBYE or NB2YE, the conditions become detrimental to the cells.

TSB already contains glucose in its composition, therefore, for the effect of glucose on biofilm formation, we only considered MHB and NB2 supplemented with 0.25% glucose

(MHBG and NB2G, respectively). We found that the addition of glucose to MHB significantly inhibited biofilm formation in microaerobic conditions ($p<0.01$), which is consistent with findings of Reeser et al., who also reported that glucose decreased biofilm formation of strain M129 cultured in MHB (40). On the other hand, there was no significant difference detected between biofilm formation in NB2 and NB2G in microaerobic conditions, and in aerobic conditions, the addition of glucose to NB2 enhanced biofilm formation ($p<0.01$) compared to cultivation in regular NB2 broth in aerobic conditions (Figure 4.5).

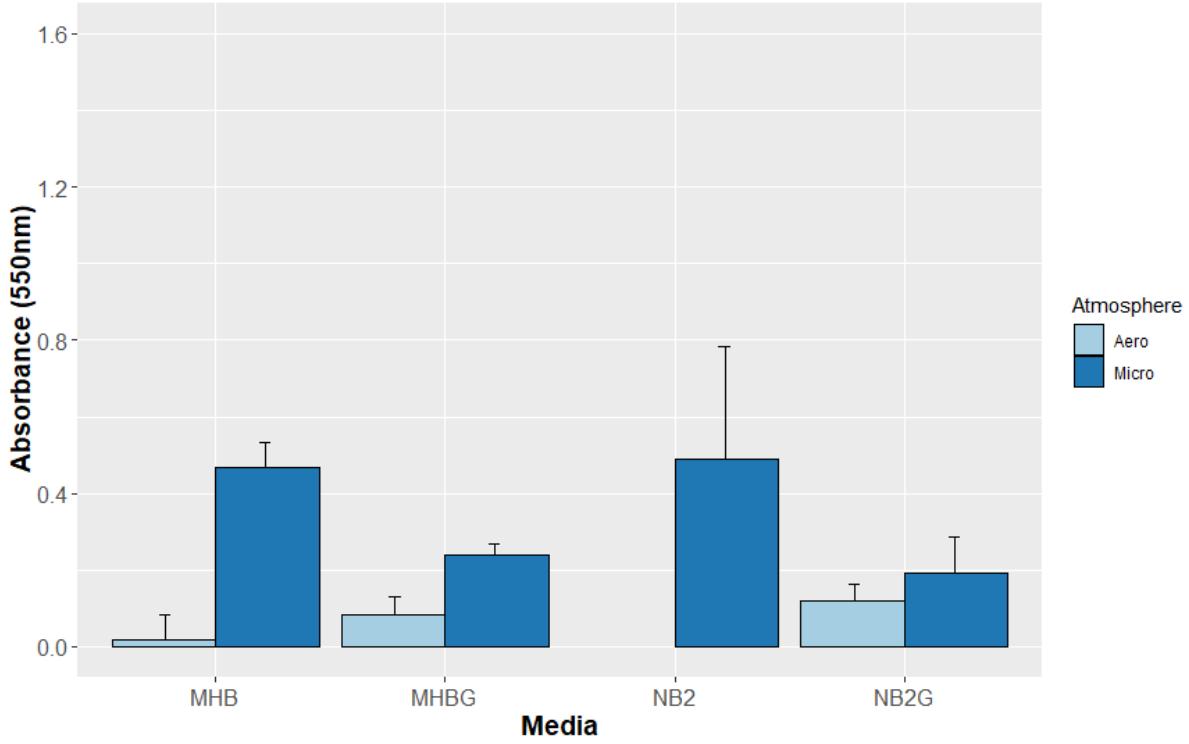

Figure 4.5 *Absorbance of biofilms measured by crystal violet assay for regular (MHB, NB2) and glucose supplemented (MHBG, NB2G) broths. The absorbance was calculated by the difference between absorbance measured in biofilm cultures and the absorbance measured for negative controls. The means were obtained from triplicate measurements of three independent experiments and the error bars represent two-sided 95% Confidence Intervals obtained from the t-distribution. The dark blue bars represent biofilm cultivated in microaerobic conditions and light blue bars represent biofilm cultivated in aerobic conditions.*

Reeser et al. attributed the inhibitory effect of glucose on biofilm formation to its characteristics as an osmolyte. Addition of other osmolytes (NaCl or sucrose) individually was also found to inhibit biofilm formation of *C. jejuni* and furthermore, cells exposed to these osmolytes were found to assume a coccoid shape under a microscope, signifying their stress (40). *C. jejuni* is believed to be unable to metabolise glucose and most of other carbohydrates (285), therefore, it is likely that its effect on *C. jejuni* cultivation may be largely contributed to its osmolyte properties. Interestingly, it has been previously reported

that mixtures of osmolytes may cancel out their individual effect on enzyme activity (286). Furthermore, it has been reported that sucrose may cancel out inhibitory effect of NaCl in salt stressed *Sinorhizobium meliloti* (287). Both NB2G and TSB contain NaCl as well as glucose, while MHBG does not contain NaCl. Hence, it is possible that the lack of NaCl in MHBG may partially explain decreased tolerance of cells to glucose in comparison to cultivation in other broths, however, further validation would be required.

4.4 Summary

Identification of environmental factors such as the impact of nutritional profile of liquid media on *C. jejuni* prevalence and biofilm formation may lead to development of new strategies to control this pathogen. For example, development of new drugs targeting specific metabolic pathways may decrease *C. jejuni* incidence in livestock and thus its transmission to humans (29). Further, it may also improve control over *C. jejuni* at the food processing stages, as the conditions these pathogens face at each stage could be manipulated to create an environment which inhibits biofilm formation of this species (40).

The evidence presented here suggests that the composition of culture medium rather than the level of nutrients may be of more importance in determining the extent of biofilm formation of *C. jejuni.* Extremely high or extremely low dilutions of a specific medium were found in some cases to inhibit biofilm formation, but not for all media considered, and in the interim values biofilm formation was found to not vary substantially.

In contrast, adding certain ingredients such as glucose or yeast extract was found to significantly inhibit biofilm formation in MHB or NB2 media. Conversely, yeast extract was found to substantially enhance biofilm formation in TSB media.

Overall, the TSBYE media, which contains both yeast extract and glucose in its composition, (two components which were found to inhibit biofilm formation in MHB or NB2 media), was found to perform best in terms of both biofilm formation and cell proliferation, which was measured by viable cell counts in biofilm supernatants. This suggests that conditions which favour growth also favour biofilm formation of *C. jejuni* ATCC33291. Furthermore, it seems that nutrient composition as well as interaction between ingredients within the medium play an important role in determining the extent of biofilm formation. Specifically, the suitability of nutrients for *C. jejuni* metabolic requirements, as well as osmolarity of the media seems to be of importance. These results confound previous reports of low nutrient media favouring biofilm formation of *C. jejuni.* It is important to note that as *C. jejuni* strains vary in their metabolism (29, 285) and sensitivity to various environmental stresses (48, 288), it is likely that different patterns could be observed for other strains.

Chapter 5 - A Genome-wide Association Study (GWAS) of Biofilm Formation in *Campylobacter jejuni* Species

5.1 Introduction

Campylobacter jejuni is a pathogenic microaerophilic bacterial species for which biofilm formation has been identified as one of the key survival mechanisms in host to host transfers, both as monocultures and in mixed species biofilms (262, 263). The transmission of *Campylobacter jejuni* to humans usually occurs through ingestion of contaminated food (undercooked meat or meat products, raw or contaminated milk) (18). Both genetic and environmental factors contribute to the biofilm formation and an understanding of their relative contribution is not clear (265). This problem is analogous to the 'nature vs nurture' debate, which has been widely investigated across the broad spectrum of life sciences.

GWAS studies aim to identify links between genetic components and a chosen trait in a given population of organisms. For example, genetic factors associated with certain diseases (289), pathogen virulence (290) or antimicrobial resistance of microorganisms (291) may be identified with the use of GWAS techniques. Successful application of GWAS methods to microbial populations has proven more challenging than genetic association studies on human populations. This has been partly attributed to the fact that bacteria reproduce clonally, rather than by mixing the DNA of two unrelated individuals. As such, in human populations, due to obligatory mixing of genetic material in every generation, mutations may appear in a diverse range of genetic backgrounds. On the other hand, in bacteria, mutations spread in a single genetic background, unless they are acquired through horizontal gene transfer mechanism. As a result, it is harder to determine whether a given mutation is responsible for a specific trait, or whether it is a combination of many genetic factors found across a given lineage (291). Current microbial GWAS methods include algorithms which strive to take into account the effect of population structure by, for example, inclusion of phylogenetic trees in their input, which estimate relatedness between the analysed organisms, however, no gold standard has been achieved among GWAS methods (292). For example, sample size has been identified as an important factor affecting performance of the GWAS methods in detecting significant genotype-phenotype associations (292). Although it has been found more challenging to uncover significant associations between the bacterial phenotype and its genotypic components, there are also considerable advantages to bacterial GWAS in comparison to human GWAS. Firstly, as bacterial organisms are less complicated than humans, it is easier to identify molecular functions of the genetic associations identified by GWAS. Furthermore, GWAS findings for bacteria can be more easily validated by e.g.,

experimental deletion or insertion of a given gene into the bacterial isolate and observing whether it adopts the expected phenotype.

This study employed two currently applied Python-based GWAS tools, Scoary and Pyseer, to search for genes and core genome SNPs associated with biofilm formation of *Campylobacter jejuni*. Both tools were applied to sequences of 49 *C. jejuni* isolates which have previously been categorised as 'upper' or 'lower' biofilm formers in a 2015 study by Pascoe et al. (265). The workflow has been based on a GWAS analysis of *Listeria monocytogenes* growth in cold conditions, presented in ref. (293), with modifications regarding usage of some bioinformatics tools.

While the sample size of 49 isolates is quite small in relation to other GWAS studies where sample sizes of thousands of isolates have been employed (93), this was the biggest sample we could obtain in which isolates have been classified in terms of their biofilm forming ability in a systematic way. The necessity to test for a specific phenotype of interest is an important factor which may limit sample size. Similarly to our study, the GWAS analysis of *L. monocytogenes* in relation to its growth in cold conditions used a sample size of 51 isolates (293).

5.2 Methodology

Selection of isolates

From a list of 102 *Campylobcacter jejuni* and *Campylobacter coli* isolates listed in a previous work by Pascoe et al. (265), the sequences of 56 *C. jejuni* isolates categorised as either upper or lower biofilm formers were recovered from PubMLST. Some isolates which fulfilled the criteria of belonging to *C. jejuni* species and being either an upper or lower biofilm former could not be recovered from the PubMLST database and were therefore excluded from the analysis (for example, cow2673 and cow53 from the ST-21 clonal complex and cow230 and CampsClin1003 from the ST-45 clonal complex). The quality metrics of the genome sequences were evaluated alongside of genome annotation in PATRIC (https://www.patricbrc.org/) which makes use of the CheckM algorithm (294). Poor quality genomes were discarded from further analysis.

The list of the remaining isolates used in the GWAS study, including details of their individual accession numbers, clonal complexes, biofilm formation ability and quality metrics can be found in Appendix B1. A summary of these isolates is given in Table 5.1.

.

Table 5.1 *Classification details of biofilm forming ability of C. jejuni isolates used in this GWAS analysis as specified in ref.* (265)*. In particular, OD_{600} represents the absorbance value of biofilms measured by crystal violet assay in the work published in ref.* (265)*.*

Biofilm Forming Ability	OD_{600}	No. isolates	No. ST-21 clonal complex isolates	No. ST-45 clonal complex isolates	Other clonal complexes
Lower	<0.201	28	6	11	11
Upper	>0.272	21	9	7	5

5.2.1 Identification of biofilm associated genes

The sequences of the isolates remaining after genome quality analysis were annotated using Prokka, producing annotations in GFF3 format. The annotations were subsequently submitted into Roary pipeline (295), made available on Galaxy Australia (https://usegalaxy.org.au/), to produce a table of gene presence-absence for the pan genome in CSV and Rtab format and a core genome alignment file. To account for the microbial clonal population structure in the GWAS analysis, a phylogenetic tree was built with RAxML-NG (Random Accelerated Maximum Likelihood Next Generation) with default settings (296), using the core genome alignment of isolates generated by Roary. The traits file was created manually, containing a binary table ranking the isolates as either high biofilm formers (1) or low biofilm formers (0). Two different GWAS tools were used, Scoary (297) and Pyseer (298), both accepting the gene presence absence table, a phenotype table and the RAxML-generated Newick tree as input.

For both Pyseer and Scoary, p-value of associations corrected by the population structure effect were recorded. Only results with p-val<0.05 were used for further analysis. The outcomes were filtered out further accordingly to recommendations in Pyseer and Scoary documentations, respectively. Namely, for Scoary analysis, those associations with naïve p-value>0.05 based on a Fisher's test were removed (297) and for Pyseer, the associations with either one of the errors: 'bad-chisq' or 'high-bse' were removed, as they were suggested to likely represent spurious associations (298, 299).

5.2.2 Identification of biofilm associated SNPs

The SNPs-GWAS was also conducted with Scoary and Pyseer, using methods described in the documentation of these tools. For SNPs-GWAS, core genome SNPs in VCF format were generated with Harvest suite software (parsnp) with recombination filtration (300), using *Campylobacter jejuni* strain NCTC11168 (Genbank accession number: AL111168.1) as a

reference sequence. The core SNPs in VCF format were converted to Scoary input format with vcf2scoary command included in the Scoary package. The obtained SNPs with p-value<0.05 were annotated with TRAMS (Tool for Rapid Annotation of Microbial SNPs) (301)

5.3 Results and discussion

Out of the 56 recovered sequences, 7 sequences (cow58, chick2253, cow3189, OxClinb21, OxClinb45, CampsClin583 and chick2197) were discarded due to poor genome quality (e.g. high contamination proportion or high contigs count) This left 49 *C. jejuni* genomes to be subjected to the GWAS analysis. Out of those, 21 were labelled as "upper" biofilm formers in the analysis by B. Pascoe (265). The genome quality details of the 49 isolates can be accessed in the Appendix B1.

5.3.1 GWAS based on presence-absence of genes

According to output generated by Roary, the pan genome of the 49 *Campylobacter jejuni* isolates consists of 4175 gene clusters (Figure 5.1), which include 1,296 core genes (present in ≥95% of isolates) and 2,879 accessory genes. In comparison, the previously reported sizes of the *C. jejuni* core genome ranged from 847 to 1,643 genes, depending on the methods used and the diversity of chosen isolates (302).

Applying the filters specified in the method for gene-phenotype association, Scoary did not report any significant hits for the 49 isolates. This can most likely be attributed to the relatively low sample number and the fact that Scoary has been reported to have a very low power (probability of detecting a true association) when the population structure filters are applied for samples of size 20-80 (297). On the other hand, Pyseer returned 14 significant associations, and two of them returned an annotated product (Table 5.2) The list of hypothetical proteins with gene-biofilm phenotype associations indicated by Pyseer can be found in Appendix B.2. Blastx was ran on the nucleotide sequences of the hypothetical proteins to assess its similarity with annotated NCTC11168 proteins, and the results have been recorded in Appendix B.2.

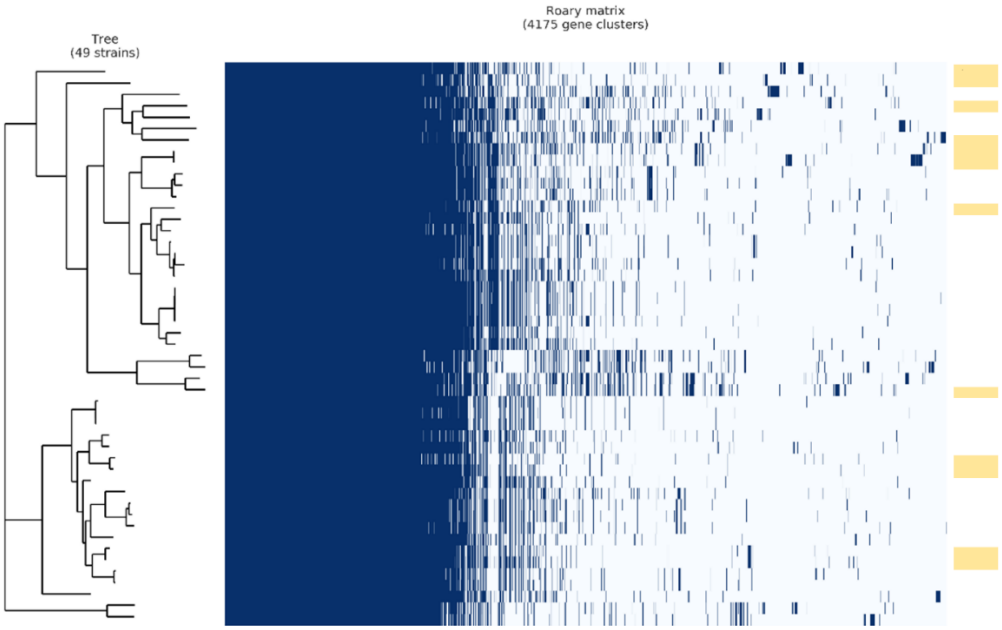

Figure 5.1 *Pan genome of 49 Campylobacter jejuni isolates generated by Roary. The dark blue and light blue shading represents gene presence and absence, respectively. On the right, the yellow shading indicates the presence of group_2452 or group_2964, which was indicated to be associated with the biofilm phenotype by Pyseer ($p<0.05$)*

Table 5.2 *In the first row - annotated genes with a statistically significant ($p<0.05$) association with biofilm formation obtained with the gene-based GWAS methods by Pyseer for 49 isolates (hypothetical proteins excluded). No associations were detected by Scoary in the gene-GWAS. (+) and (-) indicates isolates possessing and lacking the genes, respectively, and "upper" and "lower" indicate the isolates which were upper and lower biofilm formers, respectively. In the second row – the association with the biofilm phenotype obtained when combining two homologues of siaA and legI_2 which were found in the pan genome.*

Gene	**Non-unique gene name**	**(+) upper**	**(+) lower**	**(-) upper**	**(-) lower**	**Product**	**p-value**
group_2452 *group_2964*	*legI_2* *siaA*	7	5	14	23	N,N'-diacetyllegionaminic acid synthase UDP-N-acetylglucosamine 2-epimerase	0.0417
Two pairs of homologues combined	-	16	12	5	16	As above	0.0159

Included in the output presented in Table 5.2 are genes encoding for N,N'-diacetyllegionaminic acid synthase (*legl_2*) and UDP-N-acetylglucosamine 2-epimerase (*siaA*). UDP-N-acetylglucosamine 2-epimerase is believed to be involved in sialic acid biosynthesis (303). Sialic acid is a component of the *C. jejuni* lipopolysaccharide (LOS) and of the posttranslational modifications of the flagellin. It is an important component of the bacterium surface for a number of pathogens, and it has been shown to support bacterial resistance (304). On the other hand, N,N'-diacetyllegionaminic acid synthase is involved in synthesis of legionaminic acid – a compound found previously on the flagella of *C. jejuni*, which has been suggested to enhance *C. jejuni* colonization capabilities (305).

Interestingly, these two genes had the same pattern of presence/absence in the 49 isolates (Figure 5.1), suggesting a possible connection between these genes. Furthermore, it seems that these two genes are widely spread across the phylogeny (Figure 5.1). It should be noted that another pair of genes, also annotated as *siaA* and *legl_2*, was detected in the pan genome of the 49 isolates by Roary. For the second pair, the pair of genes was only present in one area of the phylogenetic tree (mostly in ST-21 isolates) and no association with the biofilm phenotype was detected ($p>0.05$) for both Pyseer and Scoary. Furthermore, the two variants of the genes did not overlap in any of the isolates. This suggests that there may exist variants of *siaA* and *legl_2* between isolates, which caused Prokka to annotate them separately. Interestingly, when the presence of the two pairs of these homologues was combined and Pyseer was ran again, this improved the strength of the association with the biofilm phenotype (Table 5.2), and a smaller proportion of isolates which did not possess one of the variants of these genes was able to form extensive biofilms.

As *siaA* and *legl_2* variants appeared together in the analysed isolates, no statement can be made as to whether one of them or both of them are more likely to be associated with the biofilm phenotype. In the literature, we were able to find experimental evidence that a *vpsA* gene, which encodes for UDP-N-acetylglucosamine 2-epimerase in *Vibrio cholera*, has been shown to affect biofilm formation of this organism, as a mutant deficient in this gene was observed to have reduced biofilm formation ability. Similarly, a *siaA* defective mutant of *Haemophilus influenza* has also exhibited inhibited biofilm formation (306). These considerations, together with our findings, suggest that at least one of the genes indicated in our analysis may affect biofilm formation ability of *C. jejuni* as well.

5.3.2 SNP-based GWAS results

The SNP-based GWAS was performed by comparing core genome SNPs with the biofilm phenotype. 71,049 SNPs in the core genome were detected, which includes SNPs present in the core genes and SNPs detected in intergenic regions. These SNPs were used as input for

this analysis, along with a Newick phylogenetic tree and the phenotype table. Again, associations were filtered accordingly to documentation specifications of respective toolkits. A test for lineage effects was performed in Pyseer which returned no hits (all p-values>0.2). Figure 5.2 represents an R-generated Manhattan plot of the position of SNPs on the reference genome (NCTC11168) and their associated p-value calculated by Pyseer. There were 758 slight associations detected by Pyseer and 9 detected by Scoary ($p<0.05$). Out of those, there were 107 and 1 strong ($p<0.01$) SNP associations reported, respectively. Five core genome SNPs were indicated by both Pyseer and Scoary to be possibly associated with the biofilm phenotype (Table 5.3), however, for all those hits Pyseer returned one of the error messages: 'high-bse' or 'bad-chisq'. Although we have discarded all other Pyseer associations which included these messages, as mutual outputs by Scoary and Pyseer were rare, we decided to make an exception and keep those associations for further investigation.

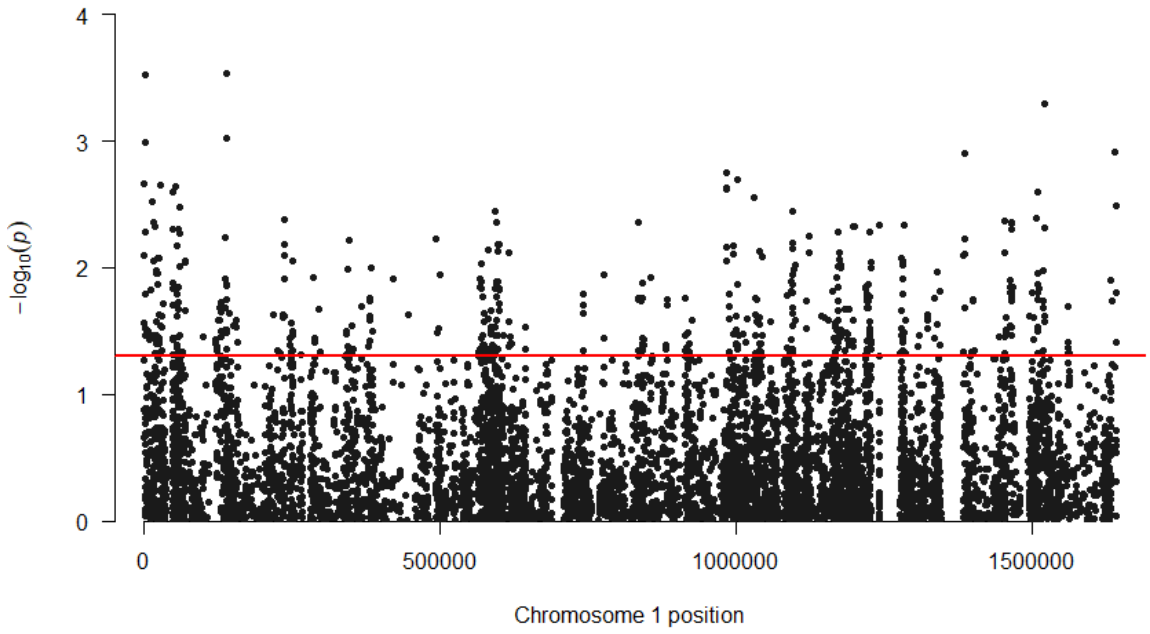

Figure 5.2 *Manhattan plot of the core genome SNPs which show an association to biofilm formation ability. Each point presents an SNP with its position on the NCTC11168 genome and its corresponding -log10(p-value), which represents the significance of the association calculated by Pyseer. The points above the red line signify SNPs associated with biofilm formation with $p<0.05$.*

All the SNPs indicated by both Scoary and Pyseer are synonymous mutations (Table 5.3), i.e., they do not alter the amino acid encoded by the changed codon. Due to this property, synonymous mutations are often referred to as silent mutations and they were reported to have been overlooked in some genetic studies (307). In recent years, however, evidence has emerged for synonymous mutations affecting human diseases and their outcomes in

human GWAS studies (307). Furthermore, synonymous mutations have been found to alter protein levels, protein structure, RNA stability or translational speed, among other things (307). Due to these considerations, synonymous mutations should not be overlooked. Among the output presented in Table 5.3, we can find *Cj0035c* – a gene encoding for an efflux protein which has previously been linked to fluoroquinolone resistance (308). In another gene, *kgtP*, two SNPs were indicated. *kgtP* encodes for alpha-ketoglutarate permease. In *C. jejuni*, KgtP is responsible for transport of citric acid cycle intermediates, which can be used as nutrient sources (285). A mutation in *kgtP* may potentially affect growth, as it is directly associated with bacterial metabolism. For example, a *kgtP* deletion mutant of *Xanthomonas oryzae pv. oryzae* was reported to be defective in both virulence and growth (309).

Other gene mutations indicated to be associated with the biofilm phenotype may affect a truncated KdpD protein (Cj0679). KdpD plays a regulatory role in potassium transport and the two component system KdpD/KdpE has been reported to be involved in pathogen virulence and survival in stressful conditions (e.g. antimicrobial stress, osmotic stress, oxidative stress) for a number of bacterial pathogens (i.e. *Staphylococcus aureus*, *Escherichia coli, Salmonella typhimurium*, *Yersinia pestis*, *Francisella* species, *Photorhab dus asymbiotica*, and mycobacteria) (310). We were not able to find a study on the effect of *kgtP* or *kdpD* mutations on *Campylobacter jejuni* biofilms, but our results suggest such an assay may be of interest.

Table 5.3 *List of SNPs identified as significantly associated with the biofilm phenotype by both tools, Scoary and Pyseer.*

Ref. pos.	Scoary pvalue	Pyseer pvalue	REF	ALT	Type	Ref codon	New codon	Ref res.	New res.	Locus tag	Gene	Product
1641351	0.0215	0.0013	C	T	synonymous	TCG	TCa	ser	ser	Cj1731c	*ruvC*	Crossover junction endodeoxyribonuclease
53763	0.0313	0.0066	A	G	synonymous	CCT	CCc	pro	pro	Cj0035c		Putative efflux protein
631981	0.0313	0.0358	G	A	synonymous	AAG	AAa	lys	lys	Cj0679	*kdpD*	Truncated KdpD protein
1546840	0.0313	0.0385	G	A	synonymous	TTG	TTa	leu	leu	Cj1619	*kgtP*	Alpha-ketoglutarate permease
1546858	0.0313	0.0385	T	C	synonymous	CTT	CTc	leu	leu	Cj1619	*kgtP*	Alpha-ketoglutarate permease

Table 5.4 *List of the first twenty SNPs with strongest evidence of association (lowest p-values given by Pyseer) to the biofilm phenotype. The complete list of associations is given in Appendix B.3.*

Ref. pos.	pvalue	REF	ALT	Type	Ref codon	New codon	Ref res.	New res.	Locus tag	Gene	Product
139381	0.0003	G	a	synonymous	GAG	GAa	glu	glu	Cj0136	*infB*	translation initiation factor IF-2
375	0.0003	C	t	synonymous	GCC	GCt	ala	ala	Cj0001	*dnaA*	chromosomal replication initiator protein
1519389	0.0005	C	t	synonymous	CGC	CGt	arg	arg	Cj1587c		putative ABC transporter
139321	0.0010	A	c	synonymous	GGA	GGc	gly	gly	Cj0136	*infB*	translation initiation factor IF-2
582	0.0010	G	a	synonymous	AAG	AAa	lys	lys	Cj0001	*dnaA*	chromosomal replication initiator protein
1638061	0.0012	T	a	nonsynonymous	TTT	TTa	phe	leu	Cj1728c		small hydrophobic protein
1386600	0.0013	T	c	synonymous	TTT	TTc	phe	phe	Cj1447c	*kpsT*	capsule polysaccharide export ATP-binding protein
982340	0.0018	A	g	synonymous	CGA	CGg	arg	arg	Cj1050c	*npdA*	NAD-dependent deacetylase
982388	0.0018	T	c	synonymous	TGT	TGc	cys	cys	Cj1050c	*npdA*	NAD-dependent deacetylase
1002617	0.0020	G	a	synonymous	TTG	TTa	leu	leu	Cj1067	*pgsA*	CDP-diacylglycerol--glycerol-3-phosphate 3-phosphatidyltransferase
1002618	0.0020	C	t	synonymous	CTA	tTA	leu	leu	Cj1067	*pgsA*	CDP-diacylglycerol--glycerol-3-phosphate 3-phosphatidyltransferase
180	0.0022	G	a	synonymous	GCG	GCa	ala	ala	Cj0001	*dnaA*	chromosomal replication initiator protein
25883	0.0022	C	t	synonymous	GCC	GCt	ala	ala	Cj0020c		cytochrome C551 peroxidase
52206	0.0023	G	a	nonsynonymous	GAT	aAT	asp	asn	Cj0034c		putative periplasmic protein
982195	0.0023	A	g	synonymous	GGA	GGg	gly	gly	Cj1049c		putative LysE family transporter protein
982057	0.0024	C	t	synonymous	GTC	GTt	val	val	Cj1049c		putative LysE family transporter protein
48436	0.0026	T	c	synonymous	TAT	TAc	Tyr	tyr	Cj0031		putative type IIS restriction/modification enzyme
1508461	0.0026	T	c	synonymous	CCT	CCc	pro	pro	Cj1577c	*nuoC*	NADH dehydrogenase I chain C
1029512	0.0028	A	g	intergenic	-	-	-	-	-	-	-
14769	0.0030	C	t	synonymous	AGC	AGt	ser	ser	Cj0009	*gltD*	glutamate synthase (NADPH) small subunit

Table 5.4 presents the SNPs indicated to be associated with biofilm phenotype by Pyseer analysis. Only the first twenty SNPs with the lowest p-values are shown – the rest of the statistically significant associations may be viewed in Appendix B.3. Similarly, all significant associations indicated by Scoary are listed in Appendix B.4.

One of the SNPs indicated to be associated with biofilm formation by Pyseer is located in the *kpsT* gene, which is involved in the transport of the polysaccharide capsule (CPS) across the cell envelope. It has been reported that removal of *kpsT* reduced colonisation ability of *Bordetella pertussis* in mice and furthermore, resulted in downregulation of some virulence genes of this species (311). Although no study has been done on effect of *kpsT* on biofilm formation or colonisation ability of *C. jejuni*, removal of another gene which resulted in disappearance of the CPS capsule in *C. jejuni* significantly reduced the ability of the mutant to attach to human epithelial cells (312). On the other hand, disruption of CPS has also been reported to not affect the ability of *C. jejuni* to colonise the chicken gut (312) or abiotic surfaces (313). Apart from potentially affecting colonisation ability and virulence, the CPS capsule has been reported to protect *C. jejuni* cells from damage when challenged with human serum, as *kpsM* mutants which did not express CPS were found significantly more susceptible (314). Pyseer has indicated three SNPs located in *kpsT* gene and three SNPs located in *kpsD* gene (which is also involved in CPS export) to be associated with the biofilm phenotype (Appendix B.3).

Another significant SNP association indicated by Pyseer was located in the *gltD* gene (Table 5.4), encoding for an enzyme which is a key part of metabolic networks of *C. jejuni* (285). In particular, it facilitates glutamate synthase activity. GltD has been previously indicated to be upregulated during biofilm growth of *Mycobacterium bovis* (315), and furthermore, utilization of glutamate or glutamine has been shown to be essential in biofilm growth of *Bacillus subtilis* (316).

A number of mutations in oxidoreductases (e.g. *thxB, nuoC, ilvC, nrdA, Cj0020c, ispH, katA*) has also been indicated to be associated with the biofilm phenotype by Pyseer ($p<0.05$). Most prominently, as many as 37 SNPs in catalase (*katA)* were included in the output of Pyseer. Catalase has been previously reported to be one of the primary components of the defence of *C. jejuni* against oxidative stress (317). Another key component protecting *C. jejuni* from oxidative stress, *ahpC,* has been reported to be regulated by *perR* and deletion of *perR* resulted in overexpression of *ahpC* and decreased biofilm formation (318). A nonsynonymous SNP in *perR* gene was also indicated to be significantly associated with the biofilm phenotype by Pyseer in our analysis ($p<0.05$).

Furthermore, SNPs in genes encoding for the flagellar hook protein (*flgE*) and a protein controlling the length of the flagellar hook (*fliK*) are included in the list of significant associations ($p<0.05$). Flagella expression and motility have been previously been indicated to be an important factor in mediating adhesion of *C. jejuni* to surfaces (8, 30), and the right hook length has been suggested to be essential for cell motility (319). Interestingly, an alias of the *flgE* gene (Cj1729c) has been previously indicated to be associated with the biofilm phenotype in a GWAS analysis performed by Pascoe et al. (265). In that analysis, 46 annotated genes were identified as possibly associated with the biofilm phenotype, and out of those, 17 genes were also identified in the present SNP GWAS analysis – among them genes suggested to play a role in sensing oxidative stress, metal uptake, nickel transport and nitrosative stress (Appendix B.5).

Finally, two SNPs (synonymous and nonsynonymous) in *rnc* gene, which encodes for ribonuclease (RNase III) and has previously been linked to virulence, were indicated by Pyseer ($p<0.05$) in our analysis. Interestingly, there's also some evidence suggesting the role of *rnc* in regulating exopolysaccharide synthesis and thus in mediating formation of the biofilm extracellular matrix of *Streptococcus mutans* (320).

5.4 Summary

A number of potential genetic associations to biofilm formation ability of *C. jejuni* has been identified in this multiple approach GWAS study. Among the returned associations, mutations in genes involved with oxygen reduction and oxygen stress response, metabolism, flagellar hook synthesis and length control, capsular polysaccharides (CPS) export or exopolysaccharides (EPS) synthesis have been identified. Those factors play a role in successful biofilm formation – for example, oxidoreductases help in *C. jejuni* survival, flagella facilitate movement and the adhesion to the surface (the latter also aided by CPS), and EPS is an important constituent of the biofilm extracellular matrix. It is therefore plausible that the candidates for biofilm-genotype associations suggested by our genome-wide association study may likely affect biofilm formation of *C. jejuni*, although more evidence is required to ascertain the strength of these associations *in vivo* (for example, by insertion/deletion experiments). The two methods used (Scoary and Pyseer) differed substantially in their output, which indicates the discrepancy between current GWAS methods and a need to test them thoroughly in order to identify an optimal one. Studies analysing and comparing the effectiveness of commonly used GWAS methods to identify true genotype-phenotype associations while minimizing the number of false positives have

indicated that a gold standard of GWAS methods is yet to be achieved (292). Potential improvements which have been recently proposed include the addition of a pre-processing step to SNP variants before subjecting them to GWAS analysis, in order to minimise the confounding effects of multiallelic sites, mismatches between the choices of a reference allele, and SNPs occurring in overlapping genes (321). Furthermore, tailoring GWAS analysis to a given species/phenotype of interest by utilizing known genotype-phenotype associations is also a promising strategy to improve the predictive power of GWAS (322).

In the GWAS analysis presented in this work, many more significant genetic associations with the biofilm phenotype have been detected among the core genome SNPs of *Campylobacter jejuni* than in the gene based GWAS. This can be partially attributed to a much smaller number of accessory genes in the 49 isolates of *C. jejuni* (2,879 gene clusters), compared to a number of detected core genome SNPs (71,049). Additionally, as the genes in the core genome are present in most of the isolates in contrast to accessory genes, the likelihood of finding a significant SNP association in those genes is higher.

It has been recognised that identification of genetic markers associated with the phenotype of interest at various stages of food production may improve control strategies in food safety (293). As biofilms are considered one of the key factors influencing survival of *C. jejuni* outside of the host (40, 48, 262), GWAS studies, which may help identify genetic markers associated with biofilm formation of *C. jejuni*, have a potential to help reduce the burden of this pathogen. Apart from informing control strategies at food production stages, GWAS studies may help identify targets for drug development, as novel strategies targeting very specific cell functions such as glutamate or glutamine metabolism have been recently introduced in this area to target biofilm formation of other problematic organisms (316).

Chapter 6 - General Conclusions and Future Directions

6.1 Major findings of this study

The main goal of this study was to investigate the effects of environmental and genetic factors on biofilm formation of *C. jejuni*. Three separate studies were designed which used different approaches, i.e., mathematical modelling, laboratory analysis and bioinformatics techniques. The components of the study were designed in order to address specific questions related to observed phenomena of biofilm formation of this species. In particular, the effects of nutrients and oxygen stress on *C. jejuni* biofilm formation were assessed by cultivating *C. jejuni* biofilms in the laboratory, and with the use of a mathematical model of biofilm formation which we proposed in this work. The model presented here is to the authors' knowledge the first attempt to apply mathematical modelling to study *C. jejuni* biofilm formation. A genome-wide association study (GWAS) was also conducted to reveal potential genetic factors associated with biofilm formation of *C. jejuni.* The results obtained in this study may inform the development of new strategies designed to control *C. jejuni* biofilms and thus ease the burden of human illness caused by this pathogen.

One of the important takeaways from this research project is that while it has been postulated previously that lower nutrient media may promote biofilm formation, in our experimental results we have found that rather than the level of nutrients, the type of nutrients and the interactions between them may play a bigger role in determining biofilm formation of *C. jejuni*. Specifically, we have identified a highly nutritious medium (TBSYE) which outperformed a lower nutrient medium (MHB) in cell growth in both planktonic and sessile states. Furthermore, we saw that addition of yeast extract to TSB promoted biofilm formation of *C. jejuni,* while it significantly inhibited biofilm formation when added to MHB or NB2.

Lysis has been stated before to be an important process in biofilm formation of *C. jejuni*, as it is believed to facilitate the release of eDNA which is considered to be a major component of *C. jejuni* biofilms. As such, as long as lysis does not surpass growth in the population, it may be beneficial for biofilm formation. In the mathematical model proposed in this work, one of the key assumptions was that increased nutrient levels increase lysis rate. This assumption, which was created based on the observed inhibition of both lysis and biofilm formation in starvation conditions, led to a prediction of the existence of an optimal nutrient concentration for which the effective growth rate is maximised. The assumption of increased lysis rate with increased nutrient levels may be expanded to include any chemical which promotes growth of *C. jejuni* in small amounts but inhibits its growth in large amounts. One such chemical identified with experimental results may be starch – an ingredient of MHB

which has been previously suggested to be beneficial to growth of *C. jejuni* populations in small amounts only. In our experiments, biofilm growth was inhibited in MHBx2 compared to MHB, which may likely be attributable to over optimal concentration of starch in MHBx2.

The model further predicts that the value of optimal nutrient concentrations may depend on the size of the inoculum, as while the survival probability of individual cells may be higher in lower nutrient conditions, the successful colonies are able to sustain a larger population of cells in higher nutrient conditions. This prediction indicates the importance of inoculum size in determining whether an invasion might be successful or not for given environmental conditions. This phenomenon arises from the assumption implemented in the model that mature biofilms act as a physical barrier in which cells experience different microenvironments to those found outside of the biofilm. As such, the environment within the biofilm may be pushed towards a regime in which the effective growth rate is positive, by for example, reducing the exposure of the cells to atmospheric oxygen. When an inoculum is large enough in high oxygen conditions, some cells may survive due to oxygen uptake and lysis of other cells which contributes to biofilm formation, among other potential reasons such as active release of extracellular material by the cells.

Our biofilm model also predicted a regime in which nearly equivalent numbers of live cells were present in biofilms cultivated in high oxygen conditions and in low oxygen conditions, while the density of biofilms in aerobic conditions was higher. On the other hand, when the exposure of cells to oxygen was high enough, biofilm formation was inhibited. Therefore, the previously observed patterns of enhanced or inhibited biofilm formation in aerobic conditions versus microaerobic conditions were both predicted by the model. Finally, studying the effect of the extracellular disruption mechanism through numerical simulations suggests that it is possible for a wide range of parameters to obtain equivalent biofilms even under oxygen stress with and without activating extracellular matrix disruption, which may explain the lack of association between the presence of DNAse encoding genes in a large proportion of *C. jejuni* isolates across its phylogeny and their biofilm formation ability.

The final contribution of this study is the list of potential genetic associations to biofilm formation ability of *C. jejuni*, which has been identified with the use of multiple genome-wide association (GWAS) techniques. Among the returned associations, mutations in genes involved with oxygen reduction and oxygen stress response, metabolism, flagellar hook synthesis and length control, capsular polysaccharides (CPS) export or exopolysaccharides

(EPS) synthesis have been identified. Further analysis of the identified associations would be necessary to confirm their effect on biofilm formation of *C. jejuni*.

6.2 Future directions

6.2.1 Mathematical modelling approach to study Campylobacter jejuni *biofilms*

As a result of this study, we provided a framework for future mathematical modelling studies of *C. jejuni* biofilm formation (Chapter 3). The model presented here could be improved by increasing the number of dimensions to 3, increasing the system space in order to study more mature biofilms, or by an introduction of another species which exhibits different properties. Furthermore, once enough information about *C. jejuni* metabolism is established, additional mechanisms such as release of metabolic products and utilization of different types of nutrients could be incorporated into this model and the effect of these mechanisms on the biofilm population could be assessed. In its current stage, the model is already versatile and more interesting predictions may be obtained with time by studying more of its components.

6.2.2 The influence of media and atmospheric conditions on Campylobacter jejuni *biofilm formation*

Identification of environmental factors such as the impact of nutritional profile of liquid media on *C. jejuni* prevalence and biofilm formation may lead to the development of new strategies to control this pathogen. Specifically, by improving safety measures implemented at food processing stages as well as measures aimed to control *C. jejuni* colonisation of livestock. Our experimental assay of biofilm formation under various environmental conditions (Chapter 4) suggests that the type of ingredients and the interactions between them in the media plays an important role in affecting biofilm formation ability of *C. jejuni.* This result confounds previous postulates about lower nutrient media being beneficial for biofilm formation of *C. jejuni.* Further studies on specific ingredients and their ability to promote or inhibit biofilm formation of *C. jejuni* at various concentrations may provide a better understanding of biofilm formation or persistence of *C. jejuni* outside of the host.

6.2.3 The influence of genetic factors on biofilm formation of Campylobacter jejuni

Identification of genetic markers associated with biofilm formation has a potential to help reduce the burden of *C. jejuni* through informing control strategies at food production stages or identification of targets for drug development. Construction of single gene-knockout mutants, gene complementation, or introducing single point mutations for the genes and

SNPs identified in our study (Chapter 5), could improve our understanding on the role of these mutations in affecting biofilm formation.

Appendices

Appendix A – Supplementary Material for Chapter 3

A.1 Algorithm for vital cell dynamics

We assume that vital transition events define a Poisson point process such that, if the system reached a configuration field $\{\sigma(i,t)\}$ after a cell transition at time *t*, it will remain in such a state during a period of time Δt with probability $e^{-\Lambda(t)\Delta t}$. Here,

$$\Lambda(t) = \sum_{i\in\mathcal{B}(t)} \sum_{k\in\mathcal{E}} \lambda_k(i,t), \tag{A.1}$$

is the decay rate of the state $\{\sigma(i,t)\}$ accounting for all the possible transitions, $\mathcal{E} = \{g, l, d\}$, of the patches occupied by cells, $\mathcal{B}(t) = \{i|\sigma(i,t) = C\}$. Within this formulation, the time Δt between two consecutive cell transitions is a random variable given by the distribution function $1 - e^{-\Lambda(t)\Delta t}$. In practice, the time Δt until the next cell event is calculated as

$$\Delta t = -\frac{\ln(1-u)}{\Lambda(t)}, \tag{A.2}$$

where u is a uniformly distributed random number in [0,1]. A patch *i* occupied by a cell at time *t* is selected to undergo a transition $k \in \mathcal{E}$ after time Δt with probability $\Lambda(t)^{-1}\lambda_k(i,t)$.

A.2 Steady State Difference Approximation

Let $S_k(x,y)$ denote a given nutrient concentration at *x* and *y* coordinates in the 2D model at a given time, and *D* the diffusion coefficient within the biofilm at time *t*. Assuming that the function $S_k(x,y)$ is both continuous and differentiable in the interval $x \in [x-1, x+1]$, Taylor expansion can be used to calculate the difference approximation of $S_k(x,y)$ about the point (x,y).

$$S_k(x+1,y) = S_k(x,y) + \frac{\partial S_k(x,y)}{\partial x} + \frac{1}{2}\frac{\partial^2 S_k(x,y)}{\partial x^2}$$

$$S_k(x-1,y) = S_k(x,y) - (\Delta x)\frac{\partial S_k(x,y)}{\partial x} + (\Delta x)^2\frac{1}{2}\frac{\partial^2 S_k(x,y)}{\partial x^2}$$

$$\frac{\partial^2 S_k(x,y)}{\partial x^2} = S_k(x+1,y) + S_k(x-1,y) - 2S_k(x,y)$$

Similarly,

$$\frac{\partial^2 S_k(x,y)}{\partial y^2} = S_k(x,y+1) + S_k(x,y-1) - 2S_k(x,y)$$

Finally, from Eq.(10), given that $\frac{\partial S_k(i,t)}{\partial t} = 0$, we obtain

$$\nabla^2 S_k(x,y) = S_k(x,y+1) + S_k(x,y-1) - 4S_k(x,y) + S_k(x+1,y) + S_k(x-1,y) = \frac{u_k(x,y)}{D} \tag{A.3}$$

It is important to note that Eq. (A.3) holds only for non-boundary sites, which are located below the biofilm boundary layer. In the presently described model, there are closed boundary conditions on the bottom (i.e. no flux through these boundaries), periodic boundaries on the sides of the system and a perfect source of nutrients above the biofilm boundary layer (i.e. the Infinite Reservoir, where $S_k(x,y) = const.$) At the closed boundaries, there are fewer directions to which the nutrient can escape to from the boundary site, and fewer directions from which the boundary site can obtain nutrients. Taking this into consideration, the steady state of the boundary sites is maintained when the following expression holds for all sites below the biofilm boundary layer:

$$\left[\sum_{j \in Nei(i)} S_k(j)\right] - \|\mathcal{N}(i)\| S_k(i) = \frac{u_k(i)}{D} \tag{A.4}$$

The notation change from Cartesian coordinates to site numbers has been chosen to shorten the length of the expression. Eq. (A.4) is an expression for concentration at one site only. In order to solve the concentration field for the whole 2-dimensional system of size $L \times H$, a set of equations which contains Eq. (A.4) for each site within must be solved instantaneously. Thus, a concentration matrix develops:

$$\begin{bmatrix} -\|\mathcal{N}(1)\| & \gamma_{1,2} & \cdots & \gamma_{1,LH} \\ \vdots & \ddots & \cdots & \vdots \\ \gamma_{LH,1} & \cdots & \gamma_{LH,LH-1} & -\|\mathcal{N}(LH)\| \end{bmatrix} \begin{bmatrix} S_k(1) \\ \vdots \\ S_k(LH) \end{bmatrix} = \frac{1}{D} \begin{bmatrix} u_k(1) \\ \vdots \\ u_k(LH) \end{bmatrix} \tag{A.5}$$

Where

$$\gamma_{i,j} = \begin{cases} 1 \ if \ j \in \mathcal{N}(i) \\ 0 \ otherwise \end{cases}$$

Introduction of the boundary layer leads to further modifications, as above the biofilm boundary layer, the concentration is held constant at S_k^H, and it so cannot be subjected to the diffusion-relaxation equations, as $\nabla^2 C = 0$ in this region. However, these sites contribute to the overall concentration distribution, so they must be included in the concentration matrix,

as constant values.

Let *s* denote the lowest site above the boundary layer, i.e. the first site at which the concentration of the compound is held constant. Then, Eq. (A.5) must be modified as follows:

$$\begin{bmatrix} -\|\mathcal{N}(1)\| & \dots & \gamma_{1,s} & \dots & \gamma_{1,LH} \\ \vdots & \ddots & \dots & \dots & \dots \\ \gamma_{s-1,1} & \dots & -\|\mathcal{N}(s-1)\| & \dots & \gamma_{s-1,LH} \\ \delta_{s,1} & \dots & \dots & \dots & \delta_{s,LH} \\ \vdots & \dots & \dots & \ddots & \vdots \\ \delta_{LH,1} & \dots & \dots & \dots & \delta_{LH,LH} \end{bmatrix} \boldsymbol{S_k} = \frac{1}{D} \begin{bmatrix} u_k(1) \\ \vdots \\ u_k(s) \\ \mathrm{S_k^H} \\ \vdots \\ \mathrm{S_k^H} \end{bmatrix} \tag{A.6}$$

Where

$$\boldsymbol{S_k} = \begin{bmatrix} S_k(1) \\ \vdots \\ S_k(s) \\ \vdots \\ S_k(LH) \end{bmatrix}$$

$\delta_{i,j}$ is the Kronecker delta, i.e.:

$$\delta_{i,j} = \begin{cases} 1 \ if \ j = i \\ 0 \ otherwise \end{cases}$$

It is easy to see that Eq. (A.6) will return $S_k(i) = \mathrm{S_k^H}$ for all $i > s$, thus satisfying the boundary conditions imposed upon the system.

Finally, the solution of the state of the concentration field for the given state results in the following concentration field:

$$\boldsymbol{S_k} = \frac{1}{D} \begin{bmatrix} -\|Nei(1)\| & \dots & \dots & \dots & \gamma_{1,LH} \\ \vdots & \ddots & \dots & \dots & \dots \\ \gamma_{s-1,1} & \dots & -\|Nei(s-1)\| & \dots & \gamma_{s-1,LH} \\ \delta_{s,1} & \dots & \dots & \dots & \delta_{s,LH} \\ \vdots & \dots & \dots & \ddots & \vdots \\ \delta_{LH,1} & \dots & \dots & \dots & \delta_{LH,LH} \end{bmatrix}^{-1} \begin{bmatrix} u_k(1) \\ \vdots \\ u_k(s) \\ \mathrm{S_k^H} \\ \vdots \\ \mathrm{S_k^H} \end{bmatrix} \tag{A.7}$$

One must be careful when using Eq. (A.7) for calculating the concentration field, as it is capable of returning negative concentration values which are not realistic. This error stems from the simplifying assumption that the uptake rate of cells is constant in the time interval [*t*-1, *t*], which can lead to an overconsumption effect.

This can be rectified by setting a lower boundary for concentration, so that whenever Eq. (A.7) predicts a negative value of the compound at a given site, the result is corrected to concentration of 0.

Appendix B – Supplementary Material for Chapter 5

B.1 List of isolates, their biofilm forming ability and genome quality details

C. jejuni isolate	PubMLST ID	MLST clonal complex	Biofilm formation (Ref. (265))	Biofilm (OD600)	Coarse consistency(%)	Fine consistency (%)	Completeness (%)	Contamination(%)	Contig count	DNA size (bp)	Contigs N50 (bp)
starling177	60474	ST-177	lower	0.134	98.9	98.9	100	0	52	1582720	78095
cow273	60644	ST-206	upper	0.278	99.1	98.7	100	0	62	1715399	86763
cow3205	60655	ST-206	lower	0.072	99.1	98.8	100	0	64	1720953	104138
cow206	60626	ST-206	lower	0.173	99.1	98.8	100	0	62	1671619	84503
chick1360	57770	ST-21	upper	0.301	99	98.8	100	0	67	1693941	77452
chickc21	57795	ST-21	upper	0.559	98.9	98.7	100	0	139	1679349	26520
cow518	60632	ST-21	upper	0.523	98.9	98.7	100	0	60	1705325	102715
cow3201	60653	ST-21	lower	0.163	99.1	97.9	100	1.9	114	1629692	72911
cowa21	60634	ST-21	upper	0.404	98.9	98.7	100	0	100	1658436	50185
cowb21	60646	ST-21	upper	0.28	99	99	100	0	130	1659711	32912
CampsClin21	60469	ST-21	lower	0.153	98.9	98.7	100	0	65	1656471	86827
chick53	57785	ST-21	lower	0.197	99.1	98.9	100	0	102	1651079	51915
chick104	57768	ST-21	lower	0.112	99	98.9	100	0	112	1761202	51191
cow104	60651	ST-21	lower	0.103	99	99	100	0	82	1762939	55677
chick19	60637	ST-21	upper	0.501	99	99	100	0	84	1689713	65558
chick50	57783	ST-21	upper	0.289	99	98.8	100	0	63	1692341	121354
chick1086	57769	ST-21	upper	0.349	99	98.8	100	0	65	1692435	127697
chick266	60639	ST-21	lower	0.177	99	98.8	100	0	76	1693845	77785
CampsClin883	57766	ST-21	upper	0.277	99.1	99	100	0	113	1667560	43442
chick354	57782	ST-257	lower	0.066	99	98.7	100	0	75	1697013	70536
chick353	57781	ST-353	lower	0.177	98.7	98.3	100	0	129	1776210	61083
chickb354	57793	ST-354	lower	0.16	99	98.7	100	0	143	1688706	41754
cow3583	60642	ST-42	lower	0.194	99	99	100	0	93	1654563	60789
chick51	57784	ST-443	lower	0.082	99.1	98.7	100	0.9	58	1714044	101326
chick2219	57776	ST-45	upper	0.317	99	99	100	0	84	1616482	56181

chickc45	57796	ST-45	upper	0.303	99	98.5	100	3.6	343	1595762	8702
cow137	60461	ST-45	upper	0.308	98.9	98.6	100	0	192	1626334	31482
cow334	60629	ST-45	upper	0.4	99	98.9	100	0	102	1616567	59062
cow3207	60463	ST-45	upper	0.603	99	99	100	0	70	1641790	61328
duck45	60483	ST-45	upper	1.005	99	99	100	0	60	1616162	81477
CampsClin11	57761	ST-45	lower	0.176	98.8	98.8	100	0	92	1650105	60731
chick1003	57767	ST-45	lower	0.201	98.9	98.7	100	0	106	1617200	43134
chick2048	57773	ST-45	lower	0.176	99	98.4	100	0.9	163	1631119	26959
chickb45	57794	ST-45	lower	0.087	99	99	100	0	80	1649834	64921
chickd45	57797	ST-45	lower	0.087	99	99	100	0	55	1618948	78275
cow3214	60464	ST-45	lower	0.08	98.9	98.8	100	0	115	1654338	48186
cowb45	60648	ST-45	lower	0.153	98.9	98.9	100	0	64	1603131	118082
cowd45	60650	ST-45	lower	0.127	98.9	98.9	100	0	98	1607691	48873
goose137	60479	ST-45	lower	0.109	99	99	100	0	53	1600143	110320
OxClina45	57799	ST-45	lower	0.118	99	99	100	0	66	1621668	62364
CAMP45	57759	ST-45	lower	0.161	99	98.9	100	0	108	1596969	28256
chick11	60640	ST-45	upper	0.384	98.8	98.8	100	0	120	1645238	36661
chick1079	60465	ST-573	lower	0.154	98.9	98.6	100	0	273	1838061	38231
chick573	60635	ST-573	upper	0.279	98.9	98.6	100	0	251	1838022	50424
chick574	57786	ST-574	upper	0.758	99	99	100	0	80	1743461	78544
CAMP61	60481	ST-61	upper	0.29	99.1	98.7	100	0.9	86	1551351	29685
chick25	57779	ST-661	lower	0.191	99	98.9	100	0	113	1698035	148299
chick814	60466	ST-661	lower	0.096	99.1	99	100	0	152	1759789	83080
starling1020	60476	ST-682	upper	0.62	99	98.9	100	0	61	1578916	75294

B.2 List of significant gene-GWAS associations (p<0.05) indicated by Pyseer – hypothetical proteins.

Gene	Product	NCTC11168 similar gene ID	Identity (%)	Product of the NCTC11168 gene	p-value	(+) upper	(+) lower	(-) upper	(-) lower
group_492	Putative protein	Cj1069	45	Hypothetical protein	0.00609	13	25	8	3
group_915	Hypothetical protein	Cj0564	92	Membrane protein	0.00159	14	24	7	4
smc_2	Hypothetical protein	Cj0019c	40	Hypothetical protein	0.0257	15	22	6	6
group_1497	Hypothetical protein	Cj1298	44	GDP-4-amino-4,6-dideoxy-alpha-D-acetylglucosamine N-acetyltransferase	0.0305	4	11	17	17
group_493	putative protein	Cj1069	47	Hypothetical protein	0.0254	9	5	12	23
group_542	Hypothetical protein	Cj1381	79	Putative lipoprotein	0.049	5	9	16	19
group_1072	Hypothetical protein	-	-	-	0.00619	8	5	13	23
group_1073	Hypothetical protein	-	-	-	0.0151	8	5	13	23
group_1662	Hypothetical protein	-	-	-	0.00872	6	7	15	21
group_2206	Hypothetical protein	-	-	-	0.0458	7	5	14	23
group_526	Hypothetical protein	Cj1302	44	HAD-superfamily phosphatase, subfamily IIIC	0.00836	4	8	17	20
group_209	Hypothetical protein	-	81	Oxygen-insensitive NAD(P)H nitroreductase/ Dihydropteridine reductase (EC 1.5.1.34)	0.0182	4	3	17	25

B.3 List of significant SNP associations (p<0.05) indicated by Pyseer

P-value	Position	Feature	Locus tag	Gene	Product	Start	End	Strand	Ref. base	Ref. codon	Ref. res.	SNP type	New base	New codon	New res.
0.000289	139381	CDS	Cj0136	infB	translation initiation factor IF-2	136856	139471	1	G	GAG	E	synonymous	a	GAa	E
0.000297	375	CDS	Cj0001	dnaA	chromosomal replication initiator protein	1	1323	1	C	GCC	A	synonymous	t	GCt	A
0.000505	151938	CDS	Cj1587c		putative ABC transporter	151800	151964	-1	C	CGC	R	synonymous	t	CGt	R
0.000956	139321	CDS	Cj0136	infB	translation initiation factor IF-2	136856	139471	1	A	GGA	G	synonymous	c	GGc	G
0.00102	582	CDS	Cj0001	dnaA	chromosomal replication initiator protein	1	1323	1	G	AAG	K	synonymous	a	AAa	K
0.00121	163806	CDS	Cj1728c		small hydrophobic protein	163792	163807	-1	T	TTT	F	nonsynonymous	a	TTa	L
0.00126	138660	CDS	Cj1447c	kpsT	capsule polysaccharide export ATP-binding protein	138624	138692	-1	T	TTT	F	synonymous	c	TTc	F
0.00179	982340	CDS	Cj1050c	npdA	NAD-dependent deacetylase	982247	982948	-1	A	CGA	R	synonymous	g	CGg	R
0.00179	982388	CDS	Cj1050c	npdA	NAD-dependent deacetylase	982247	982948	-1	T	TGT	C	synonymous	c	TGc	C
0.00204	100261	CDS	Cj1067	pgsA	CDP-diacylglycerol--glycerol-3-phosphate 3-phosphatidyltransferase	100256	100310	1	G	TTG	L	synonymous	a	TTa	L
0.00204	100261	CDS	Cj1067	pgsA	CDP-diacylglycerol--glycerol-3-phosphate 3-phosphatidyltransferase	100256	100310	1	C	CTA	L	synonymous	t	tTA	L
0.00218	180	CDS	Cj0001	dnaA	chromosomal replication initiator protein	1	1323	1	G	GCG	A	synonymous	a	GCa	A
0.00223	25883	CDS	Cj0020c		cytochrome C551 peroxidase	25433	26347	-1	C	GCC	A	synonymous	t	GCt	A
0.00229	52206	CDS	Cj0034c		putative periplasmic protein	51967	52668	-1	G	GAT	D	nonsynonymous	a	aAT	N
0.00234	982195	CDS	Cj1049c		putative LysE family transporter protein	981655	982254	-1	A	GGA	G	synonymous	g	GGg	G
0.00241	982057	CDS	Cj1049c		putative LysE family transporter protein	981655	982254	-1	C	GTC	V	synonymous	t	GTt	V

0.00255	48436	CDS	Cj0031		putative type IIS restriction/modification enzyme	46424	50156	1	T	TAT	Y	synonymous	c	TAc	Y
0.00255	1508461	CDS	Cj1577c	nuoC	NADH dehydrogenase I chain C	1508131	1508925	-1	T	CCT	P	synonymous	c	CCc	P
0.00278	1029512	intergenic							A				g		
0.00304	14769	CDS	Cj0009	gltD	glutamate synthase (NADPH) small subunit	14398	15843	1	C	AGC	S	synonymous	t	AGt	S
0.00325	1641342	CDS	Cj1731c	ruvC	crossover junction endodeoxyribonuclease	1640904	1641386	-1	T	TGT	C	synonymous	c	TGc	C
0.00335	59202	CDS	Cj0040		hypothetical protein Cj0040	59154	59477	1	C	CTA	L	synonymous	t	tTA	L
0.00361	592211	CDS	Cj0630c	holA	putative DNA polymerase III, delta subunit	591537	592502	-1	A	ACG	T	nonsynonymous	g	gCG	A
0.00361	1093736	CDS	Cj1161c		putative cation-transporting ATPase	1091795	1093894	-1	G	AAG	K	synonymous	a	AAa	K
0.00406	1505523	CDS	Cj1573c	nuoG	putative NADH dehydrogenase I chain G	1503531	1505993	-1	C	GGC	G	synonymous	a	GGa	G
0.00415	235804	CDS	Cj0256		putative sulfatase family protein	234599	236137	1	C	GAC	D	synonymous	t	GAt	D
0.00425	1453196	CDS	Cj1517	moaD	putative molybdopterin converting factor,subunit 1	1453143	1453364	1	A	GAA	E	synonymous	g	GAg	E
0.00433	15042	CDS	Cj0009	gltD	glutamate synthase (NADPH) small subunit	14398	15843	1	T	AGT	S	synonymous	c	AGc	S
0.00436	1463552	CDS	Cj1531	dapF	diaminopimelate epimerase	1463156	1463905	1	A	AAC	N	nonsynonymous	g	gAC	D
0.00436	1463575	CDS	Cj1531	dapF	diaminopimelate epimerase	1463156	1463905	1	A	GTA	V	synonymous	g	GTg	V
0.00446	594016	CDS	Cj1631c		putative ribonuclease	592495	594429	-1	T	ACT	T	synonymous	c	ACc	T
0.00446	594163	CDS	Cj1631c		putative ribonuclease	592495	594429	-1	C	GTC	V	synonymous	t	GTt	V
0.00446	594340	CDS	Cj1631c		putative ribonuclease	592495	594429	-1	T	ACT	T	synonymous	c	ACc	T
0.00442	834290	CDS	Cj0894c	ispH	4-hydroxy-3-methylbut-2-enyl diphosphate reductase	834005	834838	-1	C	AAC	N	synonymous	t	AAt	N
0.00442	834521	CDS	Cj0894c	ispH	4-hydroxy-3-methylbut-2-enyl diphosphate reductase	834005	834838	-1	T	AGT	S	synonymous	c	AGc	S
0.00449	1464118	CDS	Cj1532		putative periplasmic protein	1463880	1464611	1	T	ATA	I	nonsynonymous	c	AcA	T
0.00461	1283352	CDS	Cj1351	pldA	phospholipase A	1282933	1283922	1	A	GGA	G	synonymous	g	GGg	G
0.00463	1241088	CDS	Cj1311	pseF	putative acylneuraminate cytidylyltransferase	1240966	1241664	1	T	GGT	G	synonymous	c	GGc	G

0.00467	17871	CDS	Cj0013	ilvD	dihydroxy-acid dehydratase	17563	19239	1	G	ACG	T	synonymous	a	ACa	T
0.00469	119734	CDS	Cj1267c	hydA	Ni/Fe-hydrogenase small chain	1197214	1198353	-1	C	ACC	T	synonymous	t	ACt	T
0.00469	119741	CDS	Cj1267c	hydA	Ni/Fe-hydrogenase small chain	1197214	1198353	-1	G	GGG	G	synonymous	a	GGa	G
0.00469	119805	intergenic							A				g		
0.00486	152122	CDS	Cj1589		conserved hypothetical protein Cj1189	1521072	1521869	1	C	CAC	H	synonymous	t	CAt	H
0.00493	47470	CDS	Cj0031		putative type IIS restriction/modification enzyme	46424	50156	1	T	GTT	V	synonymous	c	GTc	V
0.00493	57250	CDS	Cj0039c	typA	GTP-binding protein TypA homolog	57211	59019	-1	C	GAC	D	synonymous	t	GAt	D
0.00493	57286	CDS	Cj0039c	typA	GTP-binding protein TypA homolog	57211	59019	-1	A	ACA	T	synonymous	g	ACg	T
0.00496	146460	CDS	Cj1532		putative periplasmic protein	1463880	1464611	1	C	ACC	T	synonymous	a	ACa	T
0.00519	122573	CDS	Cj1293	pseB	UDP-GlcNAc-specific C4,6 dehydratase/C5 epimerase	1224849	1225853	1	G	GTG	V	synonymous	a	GTa	V
0.00525	414	CDS	Cj0001	dnaA	chromosomal replication initiator protein	1	1323	1	A	CCA	P	synonymous	t	CCt	P
0.00528	117180	CDS	Cj1243		uroporphyrinogen decarboxylase	1170922	1171944	1	C	GGC	G	synonymous	t	GGt	G
0.00528	117113	CDS	Cj1243		uroporphyrinogen decarboxylase	1170922	1171944	1	C	ACC	T	synonymous	t	ACt	T
0.00535	59228	CDS	Cj0040		hypothetical protein Cj0040	59154	59477	1	G	TTG	L	synonymous	a	TTa	L
0.00561	112234	CDS	Cj1194		possible phosphate permease	1122289	1123815	1	T	TCT	S	synonymous	g	TCg	S
0.00575	136779	CDS	Cj0135		conserved hypothetical protein Cj0135	136612	136869	1	A	TTA	L	synonymous	g	TTg	L
0.0059	492958	CDS	Cj0530		putative periplasmic protein	491390	493960	1	T	AGT	S	synonymous	c	AGc	S
0.00596	138456	CDS	Cj1444c	kpsD	capsule polysaccharide export system periplasmic protein	1383486	1385144	-1	C	GGC	G	synonymous	t	GGt	G
0.00611	346058	CDS	Cj0378c		putative ferric reductase-like transmembrane protein. Functional classification-Membranes,lipoproteins and porins	345635	346174	-1	C	ACC	T	synonymous	t	ACt	T
0.00638	109457	CDS	Cj1163c		putative cation transport protein	1094078	1095028	-1	G	TTG	L	synonymous	a	TTa	L

0.0065	235847	CDS	Cj0256		putative sulfatase family protein	234599	236137	1	A	AGC	S	MNP	g		
0.0065	235848	CDS	Cj0256		putative sulfatase family protein	234599	236137	1	G	AGC	S	MNP	a		
0.0065	235849	CDS	Cj0256		putative sulfatase family protein	234599	236137	1	C	AGC	S	nonsynonymous	t	gat	D
0.00651	597935	CDS	Cj0636		NOL1/NOP2/sun family protein	597828	598667	1	T	AAT	N	nonsynonymous	a	AAa	K
0.00651	598094	CDS	Cj0636		NOL1/NOP2/sun family protein	597828	598667	1	T	AGT	S	nonsynonymous	g	AGg	R
0.00664	993744	CDS	Cj1057c		putative coiled-coil protein	993678	993851	-1	A	GAA	E	synonymous	g	GAg	E
0.00664	993753	CDS	Cj1057c		putative coiled-coil protein	993678	993851	-1	G	GAG	E	synonymous	a	GAa	E
0.00664	993759	CDS	Cj1057c		putative coiled-coil protein	993678	993851	-1	C	TAC	Y	synonymous	t	TAt	Y
0.00674	54695	CDS	Cj0036		hypothetical protein Cj0036	53970	55319	1	C	AGC	S	synonymous	t	AGt	S
0.00682	982727	CDS	Cj1050c	npdA	NAD-dependent deacetylase	982247	982948	-1	G	CAG	Q	synonymous	a	CAa	Q
0.0071	1094213	CDS	Cj1163c		putative cation transport protein	1094078	1095028	-1	T	TCT	S	synonymous	c	TCc	S
0.00719	581069	CDS	Cj0621		hypothetical protein Cj0621	580326	581846	1	A	TTA	L	synonymous	g	TTg	L
0.00739	597895	CDS	Cj0636		NOL1/NOP2/sun family protein	597828	598667	1	C	ACT	T	nonsynonymous	t	AtT	I
0.0075	1038972	CDS	Cj1104	ispE	putative 4-diphosphocytidyl-2-C-methyl-D-erythritol kinase	1038334	1039101	1	C	AAC	N	synonymous	t	AAt	N
0.00753	1122826	CDS	Cj1194		possible phosphate permease	1122289	1123815	1	A	ATA	I	nonsynonymous	g	gTA	V
0.00753	1174230	CDS	Cj1246c	uvrC	excinuclease ABC subunit C	1174038	1175840	-1	T	ATT	I	synonymous	c	ATc	I
0.00769	616190	pseudogene	Cj0654c		pseudogene (putative transmembrane transport protein)	615014	616485	-1	G	ATG	M	nonsynonymous	a	ATa	I
0.00769	1453169	CDS	Cj1517	moaD	putative molybdopterin converting factor,subunit 1	1453143	1453364	1	T	CCT	P	synonymous	c	CCc	P
0.00784	993902	CDS	Cj1058c	guaB	inosine-5'-monophosphate dehydrogenase	993848	995305	-1	A	GTA	V	synonymous	g	GTg	V
0.00786	1386582	CDS	Cj1447c	kpsT	capsule polysaccharide export ATP-binding protein	1386264	1386926	-1	A	GAA	E	synonymous	g	GAg	E
0.00793	1383945	CDS	Cj1444c	kpsD	capsule polysaccharide export system periplasmic protein	1383486	1385144	-1	T	GGT	G	synonymous	c	GGc	G

0.0080 1	168	CDS	Cj0001	dnaA	chromosomal replication initiator protein	1	1323	1	C	GGC	G	synonymous	t	GGt	G
0.0080 8	235758	CDS	Cj0256		putative sulfatase family protein	234599	236137	1	G	AGC	S	nonsynonymous	a	AaC	N
0.0081 4	1043364	CDS	Cj1110c		putative MCP-type signal transduction protein	1043223	1044512	-1	C	CGC	R	synonymous	t	CGt	R
0.0085 1	23287	CDS	Cj0017c	dsbI	disulphide bond formation protein	21854	23380	-1	A	ATA	I	nonsynonymous	g	gTA	V
0.0085 1	23288	CDS	Cj0017c	dsbI	disulphide bond formation protein	21854	23380	-1	G	TTG	L	nonsynonymous	t	TTt	F
0.0085 1	23288	sig_peptide	Cj0017c	dsbI		23288	23380		G				t		
0.0085 1	23799	CDS	Cj0019c		putative MCP-domain signal transduction protein	23665	25443	-1	T	TTA	L	synonymous	c	cTA	L
0.0085 1	24796	CDS	Cj0019c		putative MCP-domain signal transduction protein	23665	25443	-1	C	CTC	L	synonymous	t	CTt	L
0.0085 1	1277638	CDS	Cj1345c		putative periplasmic protein	1277455	1278849	-1	C	GAC	D	synonymous	t	GAt	D
0.0086 3	1172843	CDS	Cj1244		putative radical SAM domain protein	1171941	1172843	1	G	TAG	*	synonymous	a	TAa	*
0.0086 3	1172843	CDS	Cj1245c		putative membrane protein	1172840	1174036	-1	C	GCC	A	synonymous	t	GCt	A
0.0087 6	68856	CDS	Cj0053c	trmU	tRNA (5-methylaminomethyl-2-thiouridylate)-methyltransferase	68532	69548	-1	A	AGA	R	synonymous	g	AGg	R
0.0088 5	1281904	CDS	Cj1349c		putative fibronectin/fibrinogen-binding protein	1280992	1282299	-1	C	CAC	H	synonymous	t	CAt	H
0.0088 7	982043	CDS	Cj1049c		putative LysE family transporter protein	981655	982254	-1	T	TTT	F	nonsynonymous	c	TcT	S
0.0089 3	249294	CDS	Cj0271		bacterioferritin comigratory protein homolog	248950	249405	1	C	TAC	Y	synonymous	t	TAt	Y
0.0089 3	249300	CDS	Cj0271		bacterioferritin comigratory protein homolog	248950	249405	1	T	GGT	G	synonymous	c	GGc	G
0.0089 3	249301	CDS	Cj0271		bacterioferritin comigratory protein homolog	248950	249405	1	C	CTT	L	MNP	a		
0.0089 3	249303	CDS	Cj0271		bacterioferritin comigratory protein homolog	248950	249405	1	T	CTT	L	nonsynonymous	c	aTc	I
0.0089 4	16990	CDS	Cj0012c	rrc	non-haem iron protein	16756	17403	-1	C	GGC	G	synonymous	t	GGt	G
0.0090 4	68904	CDS	Cj0053c	trmU	tRNA (5-methylaminomethyl-2-thiouridylate)-methyltransferase	68532	69548	-1	A	GTA	V	synonymous	g	GTg	V

0.00909	1227523	CDS	Cj1295		conserved hypothetical protein Cj1295	1226978	1228285	1	C	AAC	N	synonymous	t	AAt	N
0.00909	1227611	CDS	Cj1295		conserved hypothetical protein Cj1295	1226978	1228285	1	T	TTA	L	synonymous	c	cTA	L
0.00927	569066	CDS	Cj0609c		possible periplasmic protein	568946	570124	-1	G	GGG	G	synonymous	a	GGa	G
0.00952	1098810	CDS	Cj1170c	omp50	50 kda outer membrane protein precursor	1098615	1100036	-1	G	CCG	P	synonymous	a	CCa	P
0.00955	1279850	CDS	Cj1346c		1-deoxy-D-xylulose 5-phosphate reductoisomerase	1278851	1279921	-1	C	CCC	P	MNP	t		
0.0096	1175700	CDS	Cj1246c	uvrC	excinuclease ABC subunit C	1174038	1175840	-1	T	GCT	A	synonymous	c	GCc	A
0.0097	59561	CDS	Cj0041	fliK	putative flagellar hook-length control protein	59493	61289	1	C	AGC	S	synonymous	t	AGt	S
0.00971	59567	CDS	Cj0041	fliK	putative flagellar hook-length control protein	59493	61289	1	T	TCT	S	synonymous	c	TCc	S
0.00984	1171296	CDS	Cj1243		uroporphyrinogen decarboxylase	1170922	1171944	1	A	ACA	T	synonymous	t	ACt	T
0.00994	381812	CDS	Cj0415		putative GMC oxidoreductase subunit	381667	383388	1	A	TAT	Y	nonsynonymous	t	TtT	F
0.0101	1174961	CDS	Cj1246c	uvrC	excinuclease ABC subunit C	1174038	1175840	-1	C	CAA	Q	nonsynonymous	a	aAA	K
0.0101	1227559	CDS	Cj1295		conserved hypothetical protein Cj1295	1226978	1228285	1	C	TGC	C	synonymous	t	TGt	C
0.0101	1227562	CDS	Cj1295		conserved hypothetical protein Cj1295	1226978	1228285	1	C	CAC	H	synonymous	t	CAt	H
0.0103	1097976	CDS	Cj1168c		putative integral membrane protein (dedA homolog)	1097694	1098296	-1	T	CAT	H	synonymous	c	CAc	H
0.0104	343224	CDS	Cj0376		putative periplasmic protein	343102	343926	1	A	GCA	A	synonymous	g	GCg	A
0.0104	343282	CDS	Cj0376		putative periplasmic protein	343102	343926	1	G	GTT	V	nonsynonymous	a	aTT	I
0.0105	19999	CDS	Cj0015c		hypothetical protein Cj0015c	19867	21093	-1	T	AGT	S	synonymous	c	AGc	S
0.0105	1518963	CDS	Cj1587c		putative ABC transporter	1518009	1519640	-1	T	GGT	G	synonymous	c	GGc	G
0.0108	1337845	CDS	Cj1403c	gapA	glyceraldehyde 3-phosphate dehydrogenase	1337347	1338345	-1	C	GGC	G	synonymous	t	GGt	G
0.0111	19245	intergenic							T				a		
0.0111	20260	CDS	Cj0015c		hypothetical protein Cj0015c	19867	21093	-1	T	ATT	I	synonymous	c	ATc	I
0.0111	20263	CDS	Cj0015c		hypothetical protein Cj0015c	19867	21093	-1	T	CTT	L	synonymous	g	CTg	L

0.0111	20287	CDS	Cj0015c		hypothetical protein Cj0015c	19867	21093	-1	C	ATC	I	synonymous	t	ATt	I
0.0111	20311	CDS	Cj0015c		hypothetical protein Cj0015c	19867	21093	-1	C	TAC	Y	synonymous	t	TAt	Y
0.0111	20338	CDS	Cj0015c		hypothetical protein Cj0015c	19867	21093	-1	A	ACA	T	synonymous	g	ACg	T
0.0111	21292	CDS	Cj0016		putative transcriptional regulatory protein	21159	21833	1	A	AAA	K	nonsynonymous	g	AgA	R
0.0111	21365	CDS	Cj0016		putative transcriptional regulatory protein	21159	21833	1	G	AAG	K	synonymous	a	AAa	K
0.0112	1508025	CDS	Cj1576c	nuoD	NADH dehydrogenase I chain D	1506903	1508129	-1	T	CAT	H	synonymous	c	CAc	H
0.0113	776002	CDS	Cj0828c	ilvA	threonine dehydratase biosynthetic	775149	776360	-1	G	AGA	R	nonsynonymous	a	AaA	K
0.0114	499190	CDS	Cj0534	sucD	succinyl-coA synthetase alpha chain	498354	499223	1	A	TTA	L	synonymous	g	TTg	L
0.0115	1093747	CDS	Cj1161c		putative cation-transporting ATPase	1091795	1093894	-1	A	AAA	K	nonsynonymous	g	gAA	E
0.0119	854428	CDS	Cj0917c	cstA	putative integral membrane protein (CstA homolog)	852359	854470	-1	G	GCT	A	nonsynonymous	a	aCT	T
0.0119	854428	sig_peptide	Cj0917c	cstA		854369	854470		G				a		
0.012	284984	CDS	Cj0314	lysA	diaminopimelate decarboxylase	284283	285491	1	C	TGC	C	synonymous	t	TGt	C
0.012	1168910	CDS	Cj1240c		putative periplasmic protein	1168448	1169050	-1	A	GAA	E	synonymous	g	GAg	E
0.0121	235702	CDS	Cj0256		putative sulfatase family protein	234599	236137	1	C	AGC	S	synonymous	t	AGt	S
0.0122	420739	CDS	Cj0455c		putative membrane protein	420548	420892	-1	G	GAT	D	nonsynonymous	a	aAT	N
0.0123	136506	CDS	Cj0134	thrB	putative homoserine kinase	135709	136587	1	G	TTG	L	synonymous	a	TTa	L
0.0124	567659	CDS	Cj0608		putative outer membrane efflux protein	567573	568943	1	T	ACT	T	synonymous	c	ACc	T
0.0124	1279895	CDS	Cj1346c		1-deoxy-D-xylulose 5-phosphate reductoisomerase	1278851	1279921	-1	T	AGT	S	synonymous	c	AGc	S
0.0127	1462718	CDS	Cj1530	coaE	putative dephospho-CoA kinase	1462554	1463159	1	T	AGT	S	synonymous	c	AGc	S
0.0127	1631284	CDS	Cj1719c	leuA	2-isopropylmalate synthase	1630393	1631928	-1	T	AGT	S	synonymous	c	AGc	S
0.0129	569645	CDS	Cj0609c		possible periplasmic protein	568946	570124	-1	A	AAA	K	synonymous	g	AAg	K
0.0129	569665	CDS	Cj0609c		possible periplasmic protein	568946	570124	-1	C	CTA	L	synonymous	t	tTA	L

0.0129	569697	CDS	Cj0609c		possible periplasmic protein	568946	570124	-1	C	ACT	T	nonsynonymous	t	Atc	I
0.0129	1093370	CDS	Cj1161c		putative cation-transporting ATPase	1091795	1093894	-1	T	AGT	S	synonymous	c	AGc	S
0.013	594015	CDS	Cj0631c		putative ribonuclease	592495	594429	-1	G	GCC	A	nonsynonymous	a	aCC	T
0.013	596358	CDS	Cj0633		putative periplasmic protein	595574	596656	1	C	GCT	A	nonsynonymous	t	GtT	V
0.0132	48333	CDS	Cj0031		putative type IIS restriction/modification enzyme	46424	50156	1	G	GGG	G	nonsynonymous	a	GaG	E
0.0133	840467	CDS	Cj0899c	thiJ	4-methyl-5(beta-hydroxyethyl)-thiazole monophosphate synthesis protein	839939	840508	-1	A	GAA	E	synonymous	g	GAg	E
0.0133	840574	CDS	Cj0900c		small hydrophobic protein	840501	840680	-1	A	CAT	H	nonsynonymous	g	CgT	R
0.0133	840579	CDS	Cj0900c		small hydrophobic protein	840501	840680	-1	C	TTC	F	synonymous	t	TTt	F
0.0134	138943	CDS	Cj0136	infB	translation initiation factor IF-2	136856	139471	1	G	GAG	E	synonymous	a	GAa	E
0.0134	1002611	CDS	Cj1067	pgsA	CDP-diacylglycerol--glycerol-3-phosphate 3-phosphatidyltransferase	1002567	1003103	1	G	GCG	A	synonymous	a	GCa	A
0.0135	1222333	CDS	Cj1290c	accC	biotin carboxylase	1222303	1223634	-1	T	AAT	N	synonymous	c	AAc	N
0.0135	1507587	CDS	Cj1576c	nuoD	NADH dehydrogenase I chain D	1506903	1508129	-1	G	GAG	E	synonymous	a	GAa	E
0.0135	1507632	CDS	Cj1576c	nuoD	NADH dehydrogenase I chain D	1506903	1508129	-1	G	ACG	T	synonymous	a	ACa	T
0.0136	22250	CDS	Cj0017c	dsbI	disulphide bond formation protein	21854	23380	-1	A	GCA	A	synonymous	g	GCg	A
0.0137	1095494	CDS	Cj1165c		putative integral membrane protein	1095437	1095931	-1	C	ACC	T	synonymous	t	ACt	T
0.0138	139605	CDS	Cj0137	rbfA	putative ribosome-binding factor A	139468	139830	1	C	TAC	Y	synonymous	t	TAt	Y
0.014	593812	CDS	Cj0631c		putative ribonuclease	592495	594429	-1	T	AGT	S	synonymous	c	AGc	S
0.0141	1466535	CDS	Cj1535c	pgi	putative glucose-6-phosphate isomerase	1466289	1467509	-1	T	AGT	S	synonymous	c	AGc	S
0.0142	57568	CDS	Cj0039c	typA	GTP-binding protein TypA homolog	57211	59019	-1	G	GCG	A	synonymous	t	GCt	A
0.0142	1218573	CDS	Cj1286c	upp	uracil phosphoribosyltransferase	1218134	1218760	-1	C	ACA	T	nonsynonymous	t	AtA	I
0.0143	1171195	CDS	Cj1243		uroporphyrinogen decarboxylase	1170922	1171944	1	G	GTT	V	nonsynonymous	a	aTT	I

0.0144	567625	CDS	Cj0608		putative outer membrane efflux protein	567573	568943	1	T	GTA	V	nonsynonymous	c	GcA	A
0.0144	1462804	CDS	Cj1530	coaE	putative dephospho-CoA kinase	1462554	1463159	1	T	ATC	I	nonsynonymous	c	AcC	T
0.0145	47770	CDS	Cj0031		putative type IIS restriction/modification enzyme	46424	50156	1	A	TTA	L	synonymous	g	TTg	L
0.0145	1510596	CDS	Cj1581c		putative peptide ABC-transport system ATP-binding protein	1510567	1511277	-1	A	AAT	N	nonsynonymous	g	gAT	D
0.0146	591921	CDS	Cj0630c	holA	putative DNA polymerase III, delta subunit	591537	592502	-1	A	TTA	L	synonymous	g	TTg	L
0.0146	1218503	CDS	Cj1286c	upp	uracil phosphoribosyltransferase	1218134	1218760	-1	T	AGT	S	synonymous	c	AGc	S
0.0147	569696	CDS	Cj0609c		possible periplasmic protein	568946	570124	-1	T	ACT	T	MNP	c		
0.0147	598325	CDS	Cj0636		NOL1/NOP2/sun family protein	597828	598667	1	A	TCA	S	synonymous	g	TCg	S
0.0147	598640	CDS	Cj0636		NOL1/NOP2/sun family protein	597828	598667	1	T	TTT	F	synonymous	c	TTc	F
0.0147	598649	CDS	Cj0636		NOL1/NOP2/sun family protein	597828	598667	1	G	AAG	K	synonymous	a	AAa	K
0.0148	12467	CDS	Cj0007	gltB	glutamate synthase (NADPH) large subunit	8144	12634	1	A	ACT	T	nonsynonymous	g	gCT	A
0.015	598817	intergenic							A				t		
0.0152	1092035	CDS	Cj1161c		putative cation-transporting ATPase	1091795	1093894	-1	A	GGA	G	synonymous	g	GGg	G
0.0152	1092074	CDS	Cj1161c		putative cation-transporting ATPase	1091795	1093894	-1	G	ATG	M	nonsynonymous	a	ATa	I
0.0152	1344736	CDS	Cj1412c		putative integral membrane protein	1343908	1345029	-1	A	GGA	G	synonymous	g	GGg	G
0.0153	1160982	CDS	Cj1232		hypothetical protein Cj1232	1160803	1161135	1	C	ACC	T	synonymous	t	ACt	T
0.0153	1160985	CDS	Cj1232		hypothetical protein Cj1232	1160803	1161135	1	T	TAT	Y	synonymous	c	TAc	Y
0.0156	1094261	CDS	Cj1163c		putative cation transport protein	1094078	1095028	-1	A	GTA	V	synonymous	g	GTg	V
0.0156	1094627	CDS	Cj1163c		putative cation transport protein	1094078	1095028	-1	C	ATC	I	synonymous	t	ATt	I
0.0156	1499062	CDS	Cj1567c	nuoM	NADH dehydrogenase I chain M	1497800	1499287	-1	G	GTA	V	nonsynonymous	a	aTA	I
0.0156	1508758	CDS	Cj1577c	nuoC	NADH dehydrogenase I chain C	1508131	1508925	-1	G	GAG	E	synonymous	a	GAa	E

0.0157	1463251	CDS	Cj1531	dapF	diaminopimelate epimerase	1463156	1463905	1	T	TGT	C	synonymous	c	TGc	C
0.0157	1463266	CDS	Cj1531	dapF	diaminopimelate epimerase	1463156	1463905	1	T	GGT	G	synonymous	c	GGc	G
0.0157	1463274	CDS	Cj1531	dapF	diaminopimelate epimerase	1463156	1463905	1	C	GCG	A	nonsynonymous	g	GgG	G
0.0157	1463611	CDS	Cj1531	dapF	diaminopimelate epimerase	1463156	1463905	1	G	ATG	M	nonsynonymous	a	ATa	I
0.0157	1641000	CDS	Cj1731c	ruvC	crossover junction endodeoxyribonuclease	1640904	1641386	-1	G	AGG	R	synonymous	a	AGa	R
0.0158	1093793	CDS	Cj1161c		putative cation-transporting ATPase	1091795	1093894	-1	T	AGT	S	synonymous	c	AGc	S
0.016	1221038	CDS	Cj1288c	gltX2	glutamyl-tRNA synthetase	1219994	1221385	-1	C	GCC	A	synonymous	t	GCt	A
0.0161	1650	CDS	Cj0002	dnaN	DNA polymerase III, beta chain	1483	2550	1	A	AAA	K	synonymous	g	AAg	K
0.0162	1279774	CDS	Cj1346c		1-deoxy-D-xylulose 5-phosphate reductoisomerase	1278851	1279921	-1	T	TCC	S	nonsynonymous	g	gCC	A
0.0163	54587	CDS	Cj0036		hypothetical protein Cj0036	53970	55319	1	T	GCT	A	synonymous	c	GCc	A
0.0163	741068	CDS	Cj0790	purU	formyltetrahydrofolate deformylase	740346	741170	1	A	CAA	Q	synonymous	g	CAg	Q
0.0163	987961	CDS	Cj1052c	mutS	putative mismatch repair protein	987019	989229	-1	C	ACC	T	synonymous	t	ACt	T
0.0163	987964	CDS	Cj1052c	mutS	putative mismatch repair protein	987019	989229	-1	C	ATC	I	synonymous	t	ATt	I
0.0163	1464113	CDS	Cj1532		putative periplasmic protein	1463880	1464611	1	C	AAC	N	synonymous	t	AAt	N
0.0163	1464463	CDS	Cj1532		putative periplasmic protein	1463880	1464611	1	A	GAA	E	nonsynonymous	c	GcA	A
0.0166	1085699	intergenic							T				c		
0.0168	1220900	CDS	Cj1288c	gltX2	glutamyl-tRNA synthetase	1219994	1221385	-1	G	GTG	V	synonymous	a	GTa	V
0.0168	1463692	CDS	Cj1531	dapF	diaminopimelate epimerase	1463156	1463905	1	C	CGC	R	synonymous	t	CGt	R
0.0168	1463942	CDS	Cj1532		putative periplasmic protein	1463880	1464611	1	G	AAG	K	synonymous	a	AAa	K
0.0168	1463944	CDS	Cj1532		putative periplasmic protein	1463880	1464611	1	G	AGT	S	nonsynonymous	a	AaT	N
0.0168	1463948	CDS	Cj1532		putative periplasmic protein	1463880	1464611	1	A	GGA	G	synonymous	c	GGc	G
0.0169	593818	CDS	Cj0631c		putative ribonuclease	592495	594429	-1	C	GGC	G	synonymous	t	GGt	G

0.017	572349	CDS	Cj0611c		putative acyltransferase family protein	571122	572498	-1	A	TTA	L	synonymous	g	TTg	L
0.0171	1224463	CDS	Cj1292	dcd	putative deoxycytidine triphosphate deaminase	1224239	1224799	1	C	GGC	G	synonymous	t	GGt	G
0.0171	1224580	CDS	Cj1292	dcd	putative deoxycytidine triphosphate deaminase	1224239	1224799	1	C	GGC	G	synonymous	t	GGt	G
0.0172	381341	CDS	Cj0414		putative oxidoreductase subunit	380937	381665	1	C	GGC	G	synonymous	t	GGt	G
0.0173	834299	CDS	Cj0894c	ispH	4-hydroxy-3-methylbut-2-enyl diphosphate reductase	834005	834838	-1	A	AGA	R	synonymous	g	AGg	R
0.0173	834431	CDS	Cj0894c	ispH	4-hydroxy-3-methylbut-2-enyl diphosphate reductase	834005	834838	-1	A	GTA	V	synonymous	g	GTg	V
0.0173	913874	CDS	Cj0980		putative peptidase	913030	914298	1	A	GAT	D	nonsynonymous	g	GgT	G
0.0175	840368	CDS	Cj0899c	thiJ	4-methyl-5(beta-hydroxyethyl)-thiazole monophosphate synthesis protein	839939	840508	-1	G	TTG	L	synonymous	a	TTa	L
0.0176	1031294	CDS	Cj1097		putative transmembrane transport protein	1030953	1032176	1	T	ACT	T	synonymous	c	ACc	T
0.0176	1333491	CDS	Cj1399c	hydA2	putative Ni/Fe-hydrogenase small subunit	1333149	1334642	-1	A	GAA	E	synonymous	g	GAg	E
0.0177	1035306	CDS	Cj1101		ATP-dependent DNA helicase	1035226	1037301	1	C	GGC	G	synonymous	t	GGt	G
0.0177	1221812	CDS	Cj1289		possible periplasmic protein	1221450	1222265	1	G	AAG	K	synonymous	a	AAa	K
0.0177	1398638	CDS	Cj1462	flgI	flagellar P-ring protein	1398468	1399514	1	T	TCT	S	synonymous	c	TCc	S
0.0178	1391632	CDS	Cj0136	infB	translation initiation factor IF-2	1368556	1394701	1	C	AGC	S	synonymous	t	AGt	S
0.0178	1221122	CDS	Cj1288c	gltX2	glutamyl-tRNA synthetase	1219994	1221385	-1	T	GTT	V	synonymous	c	GTc	V
0.0179	883082	CDS	Cj0942c	secA	preprotein translocase SecA subunit	880838	883426	-1	A	TTA	L	synonymous	g	TTg	L
0.0179	883325	CDS	Cj0942c	secA	preprotein translocase SecA subunit	880838	883426	-1	C	AAC	N	nonsynonymous	a	AAa	K
0.0179	1508521	CDS	Cj1577c	nuoC	NADH dehydrogenase I chain C	1508131	1508925	-1	G	GGG	G	synonymous	a	GGa	G
0.018	834731	CDS	Cj0894c	ispH	4-hydroxy-3-methylbut-2-enyl diphosphate reductase	834005	834838	-1	C	CAC	H	synonymous	t	CAt	H
0.018	834915	CDS	Cj0895c	aroA	3-phosphoshikimate 1-carboxyvinyltransferase	834828	836114	-1	C	ATC	I	synonymous	t	ATt	I
0.0181	1219624	CDS	Cj1287c		malate oxidoreductase	1218757	1219992	-1	G	CCG	P	synonymous	a	CCa	P

0.0181	1219912	CDS	Cj1287c		malate oxidoreductase	1218757	1219992	-1	T	AGT	S	synonymous	c	AGc	S
0.0181	1220036	CDS	Cj1288c	gltX2	glutamyl-tRNA synthetase	1219994	1221385	-1	T	TGT	C	synonymous	c	TGc	C
0.0181	1220049	CDS	Cj1288c	gltX2	glutamyl-tRNA synthetase	1219994	1221385	-1	G	GGG	G	nonsynonymous	a	GaG	E
0.0181	1220261	CDS	Cj1288c	gltX2	glutamyl-tRNA synthetase	1219994	1221385	-1	G	GTG	V	synonymous	a	GTa	V
0.0181	1220300	CDS	Cj1288c	gltX2	glutamyl-tRNA synthetase	1219994	1221385	-1	C	AGC	S	synonymous	t	AGt	S
0.0181	1220516	CDS	Cj1288c	gltX2	glutamyl-tRNA synthetase	1219994	1221385	-1	T	GAT	D	synonymous	c	GAc	D
0.0181	1221719	CDS	Cj1289		possible periplasmic protein	1221450	1222265	1	C	ACC	T	synonymous	g	ACg	T
0.0182	1464284	CDS	Cj1532		putative periplasmic protein	1463880	1464611	1	C	AAC	N	synonymous	t	AAt	N
0.0182	1464393	CDS	Cj1532		putative periplasmic protein	1463880	1464611	1	G	GAC	D	nonsynonymous	t	tAC	Y
0.0182	1464488	CDS	Cj1532		putative periplasmic protein	1463880	1464611	1	G	GGG	G	synonymous	a	GGa	G
0.0183	838376	CDS	Cj0896c	pheT	phenylalanyl-tRNA synthetase beta chain	836111	838432	-1	A	TCA	S	synonymous	g	TCg	S
0.0183	1118479	CDS	Cj1190c	cetA	bipartate energy taxis response protein cetA	1117705	1119084	-1	T	GGT	G	synonymous	c	GGc	G
0.0183	1398806	CDS	Cj1462	flgI	flagellar P-ring protein	1398468	1399514	1	T	GGT	G	synonymous	c	GGc	G
0.0183	1634533	CDS	Cj1724c		putative GTP cyclohydrolase I	1634296	1634679	-1	A	GCA	A	synonymous	g	GCg	A
0.0184	380444	CDS	Cj0413		putative periplasmic protein	380097	380801	1	A	TCA	S	synonymous	g	TCg	S
0.0184	1464365	CDS	Cj1532		putative periplasmic protein	1463880	1464611	1	A	AAA	K	synonymous	g	AAg	K
0.0184	1519035	CDS	Cj1587c		putative ABC transporter	1518009	1519640	-1	T	GAT	D	synonymous	c	GAc	D
0.0185	1172012	CDS	Cj1244		putative radical SAM domain protein	1171941	1172843	1	T	AGT	S	synonymous	c	AGc	S
0.0188	57538	CDS	Cj0039c	typA	GTP-binding protein TypA homolog	57211	59019	-1	A	TTA	L	synonymous	g	TTg	L
0.0189	1291933	CDS	Cj0128c		putative inositol monophosphatase family protein	1290778	1298000	-1	C	GCA	A	nonsynonymous	t	GtA	V
0.019	599378	intergenic							G				a		

0.0192	32554	CDS	Cj0025c		putative sodium:dicarboxylate family transmembrane symporter	32134	33519	-1	C	ACC	T	synonymous	t	ACt	T
0.0195	741452	CDS	Cj0791c		putative aminotransferase	741167	74244 1	-1	C	TAC	Y	synonymous	t	TAt	Y
0.0195	883412	CDS	Cj0942c	secA	preprotein translocase SecA subunit	880838	88342 6	-1	T	GCT	A	synonymous	c	GCc	A
0.0197	48429	CDS	Cj0031		putative type IIS restriction/modification enzyme	46424	50156	1	A	AAC	N	nonsynonymous	g	AgC	S
0.02	1560440	CDS	Cj1634c	aroC	chorismate synthase	1559723	15608 11	-1	C	AGC	S	synonymous	t	AGt	S
0.0202	367284	CDS	Cj0402	glyA	serine hydroxymethyltransferase	367219	368463	1	T	TGT	C	synonymous	c	TGc	C
0.0202	367359	CDS	Cj0402	glyA	serine hydroxymethyltransferase	367219	368463	1	A	ACA	T	synonymous	g	ACg	T
0.0203	987256	CDS	Cj1052c	mutS	putative mismatch repair protein	987019	989229	-1	T	CTT	L	synonymous	a	CTa	L
0.0203	987262	CDS	Cj1052c	mutS	putative mismatch repair protein	987019	989229	-1	T	ACT	T	MNP	c		
0.0203	987263	CDS	Cj1052c	mutS	putative mismatch repair protein	987019	989229	-1	C	ACT	T	nonsynonymous	g	Agc	S
0.0203	987273	CDS	Cj1052c	mutS	putative mismatch repair protein	987019	989229	-1	G	GCA	A	nonsynonymous	a	aCA	T
0.0203	987274	CDS	Cj1052c	mutS	putative mismatch repair protein	987019	989229	-1	G	GTG	V	synonymous	a	GTa	V
0.0203	987283	CDS	Cj1052c	mutS	putative mismatch repair protein	987019	989229	-1	C	AGC	S	synonymous	t	AGt	S
0.0203	987316	CDS	Cj1052c	mutS	putative mismatch repair protein	987019	989229	-1	C	AGC	S	synonymous	t	AGt	S
0.0205	1386597	CDS	Cj1447c	kpsT	capsule polysaccharide export ATP-binding protein	1386264	13869 26	-1	G	GTG	V	synonymous	a	GTa	V
0.0206	136939	CDS	Cj0136	infB	translation initiation factor IF-2	136856	139471	1	A	GGA	G	synonymous	g	GGg	G
0.0206	1517776	CDS	Cj1586	cgb	single domain haemoglobin	1517567	15179 89	1	G	TTG	L	synonymous	a	TTa	L
0.0207	127995	CDS	Cj0126c		conserved hypothetical protein Cj0126c	127776	12823 7	-1	T	GAT	D	nonsynonymous	a	GAa	E
0.0207	1220251	CDS	Cj1288c	gltX2	glutamyl-tRNA synthetase	1219994	12213 85	-1	G	GTT	V	nonsynonymous	a	aTT	I
0.0207	1220405	CDS	Cj1288c	gltX2	glutamyl-tRNA synthetase	1219994	12213 85	-1	A	AAA	K	synonymous	g	AAg	K
0.0207	1220456	CDS	Cj1288c	gltX2	glutamyl-tRNA synthetase	1219994	12213 85	-1	T	AAT	N	synonymous	c	AAc	N

0.0207	1220542	CDS	Cj1288c	gltX2	glutamyl-tRNA synthetase	1219994	1221385	-1	C	CTA	L	synonymous	t	tTA	L
0.0207	1220777	CDS	Cj1288c	gltX2	glutamyl-tRNA synthetase	1219994	1221385	-1	G	GGG	G	synonymous	a	GGa	G
0.0207	1220780	CDS	Cj1288c	gltX2	glutamyl-tRNA synthetase	1219994	1221385	-1	T	CGT	R	synonymous	c	CGc	R
0.0208	1094246	CDS	Cj1163c		putative cation transport protein	1094078	1095028	-1	T	ATT	I	synonymous	c	ATc	I
0.0209	1517758	CDS	Cj1586	cgb	single domain haemoglobin	1517567	1517989	1	T	GCT	A	synonymous	c	GCc	A
0.021	1222190	CDS	Cj1289		possible periplasmic protein	1221450	1222265	1	T	AGT	S	synonymous	c	AGc	S
0.0211	1183636	CDS	Cj1253	pnp	polyribonucleotide nucleotidyltransferase	1183594	1185753	1	G	GAT	D	nonsynonymous	a	aAT	N
0.0211	1189225	CDS	Cj1259	porA	major outer membrane protein	1189121	1190395	1	A	TCA	S	synonymous	t	TCt	S
0.0211	1189228	CDS	Cj1259	porA	major outer membrane protein	1189121	1190395	1	T	GGT	G	synonymous	a	GGa	G
0.0214	56999	CDS	Cj0038c		putative poly(A) polymerase family protein	56564	57211	-1	A	TTA	L	synonymous	g	TTg	L
0.0214	294198	CDS	Cj0322	perR	peroxide stress regulator	293931	294341	1	A	ACC	T	nonsynonymous	g	gCC	A
0.0218	69123	CDS	Cj0053c	trmU	tRNA (5-methylaminomethyl-2-thiouridylate)-methyltransferase	68532	69548	-1	T	AGT	S	synonymous	c	AGc	S
0.0218	69129	CDS	Cj0053c	trmU	tRNA (5-methylaminomethyl-2-thiouridylate)-methyltransferase	68532	69548	-1	G	TTG	L	synonymous	a	TTa	L
0.0218	586219	CDS	Cj0626	hypE	hydrogenase isoenzymes formation protein	586190	587164	1	C	GGC	G	synonymous	t	GGt	G
0.0218	1278316	CDS	Cj1345c		putative periplasmic protein	1277455	1278849	-1	G	TCG	S	synonymous	a	TCa	S
0.0219	600188	CDS	Cj0639c	adk	adenylate kinase	599909	600487	-1	C	GAC	D	synonymous	t	GAt	D
0.022	1171254	CDS	Cj1243		uroporphyrinogen decarboxylase	1170922	1171944	1	A	GCA	A	synonymous	g	GCg	A
0.0223	129099	CDS	Cj0128c		putative inositol monophosphatase family protein	129078	129800	-1	A	TTA	L	synonymous	g	TTg	L
0.0229	741236	CDS	Cj0791c		putative aminotransferase	741167	742441	-1	G	ACG	T	synonymous	a	ACa	T
0.0229	741329	CDS	Cj0791c		putative aminotransferase	741167	742441	-1	T	TGT	C	synonymous	c	TGc	C

0.0229	742019	CDS	Cj0791c		putative aminotransferase	741167	742441	-1	G	TTG	L	synonymous	a	TTa	L
0.0231	143646	CDS	Cj0140		hypothetical protein Cj0140	142633	143964	1	T	AAT	N	nonsynonymous	a	AAa	K
0.0232	573358	CDS	Cj0613	pstS	putative periplasmic phosphate binding protein	573224	574219	1	C	ATC	I	synonymous	t	ATt	I
0.0232	109556	CDS	Cj1165c		putative integral membrane protein	109543	109593	-1	A	TTA	L	synonymous	g	TTg	L
0.0233	446729	CDS	Cj0479	rpoC	DNA-directed RNA polymerase beta' chain	443121	447674	1	C	GCC	A	synonymous	t	GCt	A
0.0234	218778	CDS	Cj0236c		putative integral membrane protein	218523	219218	-1	T	GGT	G	synonymous	c	GGc	G
0.0234	234077	CDS	Cj0255c	exoA	exodeoxyribonuclease	233723	234481	-1	G	AAG	K	synonymous	a	AAa	K
0.0234	103160	CDS	Cj1097		putative transmembrane transport protein	103095	103217	1	T	CTT	L	synonymous	c	CTc	L
0.0235	24372	CDS	Cj0019c		putative MCP-domain signal transduction protein	23665	25443	-1	C	CAT	H	nonsynonymous	t	tAT	Y
0.0238	132285	CDS	Cj1385	katA	catalase	132252	132395	1	C	GAC	D	synonymous	t	GAt	D
0.0238	132288	CDS	Cj1385	katA	catalase	132252	132395	1	G	TTG	L	synonymous	a	TTa	L
0.0238	132300	CDS	Cj1385	katA	catalase	132252	132395	1	C	AAC	N	synonymous	t	AAt	N
0.0238	132306	CDS	Cj1385	katA	catalase	132252	132395	1	G	CCG	P	synonymous	a	CCa	P
0.0238	132309	CDS	Cj1385	katA	catalase	132252	132395	1	A	ACA	T	synonymous	t	ACt	T
0.0238	132312	CDS	Cj1385	katA	catalase	132252	132395	1	C	TTC	F	synonymous	t	TTt	F
0.0238	132315	CDS	Cj1385	katA	catalase	132252	132395	1	C	TTC	F	synonymous	t	TTt	F
0.0239	151961	CDS	Cj1587c		putative ABC transporter	151800	151964	-1	G	AAG	K	synonymous	a	AAa	K
0.0239	151961	sig_peptide	Cj1587c			151950	151964		G				a		
0.024	55550	CDS	Cj0037c		putative cytochrome C	55343	56386	-1	A	CTA	L	synonymous	g	CTg	L
0.024	380756	CDS	Cj0413		putative periplasmic protein	380097	380801	1	T	TGT	C	synonymous	c	TGc	C
0.024	149354	CDS	Cj1565c	pflA	paralysed flagellum protein	149403	149640	-1	A	GAA	E	synonymous	g	GAg	E
0.0241	100037	CDS	Cj1062		putative CinA-like protein	999629	1000717	1	G	GTA	V	nonsynonymous	a	aTA	I

0.0242	235741	CDS	Cj0256		putative sulfatase family protein	2345999	2361377	1	A	GAA	E	synonymous	g	GAg	E
0.0242	1143058	CDS	Cj1214c		putative exporting protein	1143019	1143744	-1	G	AGG	R	synonymous	a	AGa	R
0.0243	29639	CDS	Cj0023	purB	adenylosuccinate lyase	28382	29710	1	G	GTA	V	nonsynonymous	a	aTA	I
0.0246	1039044	CDS	Cj1104	ispE	putative 4-diphosphocytidyl-2-C-methyl-D-erythritol kinase	1038334	1039101	1	G	TTG	L	synonymous	a	TTa	L
0.0249	234239	CDS	Cj0255c	exoA	exodeoxyribonuclease	233723	234481	-1	C	TGC	C	synonymous	t	TGt	C
0.025	1147239	CDS	Cj1219c		putative periplasmic protein	1146492	1149032	-1	T	GAT	D	synonymous	c	GAc	D
0.025	1454759	CDS	Cj1519	moeA2	putative molybdopterin biosynthesis protein	1453815	1455005	1	A	CAA	Q	synonymous	g	CAg	Q
0.025	1500968	CDS	Cj1568c	nuoL	NADH dehydrogenase I chain L	1499289	1501079	-1	T	TTA	L	synonymous	c	cTA	L
0.0252	580865	CDS	Cj0621		hypothetical protein Cj0621	580326	581846	1	G	AAG	K	synonymous	a	AAa	K
0.0252	987682	CDS	Cj1052c	mutS	putative mismatch repair protein	987019	989229	-1	T	GTT	V	synonymous	c	GTc	V
0.0255	1017381	CDS	Cj1085c	mfd	transcription-repair coupling factor	1015629	1018565	-1	G	TTG	L	synonymous	a	TTa	L
0.0255	1323377	CDS	Cj1385	katA	catalase	1322526	1323950	1	A	GTA	V	synonymous	g	GTg	V
0.0258	1155451	CDS	Cj1226c		putative two-component sensor (histidine kinase)	1154267	1155514	-1	A	ATA	I	nonsynonymous	g	gTA	V
0.0258	1284356	CDS	Cj1352	ceuB	enterochelin uptake permease	1284008	1284976	1	A	ATA	I	nonsynonymous	g	gTA	V
0.0259	154350	CDS	Cj0150c		aminotransferase	153670	154872	-1	T	TGT	C	nonsynonymous	c	cGT	R
0.0259	154743	CDS	Cj0150c		aminotransferase	153670	154872	-1	A	AAT	N	nonsynonymous	g	gAT	D
0.0259	595141	CDS	Cj0632	ilvC	ketol-acid reductoisomerase	594548	595570	1	T	GCT	A	synonymous	c	GCc	A
0.026	1098720	CDS	Cj1170c	omp50	50 kda outer membrane protein precursor	1098615	1100036	-1	C	ACC	T	synonymous	t	ACt	T
0.026	1322831	CDS	Cj1385	katA	catalase	1322526	1323950	1	A	GAA	E	synonymous	g	GAg	E
0.026	1322834	CDS	Cj1385	katA	catalase	1322526	1323950	1	C	CGC	R	synonymous	t	CGt	R
0.026	1322852	CDS	Cj1385	katA	catalase	1322526	1323950	1	T	GCT	A	synonymous	c	GCc	A
0.026	1322927	CDS	Cj1385	katA	catalase	1322526	1323950	1	T	GCT	A	synonymous	g	GCg	A

0.026	1322936	CDS	Cj1385	katA	catalase	1322526	1323950	1	C	TTC	F	synonymous	t	TTt	F
0.0261	124404	CDS	Cj0121		conserved hypothetical protein Cj0121	124258	124665	1	A	GTA	V	synonymous	g	GTg	V
0.0261	127596	CDS	Cj0125c		hypothetical protein Cj0125c	127402	127764	-1	G	GCT	A	nonsynonymous	a	aCT	T
0.0261	379930	CDS	Cj0412		putative ATP/GTP binding protein	378268	380097	1	A	ACT	T	nonsynonymous	g	gCT	A
0.0261	924505	CDS	Cj0992c		oxygen-independent coproporphyrinogen III oxidase	923422	924777	-1	C	ACC	T	synonymous	t	ACt	T
0.0262	1171137	CDS	Cj1243		uroporphyrinogen decarboxylase	1170922	1171944	1	C	ATC	I	synonymous	t	ATt	I
0.0262	1171564	CDS	Cj1243		uroporphyrinogen decarboxylase	1170922	1171944	1	G	GAA	E	nonsynonymous	a	aAA	K
0.0262	1196771	CDS	Cj1266c	hydB	Ni/Fe-hydrogenase large subunit	1195496	1197211	-1	G	GCG	A	synonymous	a	GCa	A
0.0264	1035288	CDS	Cj1101		ATP-dependent DNA helicase	1035226	1037301	1	C	GCC	A	synonymous	t	GCt	A
0.0265	573421	CDS	Cj0613	pstS	putative periplasmic phosphate binding protein	573224	574219	1	T	TGT	C	synonymous	c	TGc	C
0.0265	6161718	pseudogene	Cj0654c		pseudogene (putative transmembrane transport protein)	6150104	6164805	-1	G	GGG	G	synonymous	a	GGa	G
0.0265	1519374	CDS	Cj1587c		putative ABC transporter	1518009	1519640	-1	C	ATC	I	synonymous	a	ATa	I
0.0267	19227	CDS	Cj0013	ilvD	dihydroxy-acid dehydratase	17563	19239	1	A	TTA	L	synonymous	g	TTg	L
0.0267	1057826	CDS	Cj1125c	pglA	GalNAc transferase	1056839	1057969	-1	C	GGC	G	synonymous	t	GGt	G
0.0269	1224533	CDS	Cj1292	dcd	putative deoxycytidine triphosphate deaminase	1224239	1224799	1	A	AAT	N	nonsynonymous	g	gAT	D
0.0269	1224905	CDS	Cj1293	pseB	UDP-GlcNAc-specific C4,6 dehydratase/C5 epimerase	1224849	1225853	1	T	ACT	T	synonymous	c	ACc	T
0.0269	1224944	CDS	Cj1293	pseB	UDP-GlcNAc-specific C4,6 dehydratase/C5 epimerase	1224849	1225853	1	G	AAG	K	synonymous	a	AAa	K
0.0269	1224947	CDS	Cj1293	pseB	UDP-GlcNAc-specific C4,6 dehydratase/C5 epimerase	1224849	1225853	1	T	ATT	I	synonymous	c	ATc	I
0.027	1166341	CDS	Cj1237c		putative phosphatase	1165564	1166538	-1	G	GCG	A	synonymous	a	GCa	A
0.0272	247040	CDS	Cj0268c		putative transmembrane protein	246555	247643	-1	A	ACT	T	nonsynonymous	g	gCT	A
0.0274	201	CDS	Cj0001	dnaA	chromosomal replication initiator protein	1	1323	1	C	AGC	S	synonymous	t	AGt	S

0.0281	1039574	CDS	Cj1106		putative periplasmic thioredoxin	1039561	1040163	1	C	TCA	S	nonsynonymous	t	TtA	L
0.0281	1039574	sig_peptide	Cj1106			1039561	1039620		C				t		
0.0282	1338004	CDS	Cj1403c	gapA	glyceraldehyde 3-phosphate dehydrogenase	1337347	1338345	-1	A	GTA	V	synonymous	g	GTg	V
0.0286	3514557	CDS	Cj0386		putative GTP-binding protein	3514456	3528288	1	T	ATT	I	synonymous	c	ATc	I
0.0286	1462164	CDS	Cj1529c	purM	phosphoribosylformylglycin amidine cyclo-ligase	1461504	1462493	-1	C	CTC	L	synonymous	t	CTt	L
0.0286	1462170	CDS	Cj1529c	purM	phosphoribosylformylglycin amidine cyclo-ligase	1461504	1462493	-1	C	GCC	A	synonymous	t	GCt	A
0.0289	25876	CDS	Cj0020c		cytochrome C551 peroxidase	25433	26347	-1	G	GTT	V	nonsynonymous	a	aTc	I
0.0289	1094138	CDS	Cj1163c		putative cation transport protein	1094078	1095028	-1	G	TTG	L	MNP	a		
0.0289	1094140	CDS	Cj1163c		putative cation transport protein	1094078	1095028	-1	T	TTG	L	synonymous	c	cTa	L
0.029	151313	CDS	Cj0146c	trxB	thioredoxin reductase	150638	151576	-1	T	AGT	S	synonymous	c	AGc	S
0.0291	593071	CDS	Cj0631c		putative ribonuclease	592495	594429	-1	G	AAG	K	synonymous	a	AAa	K
0.0291	593132	CDS	Cj0631c		putative ribonuclease	592495	594429	-1	T	GTA	V	nonsynonymous	c	GcA	A
0.0292	30211	CDS	Cj0024	nrdA	ribonucleoside-diphosphate reductase alpha chain	29726	32095	1	A	CAA	Q	synonymous	g	CAg	Q
0.0293	1166623	CDS	Cj1238	pdxJ	pyridoxal phosphate biosynthetic protein	1166594	1167367	1	T	ATT	I	synonymous	c	ATc	I
0.0294	155227	CDS	Cj0151c		putative periplasmic protein	154884	155690	-1	T	GTA	V	nonsynonymous	c	GcA	A
0.0294	586097	CDS	Cj0625	hypD	hydrogenase isoenzymes formation protein	585102	586193	1	C	GGC	G	synonymous	t	GGt	G
0.0297	644389	CDS	Cj0689	ackA	acetate kinase	643409	644599	1	C	GGC	G	synonymous	t	GGt	G
0.0297	1464896	CDS	Cj1533c		putative helix-turn-helix containing protein	1464608	1465645	-1	A	TCA	S	synonymous	g	TCg	S
0.0297	1465061	CDS	Cj1533c		putative helix-turn-helix containing protein	1464608	1465645	-1	C	CTC	L	synonymous	t	CTt	L
0.0297	1508241	CDS	Cj1577c	nuoC	NADH dehydrogenase I chain C	1508131	1508925	-1	G	GGT	G	nonsynonymous	a	aGT	S
0.0297	1508317	CDS	Cj1577c	nuoC	NADH dehydrogenase I chain C	1508131	1508925	-1	A	GAA	E	synonymous	g	GAg	E
0.0299	393	CDS	Cj0001	dnaA	chromosomal replication initiator protein	1	1323	1	A	AAA	K	synonymous	g	AAg	K

0.0299	1278040	CDS	Cj1345c		putative periplasmic protein	1277455	1278849	-1	A	GTA	V	synonymous	g	GTg	V
0.0302	1117454	CDS	Cj1189c	cetB	bipartate energy taxis response protein cetB	1117190	1117687	-1	T	AAT	N	synonymous	c	AAc	N
0.0303	1224706	CDS	Cj1292	dcd	putative deoxycytidine triphosphate deaminase	1224239	1224799	1	G	TTG	L	synonymous	a	TTa	L
0.0305	496883	CDS	Cj0532	mdh	malate dehydrogenase	496278	497180	1	G	GAG	E	nonsynonymous	t	GAt	D
0.0305	1322684	CDS	Cj1385	katA	catalase	1322526	1323950	1	T	GCT	A	synonymous	c	GCc	A
0.0305	1322696	CDS	Cj1385	katA	catalase	1322526	1323950	1	A	GGA	G	synonymous	g	GGg	G
0.0305	1322705	CDS	Cj1385	katA	catalase	1322526	1323950	1	C	GGC	G	synonymous	a	GGa	G
0.0305	1322717	CDS	Cj1385	katA	catalase	1322526	1323950	1	T	ATT	I	synonymous	c	ATc	I
0.0305	1322720	CDS	Cj1385	katA	catalase	1322526	1323950	1	A	ACA	T	synonymous	c	ACc	T
0.0305	1322723	CDS	Cj1385	katA	catalase	1322526	1323950	1	A	GCA	A	synonymous	t	GCt	A
0.0305	1322726	CDS	Cj1385	katA	catalase	1322526	1323950	1	C	GAC	D	synonymous	t	GAt	D
0.0305	1322730	CDS	Cj1385	katA	catalase	1322526	1323950	1	A	AGT	S	MNP	t		
0.0305	1322731	CDS	Cj1385	katA	catalase	1322526	1323950	1	G	AGT	S	synonymous	c	tcT	S
0.0305	1322741	CDS	Cj1385	katA	catalase	1322526	1323950	1	C	ACC	T	synonymous	t	ACt	T
0.0305	1322744	CDS	Cj1385	katA	catalase	1322526	1323950	1	A	AAA	K	synonymous	g	AAg	K
0.0305	1322747	CDS	Cj1385	katA	catalase	1322526	1323950	1	T	GCT	A	synonymous	a	GCa	A
0.0305	1322753	CDS	Cj1385	katA	catalase	1322526	1323950	1	T	ATT	I	synonymous	a	ATa	I
0.0305	1322765	CDS	Cj1385	katA	catalase	1322526	1323950	1	C	GGC	G	synonymous	a	GGa	G
0.0305	1322769	CDS	Cj1385	katA	catalase	1322526	1323950	1	G	GTT	V	MNP	a		
0.0305	1322771	CDS	Cj1385	katA	catalase	1322526	1323950	1	T	GTT	V	nonsynonymous	a	aTa	I
0.0305	1322777	CDS	Cj1385	katA	catalase	1322526	1323950	1	A	CCA	P	synonymous	t	CCt	P
0.0305	1322778	CDS	Cj1385	katA	catalase	1322526	1323950	1	T	TTA	L	MNP	c		

0.0305	1322780	CDS	Cj1385	katA	catalase	1322526	1323950	1	A	TTA	L	synonymous	t	cTt	L
0.0305	1322783	CDS	Cj1385	katA	catalase	1322526	1323950	1	T	TTT	F	synonymous	c	TTc	F
0.0305	1322784	CDS	Cj1385	katA	catalase	1322526	1323950	1	T	TTA	L	synonymous	c	cTA	L
0.0305	1322795	CDS	Cj1385	katA	catalase	1322526	1323950	1	A	TCA	S	synonymous	t	TCt	S
0.0305	1322798	CDS	Cj1385	katA	catalase	1322526	1323950	1	A	ACA	T	synonymous	t	ACt	T
0.0305	1322801	CDS	Cj1385	katA	catalase	1322526	1323950	1	A	GTA	V	synonymous	t	GTt	V
0.0306	1374466	CDS	Cj0136	infB	translation initiation factor IF-2	1368566	1394701	1	T	GTT	V	synonymous	c	GTc	V
0.0306	1375339	CDS	Cj0136	infB	translation initiation factor IF-2	1368566	1394701	1	A	TTA	L	synonymous	g	TTg	L
0.0306	1177176	CDS	Cj1248	guaA	GMP synthase (glutamine-hydrolyzing)	1176466	1178001	1	T	AGT	S	synonymous	c	AGc	S
0.0311	4747	CDS	Cj0003	gyrB	DNA gyrase subunit B	2579	4888	1	C	AGC	S	synonymous	t	AGt	S
0.0312	1168203	CDS	Cj1239	pdxA	putative 4-hydroxythreonine-4-phosphate dehydrogenase	1167364	1168458	1	G	GTG	V	synonymous	a	GTa	V
0.0312	1454837	CDS	Cj1519	moeA2	putative molybdopterin biosynthesis protein	1453815	1455005	1	C	TGC	C	synonymous	t	TGt	C
0.0313	999727	CDS	Cj1062		putative CinA-like protein	999629	1000717	1	T	ATT	I	synonymous	c	ATc	I
0.0314	384322	CDS	Cj0418c		hypothetical protein Cj0418c	383947	384651	-1	C	GGC	G	synonymous	t	GGt	G
0.0314	384409	CDS	Cj0418c		hypothetical protein Cj0418c	383947	384651	-1	A	GCA	A	synonymous	g	GCg	A
0.0318	5232	CDS	Cj0004c		monohaem cytochrome C	4916	5257	-1	C	GCA	A	nonsynonymous	t	GtA	V
0.0318	5232	sig_peptide	Cj0004c			5207	5257		C				t		
0.0318	9346	CDS	Cj0007	gltB	glutamate synthase (NADPH) large subunit	8144	12634	1	T	GCT	A	synonymous	c	GCc	A
0.0318	129021	CDS	Cj0127c	accD	acetyl-coenzyme A carboxylase carboxyl transferase subunit beta	128227	129069	-1	G	GTT	V	nonsynonymous	a	aTT	I
0.0318	129022	CDS	Cj0127c	accD	acetyl-coenzyme A carboxylase carboxyl transferase subunit beta	128227	129069	-1	C	AGC	S	synonymous	t	AGt	S
0.0318	340764	CDS	Cj0372		putative glutathionylspermidine synthase	340018	341187	1	T	GCT	A	synonymous	c	GCc	A

0.0319	586408	CDS	Cj0626	hypE	hydrogenase isoenzymes formation protein	586190	587164	1	T	AAT	N	synonymous	c	AAc	N
0.032	1507677	CDS	Cj1576c	nuoD	NADH dehydrogenase I chain D	1506903	1508129	-1	G	GAG	E	synonymous	a	GAa	E
0.0321	249956	CDS	Cj0272		conserved hypothetical protein Cj0272	249405	250496	1	T	TTT	F	synonymous	c	TTc	F
0.0321	340458	CDS	Cj0372		putative glutathionylspermidine synthase	340018	341187	1	C	AGC	S	synonymous	t	AGt	S
0.0322	1185132	CDS	Cj1253	pnp	polyribonucleotide nucleotidyltransferase	1183594	1185753	1	C	GAC	D	synonymous	t	GAt	D
0.0322	1185147	CDS	Cj1253	pnp	polyribonucleotide nucleotidyltransferase	1183594	1185753	1	A	GGA	G	synonymous	t	GGt	G
0.0322	1283469	CDS	Cj1351	pldA	phospholipase A	1282933	1283922	1	A	TTA	L	synonymous	g	TTg	L
0.0323	139270	CDS	Cj0136	infB	translation initiation factor IF-2	136856	139471	1	T	AAT	N	synonymous	c	AAc	N
0.0323	1454798	CDS	Cj1519	moeA2	putative molybdopterin biosynthesis protein	1453815	1455005	1	T	AGT	S	synonymous	c	AGc	S
0.0323	1454942	CDS	Cj1519	moeA2	putative molybdopterin biosynthesis protein	1453815	1455005	1	G	GTG	V	synonymous	a	GTa	V
0.0325	918018	CDS	Cj0984		conserved hypothetical protein Cj0984	917836	918576	1	T	AAT	N	synonymous	c	AAc	N
0.0326	494268	CDS	Cj0531	icd	isocitrate dehydrogenase	494077	496281	1	G	GAG	E	synonymous	a	GAa	E
0.0326	494526	CDS	Cj0531	icd	isocitrate dehydrogenase	494077	496281	1	C	GAC	D	synonymous	t	GAt	D
0.0328	1225226	CDS	Cj1293	pseB	UDP-GlcNAc-specific C4,6 dehydratase/C5 epimerase	1224849	1225853	1	T	GAT	D	synonymous	c	GAc	D
0.033	1279883	CDS	Cj1346c		1-deoxy-D-xylulose 5-phosphate reductoisomerase	1278851	1279921	-1	T	AAT	N	synonymous	c	AAc	N
0.0331	340671	CDS	Cj0372		putative glutathionylspermidine synthase	340018	341187	1	T	AGT	S	synonymous	c	AGc	S
0.0331	586222	CDS	Cj0626	hypE	hydrogenase isoenzymes formation protein	586190	587164	1	T	GGT	G	synonymous	c	GGc	G
0.0332	994343	CDS	Cj1058c	guaB	inosine-5'-monophosphate dehydrogenase	993848	995305	-1	G	GTG	V	synonymous	a	GTa	V
0.0333	3298	CDS	Cj0003	gyrB	DNA gyrase subunit B	2579	4888	1	T	AGT	S	synonymous	c	AGc	S
0.0334	585983	CDS	Cj0625	hypD	hydrogenase isoenzymes formation protein	585102	586193	1	C	TAC	Y	synonymous	t	TAt	Y
0.0334	1010767	CDS	Cj1077	ctsT	putative periplasmic protein	1010738	1011040	1	C	AGC	S	synonymous	t	AGt	S

0.0334	1010823	CDS	Cj1077	ctsT	putative periplasmic protein	1010738	1011040	1	T	ATT	I	nonsynonymous	c	AcT	T
0.0334	1011144	CDS	Cj1078		putative periplasmic protein	1011037	1011699	1	T	TTT	F	synonymous	c	TTc	F
0.0334	1011402	CDS	Cj1078		putative periplasmic protein	1011037	1011699	1	C	AAC	N	synonymous	t	AAt	N
0.0334	1011411	CDS	Cj1078		putative periplasmic protein	1011037	1011699	1	A	GAA	E	synonymous	g	GAg	E
0.0334	1011456	CDS	Cj1078		putative periplasmic protein	1011037	1011699	1	T	TCT	S	synonymous	c	TCc	S
0.0334	1011462	CDS	Cj1078		putative periplasmic protein	1011037	1011699	1	G	TTG	L	synonymous	a	TTa	L
0.0334	1011478	CDS	Cj1078		putative periplasmic protein	1011037	1011699	1	G	GCA	A	nonsynonymous	a	aCA	T
0.0337	1187557	CDS	Cj1257c		putative efflux pump	1187176	1188360	-1	C	GCC	A	synonymous	t	GCt	A
0.0337	1187911	CDS	Cj1257c		putative efflux pump	1187176	1188360	-1	G	GGG	G	synonymous	a	GGa	G
0.0339	606580	CDS	Cj0644		putative TatD-related deoxyribonuclease protein	606050	606865	1	G	TTG	L	synonymous	a	TTa	L
0.0339	1464672	CDS	Cj1533c		putative helix-turn-helix containing protein	1464608	1465645	-1	A	AAT	N	nonsynonymous	g	AgT	S
0.034	15254	CDS	Cj0009	gltD	glutamate synthase (NADPH) small subunit	14398	15843	1	T	ATG	M	nonsynonymous	c	AcG	T
0.034	250289	CDS	Cj0272		conserved hypothetical protein Cj0272	249405	250496	1	A	GCA	A	synonymous	g	GCg	A
0.034	250344	CDS	Cj0272		conserved hypothetical protein Cj0272	249405	250496	1	G	GTT	V	nonsynonymous	a	aTT	I
0.034	250356	CDS	Cj0272		conserved hypothetical protein Cj0272	249405	250496	1	A	ATT	I	nonsynonymous	g	gTT	V
0.034	250360	CDS	Cj0272		conserved hypothetical protein Cj0272	249405	250496	1	A	AAA	K	nonsynonymous	g	AgA	R
0.034	1043541	CDS	Cj1110c		putative MCP-type signal transduction protein	1043223	1044512	-1	A	CAA	Q	synonymous	g	CAg	Q
0.0341	20737	CDS	Cj0015c		hypothetical protein Cj0015c	19867	21093	-1	T	GGT	G	synonymous	c	GGc	G
0.0341	59579	CDS	Cj0041	fliK	putative flagellar hook-length control protein	59493	61289	1	T	ACT	T	synonymous	g	ACg	T
0.0341	1279124	CDS	Cj1346c		1-deoxy-D-xylulose 5-phosphate reductoisomerase	1278851	1279921	-1	C	AGC	S	synonymous	t	AGt	S
0.0342	596293	CDS	Cj0633		putative periplasmic protein	595574	596656	1	T	TAT	Y	synonymous	c	TAc	Y
0.0343	593577	CDS	Cj0631c		putative ribonuclease	592495	594429	-1	C	CTA	L	synonymous	t	tTA	L

0.0343	1173332	CDS	Cj1245c		putative membrane protein	1172840	1174036	-1	T	AAT	N	synonymous	c	AAc	N
0.0343	1173349	CDS	Cj1245c		putative membrane protein	1172840	1174036	-1	A	AAA	K	nonsynonymous	g	gAA	E
0.0344	571050	CDS	Cj0610c		putative periplasmic protein	570111	571121	-1	T	AGT	S	synonymous	c	AGc	S
0.0344	571050	sig_peptide	Cj0610c			571038	571121		T				c		
0.0344	592956	CDS	Cj0631c		putative ribonuclease	592495	594429	-1	G	GTC	V	nonsynonymous	a	aTC	I
0.0345	2370	CDS	Cj0002	dnaN	DNA polymerase III, beta chain	1483	2550	1	A	GAA	E	synonymous	g	GAg	E
0.0345	912941	CDS	Cj0979c		putative secreted nuclease	912435	912962	-1	A	ATA	I	nonsynonymous	t	tTA	L
0.0345	912941	sig_peptide	Cj0979c			912843	912962		A				t		
0.0345	1279851	CDS	Cj1346c		1-deoxy-D-xylulose 5-phosphate reductoisomerase	1278851	1279921	-1	C	CCC	P	nonsynonymous	g	Cgt	R
0.0345	1279904	CDS	Cj1346c		1-deoxy-D-xylulose 5-phosphate reductoisomerase	1278851	1279921	-1	T	AGT	S	synonymous	c	AGc	S
0.0346	1003468	CDS	Cj1068		putative peptidase M50 family protein	1003100	1004206	1	A	CTA	L	synonymous	g	CTg	L
0.0346	1032662	CDS	Cj1098	pyrB	aspartate carbamoyltransferase	1032186	1033073	1	T	AAT	N	synonymous	c	AAc	N
0.0346	1032665	CDS	Cj1098	pyrB	aspartate carbamoyltransferase	1032186	1033073	1	C	TCC	S	synonymous	t	TCt	S
0.0346	1032668	CDS	Cj1098	pyrB	aspartate carbamoyltransferase	1032186	1033073	1	C	CGC	R	synonymous	t	CGt	R
0.0346	1032671	CDS	Cj1098	pyrB	aspartate carbamoyltransferase	1032186	1033073	1	T	GTT	V	synonymous	a	GTa	V
0.0346	1308351	CDS	Cj1371		putative periplasmic protein (VacJ homolog)	1308007	1308705	1	C	ATC	I	synonymous	t	ATt	I
0.0346	1308357	CDS	Cj1371		putative periplasmic protein (VacJ homolog)	1308007	1308705	1	T	GGT	G	synonymous	c	GGc	G
0.0347	618707	CDS	Cj0661c	era	GTP-binding protein Era homolog	618206	619081	-1	A	ACA	T	synonymous	g	ACg	T
0.0347	1175253	CDS	Cj1246c	uvrC	excinuclease ABC subunit C	1174038	1175840	-1	A	ATA	I	nonsynonymous	g	ATg	M
0.0348	1453358	CDS	Cj1517	moaD	putative molybdopterin converting factor,subunit 1	1453143	1453364	1	A	GGA	G	synonymous	g	GGg	G
0.0349	1175898	CDS	Cj1247c		hypothetical protein Cj1247c	1175833	1176327	-1	A	AAA	K	nonsynonymous	g	gAA	E
0.035	56992	CDS	Cj0038c		putative poly(A) polymerase family protein	56564	57211	-1	A	AAT	N	nonsynonymous	g	gAT	D

0.035	57667	CDS	Cj0039c	typA	GTP-binding protein TypA homolog	57211	59019	-1	T	AGT	S	synonymous	c	AGc	S
0.0352	98590	CDS	Cj0088	dcuA	anaerobic C4-dicarboxylate transporter	97496	98833	1	T	ACT	T	synonymous	c	ACc	T
0.0352	1224685	CDS	Cj1292	dcd	putative deoxycytidine triphosphate deaminase	1224239	1224799	1	T	GGT	G	synonymous	c	GGc	G
0.0352	1444086	CDS	Cj1508c	fdhD	FdhD protein	1443825	1444607	-1	G	CAG	Q	synonymous	a	CAa	Q
0.0355	128455	CDS	Cj0127c	accD	acetyl-coenzyme A carboxylase carboxyl transferase subunit beta	128227	129069	-1	A	GGA	G	synonymous	g	GGg	G
0.0357	568694	CDS	Cj0608		putative outer membrane efflux protein	567573	568943	1	C	AGC	S	synonymous	t	AGt	S
0.0357	568695	CDS	Cj0608		putative outer membrane efflux protein	567573	568943	1	G	GAT	D	nonsynonymous	a	aAT	N
0.0357	841059	CDS	Cj0901		putative amino acid ABC tansporter permease protein	840763	841506	1	G	CCG	P	synonymous	a	CCa	P
0.0357	841200	CDS	Cj0901		putative amino acid ABC tansporter permease protein	840763	841506	1	A	AAA	K	synonymous	g	AAg	K
0.0357	841269	CDS	Cj0901		putative amino acid ABC tansporter permease protein	840763	841506	1	T	CCT	P	synonymous	a	CCa	P
0.0359	286910	CDS	Cj0316	pheA	chorismate mutase/prephenate dehydratase	286248	287321	1	C	TAC	Y	synonymous	t	TAt	Y
0.036	775998	CDS	Cj0828c	ilvA	threonine dehydratase biosynthetic	775149	776360	-1	A	GGA	G	synonymous	g	GGg	G
0.036	1333206	CDS	Cj1399c	hydA2	putative Ni/Fe-hydrogenase small subunit	1333149	1334642	-1	C	AGC	S	synonymous	t	AGt	S
0.0361	47848	CDS	Cj0031		putative type IIS restriction/modification enzyme	46424	50156	1	T	TGT	C	synonymous	c	TGc	C
0.0362	246547	intergenic							T				c		
0.0363	25658	CDS	Cj0020c		cytochrome C551 peroxidase	25433	26347	-1	C	CGC	R	synonymous	t	CGt	R
0.0364	22223	CDS	Cj0017c	dsbI	disulphide bond formation protein	21854	23380	-1	T	ACT	T	MNP	c		
0.0364	22224	CDS	Cj0017c	dsbI	disulphide bond formation protein	21854	23380	-1	C	ACT	T	nonsynonymous	g	Agc	S
0.0364	1164440	CDS	Cj1235		putative peptidase M23 family protein	1163799	1164620	1	T	TAT	Y	synonymous	c	TAc	Y
0.0366	18378	CDS	Cj0013	ilvD	dihydroxy-acid dehydratase	17563	19239	1	A	ACA	T	synonymous	t	ACt	T
0.0367	120622	CDS	Cj0115	slyD	FKBP-type peptidyl-prolyl cis-trans isomerase	120194	120763	1	G	GCG	A	synonymous	a	GCa	A

0.0368	47525	CDS	Cj0031		putative type IIS restriction/modification enzyme	46424	50156	1	T	TTA	L	synonymous	c	cTA	L
0.0368	1502117	CDS	Cj1571c	nuoI	NADH dehydrogenase I chain I	1501889	1502530	-1	G	GTG	V	synonymous	a	GTa	V
0.0369	842876	CDS	Cj0903c		putative amino-acid transport protein	842255	843721	-1	G	TTG	L	synonymous	a	TTa	L
0.0369	1188587	CDS	Cj1258		putative phosphotyrosine protein phosphatase	1188414	1188869	1	G	AAG	K	synonymous	a	AAa	K
0.037	1281930	CDS	Cj1349c		putative fibronectin/fibrinogen-binding protein	1280992	1282299	-1	G	GAT	D	nonsynonymous	a	aAT	N
0.0372	9833	CDS	Cj0007	gltB	glutamate synthase (NADPH) large subunit	8144	12634	1	T	TTG	L	synonymous	c	cTG	L
0.0372	20236	CDS	Cj0015c		hypothetical protein Cj0015c	19867	21093	-1	A	GTA	V	synonymous	g	GTg	V
0.0372	20602	CDS	Cj0015c		hypothetical protein Cj0015c	19867	21093	-1	A	GAA	E	synonymous	g	GAg	E
0.0372	20638	CDS	Cj0015c		hypothetical protein Cj0015c	19867	21093	-1	T	AAT	N	synonymous	c	AAc	N
0.0372	20716	CDS	Cj0015c		hypothetical protein Cj0015c	19867	21093	-1	A	AGA	R	synonymous	g	AGg	R
0.0372	21131	intergenic							T				c		
0.0372	21637	CDS	Cj0016		putative transcriptional regulatory protein	21159	21833	1	G	TGT	C	nonsynonymous	a	TaT	Y
0.0372	21799	CDS	Cj0016		putative transcriptional regulatory protein	21159	21833	1	C	GCA	A	nonsynonymous	t	GtA	V
0.0372	21983	CDS	Cj0017c	dsbI	disulphide bond formation protein	21854	23380	-1	A	AAA	K	synonymous	g	AAg	K
0.0372	22022	CDS	Cj0017c	dsbI	disulphide bond formation protein	21854	23380	-1	C	AGC	S	synonymous	t	AGt	S
0.0372	22160	CDS	Cj0017c	dsbI	disulphide bond formation protein	21854	23380	-1	A	GTA	V	synonymous	g	GTg	V
0.0372	22400	CDS	Cj0017c	dsbI	disulphide bond formation protein	21854	23380	-1	C	GGC	G	synonymous	t	GGt	G
0.0372	22463	CDS	Cj0017c	dsbI	disulphide bond formation protein	21854	23380	-1	A	TCA	S	synonymous	g	TCg	S
0.0372	22481	CDS	Cj0017c	dsbI	disulphide bond formation protein	21854	23380	-1	A	GAA	E	synonymous	g	GAg	E
0.0372	22511	CDS	Cj0017c	dsbI	disulphide bond formation protein	21854	23380	-1	T	GAT	D	synonymous	c	GAc	D
0.0372	23003	CDS	Cj0017c	dsbI	disulphide bond formation protein	21854	23380	-1	A	TTA	L	synonymous	g	TTg	L

0.0372	25001	CDS	Cj0019c		putative MCP-domain signal transduction protein	23665	25443	-1	A	AAA	K	nonsynonymous	g	AgA	R
0.0373	1092188	CDS	Cj1161c		putative cation-transporting ATPase	1091795	1093894	-1	C	CAC	H	MNP	t		
0.0373	1092190	CDS	Cj1161c		putative cation-transporting ATPase	1091795	1093894	-1	C	CAC	H	nonsynonymous	t	tAt	Y
0.0373	1279823	CDS	Cj1346c		1-deoxy-D-xylulose 5-phosphate reductoisomerase	1278851	1279921	-1	C	AAC	N	synonymous	t	AAt	N
0.0374	381642	CDS	Cj0414		putative oxidoreductase subunit	380937	381665	1	G	GGT	G	nonsynonymous	a	aGT	S
0.0377	1221257	CDS	Cj1288c	gltX2	glutamyl-tRNA synthetase	1219994	1221385	-1	T	ATT	I	synonymous	c	ATc	I
0.0377	1227463	CDS	Cj1295		conserved hypothetical protein Cj1295	1226978	1228285	1	A	GCA	A	synonymous	g	GCg	A
0.0378	1189459	CDS	Cj1259	porA	major outer membrane protein	1189121	1190395	1	C	TAC	Y	synonymous	t	TAt	Y
0.0381	1167115	CDS	Cj1238	pdxJ	pyridoxal phosphate biosynthetic protein	1166594	1167367	1	T	CTT	L	synonymous	c	CTc	L
0.0381	1167116	CDS	Cj1238	pdxJ	pyridoxal phosphate biosynthetic protein	1166594	1167367	1	G	GAT	D	nonsynonymous	a	aAT	N
0.0381	1197739	CDS	Cj1267c	hydA	Ni/Fe-hydrogenase small chain	1197214	1198353	-1	C	AGC	S	synonymous	t	AGt	S
0.0381	1197910	CDS	Cj1267c	hydA	Ni/Fe-hydrogenase small chain	1197214	1198353	-1	A	ACA	T	synonymous	g	ACg	T
0.0385	22229	CDS	Cj0017c	dsbI	disulphide bond formation protein	21854	23380	-1	T	AGT	S	synonymous	c	AGc	S
0.0385	22259	CDS	Cj0017c	dsbI	disulphide bond formation protein	21854	23380	-1	C	ACC	T	synonymous	t	ACt	T
0.0385	157401	CDS	Cj0154c		putative tetrapyrrole methylase family protein	157312	158136	-1	T	TCT	S	nonsynonymous	c	cCT	P
0.0385	287460	CDS	Cj0317	hisC	histidinol-phosphate aminotransferase	287311	288405	1	T	TGT	C	synonymous	c	TGc	C
0.0386	1171506	CDS	Cj1243		uroporphyrinogen decarboxylase	1170922	1171944	1	T	GCT	A	synonymous	c	GCc	A
0.0388	20032	CDS	Cj0015c		hypothetical protein Cj0015c	19867	21093	-1	T	ATT	I	synonymous	c	ATc	I
0.0388	20095	CDS	Cj0015c		hypothetical protein Cj0015c	19867	21093	-1	C	AGC	S	synonymous	t	AGt	S
0.039	1281727	CDS	Cj1349c		putative fibronectin/fibrinogen-binding protein	1280992	1282299	-1	G	GTG	V	synonymous	a	GTa	V
0.039	1641030	CDS	Cj1731c	ruvC	crossover junction endodeoxyribonuclease	1640904	1641386	-1	T	ACT	T	synonymous	c	ACc	T
0.0391	1016295	CDS	Cj1085c	mfd	transcription-repair coupling factor	1015629	1018565	-1	C	GGC	G	synonymous	t	GGt	G

0.0393	1560316	CDS	Cj1634c	aroC	chorismate synthase	1559723	1560811	-1	G	GA	E	nonsynonymous	a	aAA	K
0.0393	1560728	CDS	Cj1634c	aroC	chorismate synthase	1559723	1560811	-1	G	ATG	M	nonsynonymous	a	ATa	I
0.0395	120565	CDS	Cj0115	slyD	FKBP-type peptidyl-prolyl cis-trans isomerase	120194	120763	1	T	CAT	H	synonymous	c	CAc	H
0.0395	919324	CDS	Cj0985c	hipO	hippurate hydrolase	918580	919731	-1	T	GGT	G	synonymous	c	GGc	G
0.0395	919360	CDS	Cj0985c	hipO	hippurate hydrolase	918580	919731	-1	C	GGC	G	synonymous	t	GGt	G
0.0396	100069	CDS	Cj1062		putative CinA-like protein	999629	1000717	1	A	GGA	G	synonymous	t	GGt	G
0.0397	18963	CDS	Cj0013	ilvD	dihydroxy-acid dehydratase	17563	19239	1	C	ACC	T	synonymous	t	ACt	T
0.0397	18978	CDS	Cj0013	ilvD	dihydroxy-acid dehydratase	17563	19239	1	T	AGT	S	synonymous	c	AGc	S
0.0397	19246	intergenic							T				a		
0.0397	19269	CDS	Cj0014c		putative integral membrane protein	19251	19775	-1	C	ACC	T	synonymous	t	ACt	T
0.0397	19927	CDS	Cj0015c		hypothetical protein Cj0015c	19867	21093	-1	C	ACC	T	synonymous	t	ACt	T
0.0397	19953	CDS	Cj0015c		hypothetical protein Cj0015c	19867	21093	-1	C	CAT	H	nonsynonymous	t	tAT	Y
0.0397	20029	CDS	Cj0015c		hypothetical protein Cj0015c	19867	21093	-1	T	GCT	A	synonymous	c	GCc	A
0.0397	20110	CDS	Cj0015c		hypothetical protein Cj0015c	19867	21093	-1	C	TGC	C	synonymous	t	TGt	C
0.0397	20596	CDS	Cj0015c		hypothetical protein Cj0015c	19867	21093	-1	A	GAA	E	synonymous	g	GAg	E
0.0397	20766	CDS	Cj0015c		hypothetical protein Cj0015c	19867	21093	-1	C	CTA	L	synonymous	t	tTA	L
0.0397	20770	CDS	Cj0015c		hypothetical protein Cj0015c	19867	21093	-1	A	GAA	E	synonymous	g	GAg	E
0.0397	20773	CDS	Cj0015c		hypothetical protein Cj0015c	19867	21093	-1	G	GGG	G	synonymous	a	GGa	G
0.0397	20776	CDS	Cj0015c		hypothetical protein Cj0015c	19867	21093	-1	T	GAT	D	nonsynonymous	a	GAa	E
0.0397	20941	CDS	Cj0015c		hypothetical protein Cj0015c	19867	21093	-1	G	GTG	V	synonymous	a	GTa	V
0.0397	21221	CDS	Cj0016		putative transcriptional regulatory protein	21159	21833	1	G	CTG	L	synonymous	a	CTa	L
0.0397	21284	CDS	Cj0016		putative transcriptional regulatory protein	21159	21833	1	A	GAA	E	synonymous	g	GAg	E

0.0397	21407	CDS	Cj0016		putative transcriptional regulatory protein	21159	21833	1	C	ATC	I	synonymous	t	ATt	I
0.0397	21614	CDS	Cj0016		putative transcriptional regulatory protein	21159	21833	1	T	GGT	G	synonymous	c	GGc	G
0.0397	21752	CDS	Cj0016		putative transcriptional regulatory protein	21159	21833	1	T	TGT	C	synonymous	c	TGc	C
0.0397	21803	CDS	Cj0016		putative transcriptional regulatory protein	21159	21833	1	A	GGA	G	synonymous	g	GGg	G
0.0397	21902	CDS	Cj0017c	dsbI	disulphide bond formation protein	21854	23380	-1	T	AAT	N	synonymous	c	AAc	N
0.0397	21986	CDS	Cj0017c	dsbI	disulphide bond formation protein	21854	23380	-1	C	GTC	V	synonymous	t	GTt	V
0.0397	22285	CDS	Cj0017c	dsbI	disulphide bond formation protein	21854	23380	-1	A	AGT	S	nonsynonymous	g	gGT	G
0.0397	22295	CDS	Cj0017c	dsbI	disulphide bond formation protein	21854	23380	-1	T	GGT	G	synonymous	c	GGc	G
0.0397	22325	CDS	Cj0017c	dsbI	disulphide bond formation protein	21854	23380	-1	C	GAC	D	synonymous	t	GAt	D
0.0397	22523	CDS	Cj0017c	dsbI	disulphide bond formation protein	21854	23380	-1	A	GTA	V	synonymous	g	GTg	V
0.0397	22553	CDS	Cj0017c	dsbI	disulphide bond formation protein	21854	23380	-1	C	AAC	N	synonymous	t	AAt	N
0.0397	22556	CDS	Cj0017c	dsbI	disulphide bond formation protein	21854	23380	-1	T	AAT	N	synonymous	c	AAc	N
0.0397	22805	CDS	Cj0017c	dsbI	disulphide bond formation protein	21854	23380	-1	T	AGT	S	synonymous	c	AGc	S
0.0397	22880	CDS	Cj0017c	dsbI	disulphide bond formation protein	21854	23380	-1	A	GCA	A	synonymous	t	GCt	A
0.0397	23318	CDS	Cj0017c	dsbI	disulphide bond formation protein	21854	23380	-1	T	ATT	I	synonymous	c	ATc	I
0.0397	23318	sig_peptide	Cj0017c	dsbI		23288	23380		T				c		
0.0397	23728	CDS	Cj0019c		putative MCP-domain signal transduction protein	23665	25443	-1	T	AAT	N	synonymous	c	AAc	N
0.0397	23740	CDS	Cj0019c		putative MCP-domain signal transduction protein	23665	25443	-1	T	TGT	C	synonymous	c	TGc	C
0.0397	23773	CDS	Cj0019c		putative MCP-domain signal transduction protein	23665	25443	-1	C	AAC	N	synonymous	t	AAt	N
0.0397	23779	CDS	Cj0019c		putative MCP-domain signal transduction protein	23665	25443	-1	A	AAA	K	synonymous	g	AAg	K
0.0397	23788	CDS	Cj0019c		putative MCP-domain signal transduction protein	23665	25443	-1	T	AAT	N	synonymous	c	AAc	N
0.0397	23869	CDS	Cj0019c		putative MCP-domain signal transduction protein	23665	25443	-1	G	TTG	L	synonymous	a	TTa	L

0.0397	23956	CDS	Cj0019c		putative MCP-domain signal transduction protein	23665	25443	-1	A	TTA	L	synonymous	g	TTg	L
0.0397	24091	CDS	Cj0019c		putative MCP-domain signal transduction protein	23665	25443	-1	A	GCA	A	synonymous	g	GCg	A
0.0397	24129	CDS	Cj0019c		putative MCP-domain signal transduction protein	23665	25443	-1	T	TTA	L	synonymous	c	cTA	L
0.0397	24133	CDS	Cj0019c		putative MCP-domain signal transduction protein	23665	25443	-1	T	ACT	T	synonymous	c	ACc	T
0.0397	24142	CDS	Cj0019c		putative MCP-domain signal transduction protein	23665	25443	-1	T	GCT	A	synonymous	c	GCc	A
0.0397	24148	CDS	Cj0019c		putative MCP-domain signal transduction protein	23665	25443	-1	C	GAC	D	synonymous	t	GAt	D
0.0397	24152	CDS	Cj0019c		putative MCP-domain signal transduction protein	23665	25443	-1	A	AAA	K	nonsynonymous	g	AgA	R
0.0397	24157	CDS	Cj0019c		putative MCP-domain signal transduction protein	23665	25443	-1	C	TCC	S	synonymous	t	TCt	S
0.0397	24160	CDS	Cj0019c		putative MCP-domain signal transduction protein	23665	25443	-1	C	AAC	N	MNP	t		
0.0397	24161	CDS	Cj0019c		putative MCP-domain signal transduction protein	23665	25443	-1	A	AAC	N	nonsynonymous	g	Agt	S
0.0397	24166	CDS	Cj0019c		putative MCP-domain signal transduction protein	23665	25443	-1	C	ATC	I	synonymous	t	ATt	I
0.0397	24184	CDS	Cj0019c		putative MCP-domain signal transduction protein	23665	25443	-1	G	GGA	A	synonymous	a	GCa	A
0.0397	24283	CDS	Cj0019c		putative MCP-domain signal transduction protein	23665	25443	-1	A	AAA	K	synonymous	g	AAg	K
0.0397	24292	CDS	Cj0019c		putative MCP-domain signal transduction protein	23665	25443	-1	A	CTA	L	synonymous	g	CTg	L
0.0397	24373	CDS	Cj0019c		putative MCP-domain signal transduction protein	23665	25443	-1	T	ACT	T	synonymous	c	ACc	T
0.0397	24742	CDS	Cj0019c		putative MCP-domain signal transduction protein	23665	25443	-1	T	TCT	S	synonymous	c	TCc	S
0.0397	24756	CDS	Cj0019c		putative MCP-domain signal transduction protein	23665	25443	-1	A	ACT	T	nonsynonymous	g	gCT	A
0.0397	25036	CDS	Cj0019c		putative MCP-domain signal transduction protein	23665	25443	-1	G	TTG	L	synonymous	a	TTa	L
0.0397	25062	CDS	Cj0019c		putative MCP-domain signal transduction protein	23665	25443	-1	C	CTA	L	synonymous	t	tTA	L
0.0397	25069	CDS	Cj0019c		putative MCP-domain signal transduction protein	23665	25443	-1	A	TTA	L	synonymous	g	TTg	L
0.0397	25111	CDS	Cj0019c		putative MCP-domain signal transduction protein	23665	25443	-1	T	TTT	F	synonymous	c	TTc	F
0.0397	25307	CDS	Cj0019c		putative MCP-domain signal transduction protein	23665	25443	-1	T	ATA	I	nonsynonymous	c	AcA	T

0.0397	840293	CDS	Cj0899c	thiJ	4-methyl-5(beta-hydroxyethyl)-thiazole monophosphate synthesis protein	839939	840508	-1	T	GCT	A	synonymous	c	GCc	A
0.0398	618742	CDS	Cj0661c	era	GTP-binding protein Era homolog	618206	619081	-1	A	ACT	T	nonsynonymous	g	gCT	A
0.0398	619189	CDS	Cj0662c	hslU	ATP-dependent Hsl protease ATP-binding subunit	619078	620397	-1	T	AGT	S	synonymous	c	AGc	S
0.0402	1039743	CDS	Cj1106		putative periplasmic thioredoxin	1039561	1040163	1	G	GTG	V	synonymous	a	GTa	V
0.0402	1040441	CDS	Cj1107	clpS	ATP-dependent Clp protease adaptor protein	1040163	1040453	1	G	GTG	V	synonymous	a	GTa	V
0.0402	1227277	CDS	Cj1295		conserved hypothetical protein Cj1295	1226978	1228285	1	C	AAC	N	synonymous	t	AAt	N
0.0405	842944	CDS	Cj0903c		putative amino-acid transport protein	842255	843721	-1	A	ATT	I	nonsynonymous	g	gTT	V
0.0405	842987	CDS	Cj0903c		putative amino-acid transport protein	842255	843721	-1	C	GCC	A	synonymous	t	GCt	A
0.0405	999267	CDS	Cj1061c	ileS	isoleucyl-tRNA synthetase	996777	999530	-1	G	GAG	E	synonymous	a	GAa	E
0.0405	1001905	CDS	Cj1066	rdxA	nitroreductase	1001837	1002442	1	A	AAA	K	nonsynonymous	t	AAt	N
0.0406	1343648	CDS	Cj1411c		putative cytochrome P450	1342550	1343911	-1	T	AGT	S	synonymous	c	AGc	S
0.0407	54719	CDS	Cj0036		hypothetical protein Cj0036	53970	55319	1	A	AAA	K	synonymous	g	AAg	K
0.0408	571830	CDS	Cj0611c		putative acyltransferase family protein	571122	572498	-1	C	AGC	S	synonymous	t	AGt	S
0.0408	571986	CDS	Cj0611c		putative acyltransferase family protein	571122	572498	-1	G	CAG	Q	synonymous	a	CAa	Q
0.0408	572604	CDS	Cj0612c	cft	ferritin	572541	573044	-1	C	CAC	H	synonymous	t	CAt	H
0.0409	246538	CDS	Cj0267c		putative integral membrane protein	246013	246543	-1	G	GAG	E	synonymous	a	GAa	E
0.041	586957	CDS	Cj0626	hypE	hydrogenase isoenzymes formation protein	586190	587164	1	G	GAG	E	synonymous	a	GAa	E
0.0411	881036	CDS	Cj0942c	secA	preprotein translocase SecA subunit	880838	883426	-1	G	GCG	A	synonymous	a	GCa	A
0.0411	1189582	CDS	Cj1259	porA	major outer membrane protein	1189121	1190395	1	C	AAC	N	synonymous	t	AAt	N
0.0414	1306665	CDS	Cj1369		putative permease	1306186	1307505	1	C	GTC	V	synonymous	a	GTa	V
0.0414	1306869	CDS	Cj1369		putative permease	1306186	1307505	1	T	GGT	G	synonymous	c	GGc	G

0.0414	1306870	CDS	Cj1369		putative permease	1306186	1307505	1	T	TTA	L	synonymous	c	cTA	L
0.0414	1307073	CDS	Cj1369		putative permease	1306186	1307505	1	A	TCA	S	synonymous	g	TCg	S
0.0414	1307076	CDS	Cj1369		putative permease	1306186	1307505	1	C	GCC	A	synonymous	t	GCt	A
0.0415	616137	pseudogene	Cj0654c		pseudogene (putative transmembrane transport protein)	615014	616485	-1	A	AAC	N	nonsynonymous	g	AgC	S
0.0418	343531	CDS	Cj0376		putative periplasmic protein	343102	343926	1	G	GGT	G	nonsynonymous	a	aGT	S
0.0418	343632	CDS	Cj0376		putative periplasmic protein	343102	343926	1	T	AGT	S	synonymous	c	AGc	S
0.0418	840352	CDS	Cj0899c	thiJ	4-methyl-5(beta-hydroxyethyl)-thiazole monophosphate synthesis protein	839939	840508	-1	A	AGC	S	nonsynonymous	g	gGC	G
0.0419	378708	CDS	Cj0412		putative ATP/GTP binding protein	378268	380097	1	C	CCC	P	synonymous	t	CCt	P
0.0421	121590	CDS	Cj0116	fabD	malonyl CoA-acyl carrier protein transacylase	120763	121683	1	T	AGT	S	synonymous	c	AGc	S
0.0421	121754	CDS	Cj0117	pfs	5'-methylthioadenosine/S-adenosylhomocysteine nucleosidase	121680	122369	1	G	ACG	T	synonymous	a	ACa	T
0.0422	119604 1	CDS	Cj1266c	hydB	Ni/Fe-hydrogenase large subunit	119549 6	119721 1	-1	G	GTC	V	nonsynonymous	a	aTC	I
0.0424	108501 9	CDS	Cj1152c	gmhB	D,D-heptose 1,7-bisphosphate phosphatase	108473 7	108529 7	-1	A	AAA	K	synonymous	g	AAg	K
0.0428	54581	CDS	Cj0036		hypothetical protein Cj0036	53970	55319	1	T	AGT	S	synonymous	c	AGc	S
0.0431	915184	CDS	Cj0981c	cjaB	putative MFS (Major Facilitator Superfamily) transport protein	914281	915528	-1	T	CGT	R	synonymous	c	CGc	R
0.0431	915187	CDS	Cj0981c	cjaB	putative MFS (Major Facilitator Superfamily) transport protein	914281	915528	-1	T	ATT	I	synonymous	a	ATa	I
0.0433	591777	CDS	Cj0630c	holA	putative DNA polymerase III, delta subunit	591537	592502	-1	T	ATT	I	synonymous	c	ATc	I
0.0433	591801	CDS	Cj0630c	holA	putative DNA polymerase III, delta subunit	591537	592502	-1	C	AAC	N	synonymous	t	AAt	N
0.0433	592002	CDS	Cj0630c	holA	putative DNA polymerase III, delta subunit	591537	592502	-1	A	TTA	L	synonymous	g	TTg	L
0.0433	592061	CDS	Cj0630c	holA	putative DNA polymerase III, delta subunit	591537	592502	-1	G	GGG	G	nonsynonymous	a	aGG	R
0.0433	592252	CDS	Cj0630c	holA	putative DNA polymerase III, delta subunit	591537	592502	-1	C	ACA	T	nonsynonymous	t	AtA	I

0.0433	592333	CDS	Cj0630c	holA	putative DNA polymerase III, delta subunit	591537	592502	-1	G	CGT	R	nonsynonymous	a	CaT	H
0.0433	592485	CDS	Cj0630c	holA	putative DNA polymerase III, delta subunit	591537	592502	-1	T	CTT	L	synonymous	c	CTc	L
0.0433	592601	CDS	Cj0631c		putative ribonuclease	592495	594429	-1	T	GTA	V	nonsynonymous	c	GcA	A
0.0433	592604	CDS	Cj0631c		putative ribonuclease	592495	594429	-1	G	CGT	R	nonsynonymous	a	CaT	H
0.0433	592669	CDS	Cj0631c		putative ribonuclease	592495	594429	-1	T	GAT	D	synonymous	c	GAc	D
0.0433	1560347	CDS	Cj1634c	aroC	chorismate synthase	1559723	1560811	-1	A	GGA	G	synonymous	g	GGg	G
0.0434	1170933	CDS	Cj1243		uroporphyrinogen decarboxylase	1170922	1171944	1	C	ATC	I	synonymous	t	ATt	I
0.0435	1092957	CDS	Cj1161c		putative cation-transporting ATPase	1091795	1093894	-1	G	AGT	S	nonsynonymous	a	AaT	N
0.0436	1390738	CDS	Cj0136	infB	translation initiation factor IF-2	1368536	1394731	1	G	AAG	K	synonymous	a	AAa	K
0.0436	644156	CDS	Cj0689	ackA	acetate kinase	643409	644599	1	A	ATT	I	nonsynonymous	g	gTT	V
0.0437	596308	CDS	Cj0633		putative periplasmic protein	595574	596656	1	G	ACG	T	synonymous	a	ACa	T
0.0439	574232	CDS	Cj0614	pstC	putative phosphate transport system permease protein	574229	575143	1	T	TTA	L	synonymous	c	cTA	L
0.0439	574232	sig_peptide	Cj0614	pstC		574229	574342		T				c		
0.0439	580908	CDS	Cj0621		hypothetical protein Cj0621	580326	581846	1	C	CTT	L	nonsynonymous	t	tTT	F
0.0439	580940	CDS	Cj0621		hypothetical protein Cj0621	580326	581846	1	G	AAG	K	synonymous	a	AAa	K
0.0439	593821	CDS	Cj0631c		putative ribonuclease	592495	594429	-1	C	AAC	N	synonymous	t	AAt	N
0.0439	594019	CDS	Cj0631c		putative ribonuclease	592495	594429	-1	C	AGC	S	synonymous	t	AGt	S
0.0439	594319	CDS	Cj0631c		putative ribonuclease	592495	594429	-1	C	CAC	H	synonymous	t	CAt	H
0.0439	594490	intergenic							A				t		
0.0439	594824	CDS	Cj0632	ilvC	ketol-acid reductoisomerase	594548	595570	1	G	GTC	V	nonsynonymous	a	aTC	I
0.044	1002034	CDS	Cj1066	rdxA	nitroreductase	1001837	1002442	1	A	CAA	Q	synonymous	g	CAg	Q
0.044	1093196	CDS	Cj1161c		putative cation-transporting ATPase	1091795	1093894	-1	C	GGC	G	synonymous	t	GGt	G

0.044	1093313	CDS	Cj1161c		putative cation-transporting ATPase	1091795	1093894	-1	T	GCT	A	synonymous	c	GCc	A
0.044	1093445	CDS	Cj1161c		putative cation-transporting ATPase	1091795	1093894	-1	G	TTG	L	synonymous	a	TTa	L
0.044	1093490	CDS	Cj1161c		putative cation-transporting ATPase	1091795	1093894	-1	A	TTA	L	synonymous	g	TTg	L
0.044	1093830	CDS	Cj1161c		putative cation-transporting ATPase	1091795	1093894	-1	G	AGA	R	nonsynonymous	a	AaA	K
0.044	1227458	CDS	Cj1295		conserved hypothetical protein Cj1295	1226978	1228285	1	G	GAT	D	nonsynonymous	a	aAT	N
0.0441	5664901	CDS	Cj0607		ABC-type transmembrane transport protein	5656456	5675701	1	A	GTA	V	synonymous	g	GTg	V
0.0441	5665300	CDS	Cj0607		ABC-type transmembrane transport protein	5656456	5675701	1	T	TCT	S	synonymous	c	TCc	S
0.0442	3516776	CDS	Cj0386		putative GTP-binding protein	3514456	3528208	1	T	GCT	A	synonymous	c	GCc	A
0.0442	8808006	CDS	Cj0941c		putative permease	8796376	8808401	-1	G	AAG	K	synonymous	a	AAa	K
0.0442	8808006	sig_peptide	Cj0941c			8807004	8808401		G				a		
0.0443	1220936	CDS	Cj1288c	gltX2	glutamyl-tRNA synthetase	1219994	1221385	-1	C	GGC	G	synonymous	t	GGt	G
0.0443	1220969	CDS	Cj1288c	gltX2	glutamyl-tRNA synthetase	1219994	1221385	-1	A	GAA	E	synonymous	g	GAg	E
0.0446	7415305	CDS	Cj0791c		putative aminotransferase	7411667	7424401	-1	G	GTT	V	nonsynonymous	a	aTT	I
0.0447	1221578	CDS	Cj1289		possible periplasmic protein	1221450	1222265	1	A	GCA	A	synonymous	g	GCg	A
0.0448	1221111	CDS	Cj0117	pfs	5'-methylthioadenosine/S-adenosylhomocysteine nucleosidase	1216810	1223609	1	C	GGC	G	synonymous	t	GGt	G
0.0448	8411779	CDS	Cj0901		putative amino acid ABC tansporter permease protein	8407663	8415006	1	T	AGT	S	synonymous	c	AGc	S
0.0448	8414400	CDS	Cj0901		putative amino acid ABC tansporter permease protein	8407663	8415006	1	T	TAT	Y	synonymous	c	TAc	Y
0.0448	1400772	CDS	Cj1466	flgK	putative flagellar hook-associated protein	1400602	1402428	1	C	GGC	G	synonymous	t	GGt	G
0.0449	28972	CDS	Cj0023	purB	adenylosuccinate lyase	28382	29710	1	A	GAA	E	synonymous	g	GAg	E
0.0449	2237004	CDS	Cj0240c	iscS	cysteine desulfurase (NifS protein homolog)	2228897	2240778	-1	C	AAC	N	synonymous	t	AAt	N
0.045	1039544	CDS	Cj1105	smpB	small protein B homolog	1039098	1039550	1	C	TAC	Y	synonymous	t	TAt	Y
0.045	1039557	intergenic							A				t		

0.045	1039576	CDS	Cj1106		putative periplasmic thioredoxin	1039561	1040163	1	G	GC	A	nonsynonymous	a	aCC	T
0.045	1039576	sig_peptide	Cj1106			1039561	1039620		G				a		
0.045	1039623	CDS	Cj1106		putative periplasmic thioredoxin	1039561	1040163	1	G	AAG	K	synonymous	a	AAa	K
0.0451	249413	CDS	Cj0272		conserved hypothetical protein Cj0272	249405	250496	1	A	GTA	V	synonymous	g	GTg	V
0.0451	249428	CDS	Cj0272		conserved hypothetical protein Cj0272	249405	250496	1	C	AGC	S	synonymous	t	AGt	S
0.0451	1183245	CDS	Cj1252		putative periplasmic protein	1181476	1183524	1	T	GGT	G	synonymous	c	GGc	G
0.0452	833667	CDS	Cj0893c	rpsA	30S ribosomal protein S1	832203	833873	-1	C	GAC	D	synonymous	t	GAt	D
0.0452	834014	CDS	Cj0894c	ispH	4-hydroxy-3-methylbut-2-enyl diphosphate reductase	834005	834838	-1	A	AGA	R	synonymous	g	AGg	R
0.0452	834377	CDS	Cj0894c	ispH	4-hydroxy-3-methylbut-2-enyl diphosphate reductase	834005	834838	-1	A	GTA	V	synonymous	g	GTg	V
0.0453	914094	CDS	Cj0980		putative peptidase	913030	914298	1	A	GCA	A	synonymous	g	GCg	A
0.0454	1518636	CDS	Cj1587c		putative ABC transporter	1518009	1519640	-1	T	CAT	H	synonymous	c	CAc	H
0.0455	1279926	CDS	Cj1347c	cdsA	phosphatidate cytidylyltransferase	1279918	1280643	-1	C	CTT	L	nonsynonymous	t	tTT	F
0.0457	563993	CDS	Cj0605		putative amidohydrolase	563205	564395	1	T	ACT	T	synonymous	c	ACc	T
0.0457	587241	CDS	Cj0627	hypA	hydrogenase expression/formation protein	587164	587508	1	G	GTG	V	synonymous	a	GTa	V
0.0457	1094273	CDS	Cj1163c		putative cation transport protein	1094078	1095028	-1	T	AGT	S	synonymous	c	AGc	S
0.0457	1383966	CDS	Cj1444c	kpsD	capsule polysaccharide export system periplasmic protein	1383486	1385144	-1	T	AGT	S	synonymous	c	AGc	S
0.0457	1398238	CDS	Cj1461		putative DNA methylase	1397726	1398409	1	T	AGT	S	synonymous	c	AGc	S
0.0458	1040512	CDS	Cj1108	clpA	ATP-dependent Clp protease ATP-binding subunit	1040450	1042579	1	T	AAT	N	synonymous	c	AAc	N
0.0458	1165697	CDS	Cj1237c		putative phosphatase	1165564	1166538	-1	C	ACG	T	nonsynonymous	t	AtG	M
0.046	63715	CDS	Cj0043	flgE	flagellar hook protein	62231	63868	1	T	GCT	A	synonymous	c	GCc	A
0.046	988639	CDS	Cj1052c	mutS	putative mismatch repair protein	987019	989229	-1	T	GGT	G	synonymous	c	GGc	G
0.0462	25874	CDS	Cj0020c		cytochrome C551 peroxidase	25433	26347	-1	T	GTT	V	MNP	c		

0.0462	247200	CDS	Cj0268c		putative transmembrane protein	246555	247643	-1	A	GTA	V	synonymous	g	GTg	V
0.0463	919078	CDS	Cj0985c	hipO	hippurate hydrolase	918580	919731	-1	C	AGC	S	synonymous	t	AGt	S
0.0464	295849	CDS	Cj0324	ubiE	ubiquinone/menaquinone biosynthesis methlytransferase	295586	296293	1	C	GGC	G	synonymous	t	GGt	G
0.0469	18417	CDS	Cj0013	ilvD	dihydroxy-acid dehydratase	17563	19239	1	A	GTA	V	synonymous	g	GTg	V
0.0469	18909	CDS	Cj0013	ilvD	dihydroxy-acid dehydratase	17563	19239	1	T	AGT	S	synonymous	c	AGc	S
0.0469	569621	CDS	Cj0609c		possible periplasmic protein	568946	570124	-1	G	AAG	K	synonymous	a	AAa	K
0.0469	1284457	CDS	Cj1352	ceuB	enterochelin uptake permease	1284008	1284976	1	T	GGT	G	synonymous	c	GGc	G
0.0469	1284520	CDS	Cj1352	ceuB	enterochelin uptake permease	1284008	1284976	1	T	GGT	G	synonymous	c	GGc	G
0.0469	1560164	CDS	Cj1634c	aroC	chorismate synthase	1559723	1560811	-1	A	ATA	I	MNP	g		
0.0469	1560166	CDS	Cj1634c	aroC	chorismate synthase	1559723	1560811	-1	A	ATA	I	nonsynonymous	g	gTg	V
0.0469	1560206	CDS	Cj1634c	aroC	chorismate synthase	1559723	1560811	-1	C	AGC	S	synonymous	t	AGt	S
0.0469	1560946	CDS	Cj1635c	rnc	ribonuclease III	1560811	1561485	-1	C	GGC	G	synonymous	t	GGt	G
0.0469	1561206	CDS	Cj1635c	rnc	ribonuclease III	1560811	1561485	-1	C	CTT	L	nonsynonymous	t	tTT	F
0.0472	24574	CDS	Cj0019c		putative MCP-domain signal transduction protein	23665	25443	-1	T	AAT	N	synonymous	c	AAc	N
0.0472	343563	CDS	Cj0376		putative periplasmic protein	343102	343926	1	A	AAA	K	synonymous	g	AAg	K
0.0473	1453496	CDS	Cj1518	moaE	putative molybdopterin converting factor,subunit 2	1453365	1453811	1	T	GGT	G	synonymous	c	GGc	G
0.0476	229668	CDS	Cj0248		hypothetical protein Cj0248	228946	229803	1	T	GGT	G	synonymous	c	GGc	G
0.0477	566641	CDS	Cj0607		ABC-type transmembrane transport protein	565646	567571	1	T	TCT	S	synonymous	c	TCc	S
0.0477	566642	CDS	Cj0607		ABC-type transmembrane transport protein	565646	567571	1	T	TTA	L	synonymous	c	cTA	L
0.0477	566686	CDS	Cj0607		ABC-type transmembrane transport protein	565646	567571	1	A	GGA	G	synonymous	t	GGt	G
0.0477	566953	CDS	Cj0607		ABC-type transmembrane transport protein	565646	567571	1	G	AAG	K	synonymous	a	AAa	K
0.0477	1277257	CDS	Cj1344c		putative glycoprotease	1276451	1277458	-1	C	CAT	H	nonsynonymous	t	tAT	Y

0.0477	1507728	CDS	Cj1576c	nuoD	NADH dehydrogenase I chain D	1506903	1508129	-1	G	GCG	A	synonymous	a	GCa	A
0.0479	55070	CDS	Cj0036		hypothetical protein Cj0036	53970	55319	1	T	TCT	S	synonymous	c	TCc	S
0.0479	349208	CDS	Cj0383c	ribH	6,7-dimethyl-8-ribityllumazine synthase	349139	349603	-1	A	AAA	K	synonymous	g	AAg	K
0.0479	1089050	CDS	Cj1156	rho	transcription termination factor	1088373	1089671	1	A	AGA	R	synonymous	g	AGg	R
0.048	1031582	CDS	Cj1097		putative transmembrane transport protein	1030953	1032176	1	G	TTG	L	synonymous	a	TTa	L
0.048	1166274	CDS	Cj1237c		putative phosphatase	1165564	1166538	-1	A	AAA	K	nonsynonymous	g	gAA	E
0.048	1166549	intergenic							A				g		
0.0482	1218461	CDS	Cj1286c	upp	uracil phosphoribosyltransferase	1218134	1218760	-1	G	GGG	G	synonymous	a	GGa	G
0.0486	581024	CDS	Cj0621		hypothetical protein Cj0621	580326	581846	1	T	TGT	C	synonymous	c	TGc	C
0.0489	591761	CDS	Cj0630c	holA	putative DNA polymerase III, delta subunit	591537	592502	-1	A	ATC	I	nonsynonymous	g	gTC	V
0.0491	64251	CDS	Cj0044c		hypothetical protein Cj0044c	63872	65743	-1	G	AGC	S	nonsynonymous	a	AaC	N
0.0491	264505	CDS	Cj0286c		hypothetical protein Cj0286c	264166	264783	-1	G	CTG	L	synonymous	t	CTt	L
0.0491	264982	CDS	Cj0287c	greA	transcription elongation factor	264934	265419	-1	A	GAA	E	synonymous	g	GAg	E
0.0491	265027	CDS	Cj0287c	greA	transcription elongation factor	264934	265419	-1	A	GGA	G	synonymous	g	GGg	G
0.0492	47919	CDS	Cj0031		putative type IIS restriction/modification enzyme	46424	50156	1	G	AGC	S	nonsynonymous	a	AaC	N
0.0493	1166410	CDS	Cj1237c		putative phosphatase	1165564	1166538	-1	T	TCT	S	synonymous	g	TCg	S
0.0494	338783	CDS	Cj0369c		putative ferredoxin domain-containing integral membrane protein	337535	338911	-1	T	CAT	H	synonymous	c	CAc	H
0.0494	1395816	CDS	Cj1458c	thiL	putative thiamin-monophosphate kinase	1395372	1396193	-1	C	AAC	N	synonymous	t	AAt	N
0.0496	1241532	CDS	Cj1311	pseF	putative acylneuraminate cytidylyltransferase	1240966	1241664	1	G	GAG	E	synonymous	a	GAa	E
0.0497	571317	CDS	Cj0611c		putative acyltransferase family protein	571122	572498	-1	C	AAC	N	synonymous	t	AAt	N
0.0498	1041606	CDS	Cj1108	clpA	ATP-dependent Clp protease ATP-binding subunit	1040450	1042579	1	G	AGA	R	nonsynonymous	a	AaA	K

0.0498	1041709	CDS	Cj1108	clpA	ATP-dependent Clp protease ATP-binding subunit	1040450	1042579	1	C	AAC	N	synonymous	t	AAt	N
0.0498	1042348	CDS	Cj1108	clpA	ATP-dependent Clp protease ATP-binding subunit	1040450	1042579	1	C	GAC	D	synonymous	t	GAt	D
0.0498	1042507	CDS	Cj1108	clpA	ATP-dependent Clp protease ATP-binding subunit	1040450	1042579	1	T	TTT	F	synonymous	c	TTc	F
0.0498	1042525	CDS	Cj1108	clpA	ATP-dependent Clp protease ATP-binding subunit	1040450	1042579	1	T	GGT	G	synonymous	c	GGc	G
0.0498	1042592	CDS	Cj1109	aat	leucyl/phenylalanyl-tRNA--protein transferase	1042554	1043201	1	G	GCG	A	synonymous	a	GCa	A

B.4 List of significant SNP associations (p<0.05) indicated by Scoary

Scoary p-value	Position	Feature	Locus tag	Gene	Product	Start	End	Strand	Ref. base	Ref. codon	Ref. res.	SNP type	New base	New codon	New res.
0.0390625	531	CDS	Cj0001	dnaA	chromosomal replication initiator protein	1	1323	1	C	AAC	N	synonymous	t	AAt	N
0.021484375	33286	CDS	Cj0025c		putative sodium:dicarboxylate family transmembrane symporter	32134	33519	-1	T	AGT	S	synonymous	c	AGc	S
0.0390625	33465	CDS	Cj0025c		putative sodium:dicarboxylate family transmembrane symporter	32134	33519	-1	G	GCT	A	nonsynonymous	a	aCT	T
0.03125	53763	CDS	Cj0035c		putative efflux protein	52665	53867	-1	T	CCT	P	synonymous	c	CCc	P
0.03125	53763	sig_peptide	Cj0035c			53748	53867		T				c		
0.03125	631981	CDS	Cj0679	kdpD	truncated KdpD protein	631667	633487	1	G	AAG	K	synonymous	a	AAa	K
0.03125	738902	CDS	Cj0788		hypothetical protein Cj0788	738771	739262	1	A	CCA	P	synonymous	g	CCg	P
0.03125	1546840	CDS	Cj1619	kgtP	alpha-ketoglutarate permease	1546496	1547755	1	G	TTG	L	synonymous	a	TTa	L
0.03125	1546858	CDS	Cj1619	kgtP	alpha-ketoglutarate permease	1546496	1547755	1	T	CTT	L	synonymous	c	CTc	L
0.021484375	1641351	CDS	Cj1731c	ruvC	crossover junction endodeoxyribonuclease	1640904	1641386	-1	G	TCG	S	synonymous	a	TCa	S

B.5 List of genes with SNPs indicated as significantly impacting the biofilm phenotype in this analysis, which were also listed as associated with the biofilm phenotype in the analysis performed by Pascoe et al. (265)**.**

Locus tag	**Gene name**	**Product**	**Proposed function** (265)	**Nonsyn. SNPs**	**Syn SNPs**	**MNPs**
Cj0146c	*trxB*	Thioredoxin reductase	Sensing oxidative stress	0	1	0
Cj1577c	*nuoC*	NADH dehydrogenase I chain C	Sensing oxidative stress	1	4	0
Cj0020c	-	Cytochrome C551 peroxidase		1	2	1
Cj0135	-	Conserved hypothetical protein Cj0135		0	1	0
Cj0136	*infB*	Translation initiation factor IF-2		0	9	0
Cj0137	*rbfA*	Putative ribosome-binding factor A		0	1	0
Cj0134	*thrB*	Putative homoserine kinase		0	1	0
Cj0150c		Aminotransferase		2	0	0
Cj0151c		Putative periplasmic protein	Metal uptake	1	0	0
Cj0267c		Membrane protein	Sensing oxidative stress	0	1	0
Cj0268c		Membrane protease family protein HP0248		1	1	0
Cj0271		Thiol peroxidase, Bcp-type	Iron uptake	1	2	1
Cj0272		Diacylglucosamine hydrolase like		3	4	0
Cj1411c		cytochrome p450		0	1	0
Cj1447c	*kpsT*	ABC transporter, permease protein KpsM		0	3	0
Cj1581c			Nickel transport	1	0	0
Cj1586	*cgb*	Single domain hemoglobin	Nitrosative stress	0	2	0

References

1. Klapper I, Dockery J. Mathematical Description of Microbial Biofilms. *SIAM Rev [Internet]*. 2010 Jan 6 [cited 2019 Feb 21];52(2):221–65. Available from: http://epubs.siam.org/doi/10.1137/080739720

2. Sheppard SK, Maiden MCJ. The evolution of *Campylobacter jejuni* and *Campylobacter coli*. *Cold Spring Harb Perspect Biol [Internet]*. 2015 Jun 22 [cited 2019 Mar 11];7(8):1–13. Available from:

http://www.ncbi.nlm.nih.gov/pubmed/26101080

3. Cody AJ, Maiden MCJ, Strachan NJC, McCarthy ND. A systematic review of source attribution of human campylobacteriosis using multilocus sequence typing. *Eurosurveillance*. 2019 Oct 24;24(43).

4. Kaakoush NO, Castaño-Rodríguez N, Mitchell HM, Man SM. Global Epidemiology of *Campylobacter* Infection. *Clin Microbiol Rev [Internet]*. 2015 Jul 1 [cited 2017 Jul 25];28(3):687–720. Available from: http://www.ncbi.nlm.nih.gov/pubmed/26062576

5. Borck Høg B, Sommer HM, Larsen LS, Sørensen AIV, David B, Hofshagen M, et al. Farm specific risk factors for *Campylobacter* colonisation in Danish and Norwegian broilers. *Prev Vet Med*. 2016 Aug 1;130:137–45.

6. Royden A, Wedley A, Merga JY, Rushton S, Hald B, Humphrey T, et al. A role for flies (Diptera) in the transmission of *Campylobacter* to broilers? *Epidemiol Infect*. 2016 Nov 1;144(15):3326–34.

7. Fonseca BB, Fernandez H, Rossi DA. *Campylobacter spp.* and Related Organisms in Poultry [Internet]. Minas Gerais, Brazil and Valdivia, Chile: Springer International Publishing; 2016 [cited 2017 Aug 29]. Available from: https://link.springer.com/content/pdf/10.1007/978-3-319-29907-5.pdf

8. Snelling WJ, Matsuda M, Moore JE, Dooley JSG. *Campylobacter jejuni*. *Lett Appl Microbiol [Internet]*. 2005 Oct 1 [cited 2019 Mar 20];41(4):297–302. Available from: http://doi.wiley.com/10.1111/j.1472-765X.2005.01788.x

9. Altekruse SF, Stern NJ, Fields PI, Swerdlow DL. *Campylobacter jejuni*- An Emerging Foodborne Pathogen. *Emerg Infect Dis*. 1999;5(1):28–35.

10. Authority(EFSA) EFS. The European Union summary report on trends and sources of zoonoses, zoonotic agents and food-borne outbreaks in 2014 [Internet]. Vol. 13, EFSA Journal. 2015 Dec [cited 2019 Mar 12]. Available from: http://doi.wiley.com/10.2903/j.efsa.2015.4329

11. Cha W, Henderson T, Collins J, Manning SD. Factors associated with increasing campylobacteriosis incidence in Michigan, 2004-2013. *Epidemiol Infect*. 2016 Nov 1;144(15):3316–25.

12. Tam C, Viviani L, Adak B, Bolton E, Dodds J, Cowden J, et al. The Second Study of Infectious Intestinal Disease in the Community (IID2 Study) Project Number: B18021 Funder: UK Food Standards Agency Final Report Report Authors [Internet]. [cited 2020 Jun 4]. Available from: www.gutfeelings.org.uk

13. European Centre for Disease Prevention and Control (ECDC), Authority(EFSA) EFS. *Campylobacter* and *Lysteria* infections still rising in the EU-say EFSA and ECDC. 2015;(December).

14. Authority(EFSA) EFS. The European Union One Health 2018 Zoonoses Report [Internet]. 2019 [cited 2020 Mar 30]. Available from: https://www.efsa.europa.eu/sites/default/files/corporate_publications/files/5926-pls-zoonoses-report-2018.pdf

15. National Notifiable Diseases Surveillance System—notification rate of campylobacteriosis. [Internet]. Australian Government Department of Health. 2020 [cited 2020 Mar 30]. Available from: http://www9.health.gov.au/cda/source/rpt_3.cfm

16. Blaser M, Wells JG, Feldman RA, Pollard RA, Allen JR. *Campylobacter* Enteritis in the United States. *Ann Intern Med [Internet]*. 1983 Mar 1 [cited 2019 Mar 13];98(3):360. Available from: http://annals.org/article.aspx?doi=10.7326/0003-4819-98-3-360

17. O'Reilly LC, Inglis TJJ, Unicomb L, Australian *Campylobacter* Subtyping Study Group. Australian multicentre comparison of subtyping methods for the investigation of *Campylobacter* infection. *Epidemiol Infect [Internet]*. 2006 Aug 25 [cited 2017 Mar 6];134(4):768–79. Available from: http://www.journals.cambridge.org/abstract_S0950268805005777

18. WHO. *Campylobacter* [Internet]. https://www.who.int/news-room/fact-sheets/detail/campylobacter. 2018 [cited 2020 Mar 30]. Available from: https://www.who.int/news-room/fact-sheets/detail/campylobacter

19. Adak GK, Meakins SM, Yip H, Lopman BA, O'Brien SJ. Disease risks from foods, England and Wales, 1996-2000. *Emerg Infect Dis [Internet]*. 2005 Mar [cited 2019 Mar 11];11(3):365–72. Available from: http://www.ncbi.nlm.nih.gov/pubmed/15757549

20. Walker LJ, Wallace RL, Smith JJ, Graham T, Saputra T, Symes S, et al. Prevalence of *Campylobacter coli* and *Campylobacter jejuni* in Retail Chicken, Beef, Lamb, and Pork Products in Three Australian States. *J Food Prot*. 2019 Dec 1;82(12):2126–34.

21. Hald B, Sommer HM, Skovgård H. Use of Fly Screens to Reduce *Campylobacter* spp. Introduction in Broiler Houses. *Emerg Infect Dis [Internet]*. 2007 Dec [cited 2019 Jun 5];13(12):1951–3. Available from:

http://wwwnc.cdc.gov/eid/article/13/12/07-0488_article.htm

22. Bahrndorff S, Rangstrup-Christensen L, Nordentoft S, Hald B. Foodborne disease prevention and broiler chickens with reduced *Campylobacter* infection. *Emerg Infect Dis*. 2013 Mar 1;19(3):425–30.

23. Lee MD, Newell DG. *Campylobacter* in Poultry: Filling an Ecological Niche. *Avian Dis*. 2006 Mar 1;50(1):1–9.

24. Murphy C, Carroll C, Jordan KN. Environmental survival mechanisms of the foodborne pathogen *Campylobacter jejuni*. *J Appl Microbiol [Internet]*. 2006 Apr 1 [cited 2017 Aug 28];100(4):623–32. Available from: http://doi.wiley.com/10.1111/j.1365-2672.2006.02903.x

25. Brown HL, Reuter M, Salt LJ, Cross KL, Betts RP, van Vliet AHM. Chicken juice enhances surface attachment and biofilm formation of *Campylobacter jejuni*. *Appl Environ Microbiol [Internet]*. 2014 Nov 15 [cited 2017 Jul 4];80(22):7053–60. Available from: http://www.ncbi.nlm.nih.gov/pubmed/25192991

26. Teh AHT, Lee SM, Dykes GA. Does *Campylobacter jejuni* form biofilms in food-related environments? *Appl Environ Microbiol [Internet]*. 2014 Sep 1 [cited 2017 Oct 18];80(17):5154–60. Available from: http://www.ncbi.nlm.nih.gov/pubmed/24928882

27. Joshua GWP, Guthrie-Irons C, Karlyshev A V., Wren BW. Biofilm formation in *Campylobacter jejuni*. *Microbiology [Internet]*. 2006;152(2):387–96. Available from: http://www.ncbi.nlm.nih.gov/pubmed/16436427

28. Humphrey S, Lacharme-Lora L, Chaloner G, Gibbs K, Humphrey T, Williams N, et al. Heterogeneity in the Infection Biology of *Campylobacter jejuni* Isolates in Three Infection Models Reveals an Invasive and Virulent Phenotype in a ST21 Isolate from Poultry. Jacobsen ID, editor. *PLoS One [Internet]*. 2015 Oct 23 [cited 2019 Jul 17];10(10):e0141182. Available from: https://dx.plos.org/10.1371/journal.pone.0141182

29. Hofreuter D. Defining the metabolic requirements for the growth and colonization capacity of *Campylobacter jejuni*. *Front Cell Infect Microbiol [Internet]*. 2014 Sep 29 [cited 2019 May 21];4:137. Available from: http://journal.frontiersin.org/article/10.3389/fcimb.2014.00137/abstract

30. Svensson SL, Pryjma M, Gaynor EC. Flagella-Mediated Adhesion and Extracellular DNA Release Contribute to Biofilm Formation and Stress Tolerance of

Campylobacter jejuni. Roop RM, editor. *PLoS One [Internet]*. 2014 Aug 28 [cited 2017 Dec 1];9(8):e106063. Available from: http://dx.plos.org/10.1371/journal.pone.0106063

31. McCarthy Z, Smith B, Fazil A, Wu J, Ryan SD, Munther D. pH dependent *C. jejuni* thermal inactivation models and application to poultry scalding. *J Food Eng [Internet]*. 2018 Apr 1 [cited 2019 May 27];223:1–9. Available from: https://www.sciencedirect.com/science/article/abs/pii/S0260877417305009

32. Metris A, Reuter M, Gaskin DJ, Baranyi J, van Vliet AH. In vivo and in silico determination of essential genes of *Campylobacter jejuni*. *BMC Genomics [Internet]*. 2011 Dec 1 [cited 2019 May 1];12(1):535. Available from: http://bmcgenomics.biomedcentral.com/articles/10.1186/1471-2164-12-535

33. Richelle A, David B, Demaegd D, Dewerchin M, Kinet R, Morreale A, et al. Towards a widespread adoption of metabolic modeling tools in biopharmaceutical industry: a process systems biology engineering perspective. *npj Syst Biol Appl [Internet]*. 2020 Dec 1 [cited 2020 Jul 9];6(1):1–5. Available from: https://doi.org/10.1038/s41540-020-0127-y

34. Damte D, Suh JW, Lee SJ, Yohannes SB, Hossain MA, Park SC. Putative drug and vaccine target protein identification using comparative genomic analysis of KEGG annotated metabolic pathways of *Mycoplasma hyopneumoniae*. *Genomics*. 2013 Jul 1;102(1):47–56.

35. Ferrarini MG, Siqueira FM, Mucha SG, Palama TL, Jobard É, Elena-Herrmann B, et al. Insights on the virulence of swine respiratory tract mycoplasmas through genome-scale metabolic modeling. *BMC Genomics [Internet]*. 2016 Dec 13 [cited 2020 Jul 9];17(1):353. Available from: http://bmcgenomics.biomedcentral.com/articles/10.1186/s12864-016-2644-z

36. Giallourou N, Medlock GL, Bolick DT, Medeiros PH, Ledwaba SE, Kolling GL, et al. A novel mouse model of *Campylobacter jejuni* enteropathy and diarrhea. Baumler AJ, editor. *PLOS Pathog [Internet]*. 2018 May 23 [cited 2019 Jun 4];14(3):e1007083. Available from: http://dx.plos.org/10.1371/journal.ppat.1007083

37. Gradel KO, Nielsen HL, Schønheyder HC, Ejlertsen T, Kristensen B, Nielsen H. Increased short- and long-term risk of inflammatory bowel disease after *Salmonella* or *Campylobacter* gastroenteritis. *Gastroenterology [Internet]*. 2009 Aug 1 [cited 2019 Jul 18];137(2):495–501. Available from: http://www.ncbi.nlm.nih.gov/pubmed/19361507

38. Bae J, Oh E, Jeon B. Enhanced transmission of antibiotic resistance in *Campylobacter jejuni* biofilms by natural transformation. *Antimicrob Agents Chemother*. 2014 Dec 1;58(12):7573–5.

39. O'Toole G, Kaplan HB, Kolter R. Biofilm Formation as Microbial Development. *Annu Rev Microbiol [Internet]*. 2000 Oct 28 [cited 2019 Feb 4];54(1):49–79. Available from: http://www.annualreviews.org/doi/10.1146/annurev.micro.54.1.49

40. Reeser RJ, Medler RT, Billington SJ, Jost BH, Joens LA. Characterization of *Campylobacter jejuni* Biofilms under Defined Growth Conditions. *Appl Environ Microbiol [Internet]*. 2007 Mar 15 [cited 2018 Feb 20];73(6):1908–13. Available from: http://aem.asm.org/content/73/6/1908.full.pdf

41. Kalmokoff M, Lanthier P, Tremblay T-L, Foss M, Lau PC, Sanders G, et al. Proteomic analysis of *Campylobacter jejuni* 11168 biofilms reveals a role for the motility complex in biofilm formation. *J Bacteriol [Internet]*. 2006 Jun 15 [cited 2019 Feb 5];188(12):4312–20. Available from: http://www.ncbi.nlm.nih.gov/pubmed/16740937

42. Feng J, Ma L, Nie J, Konkel ME, Lu X. Environmental Stress-Induced Bacterial Lysis and Extracellular DNA Release Contribute to *Campylobacter jejuni* Biofilm Formation. *Appl Environ Microbiol [Internet]*. 2018 [cited 2018 Feb 15];84(5):1–18. Available from: http://aem.asm.org/content/84/5/e02068-17.full.pdf

43. Brown HL, Hanman K, Reuter M, Betts RP, van Vliet AHM. *Campylobacter jejuni* biofilms contain extracellular DNA and are sensitive to DNase I treatment. *Front Microbiol [Internet]*. 2015 Jul 10 [cited 2017 Dec 1];6:699. Available from: http://journal.frontiersin.org/Article/10.3389/fmicb.2015.00699/abstract

44. Gaasbeek EJ, Wagenaar JA, Guilhabert MR, van Putten JPMM, Parker CT, van der Wal FJ. Nucleases encoded by the integrated elements CJIE2 and CJIE4 inhibit natural transformation of *Campylobacter jejuni*. *J Bacteriol [Internet]*. 2010 Feb 15 [cited 2019 May 22];192(4):936–41. Available from: http://www.ncbi.nlm.nih.gov/pubmed/20023031

45. Vorkapic D, Pressler K, Schild S. Multifaceted roles of extracellular DNA in bacterial physiology. *Curr Genet [Internet]*. 2016 Feb 2 [cited 2019 May 22];62(1):71–9. Available from: http://link.springer.com/10.1007/s00294-015-0514-x

46. de Vries SP, Gupta S, Baig A, Wright E, Wedley A, Jensen AN, et al. Genome-wide fitness analyses of the foodborne pathogen *Campylobacter jejuni* in in vitro and in

vivo models. *Sci Rep [Internet]*. 2017 Dec 28 [cited 2019 May 21];7(1):1251. Available from: http://www.nature.com/articles/s41598-017-01133-4

47. Reuter M, Mallett A, Pearson BM, Van Vliet AHM. Biofilm formation by *Campylobacter jejuni* is increased under aerobic conditions. *Appl Environ Microbiol*. 2010;76(7):2122–8.

48. Teh AHT, Lee SM, Dykes GA. The influence of dissolved oxygen level and medium on biofilm formation by *Campylobacter jejuni*. *Food Microbiol [Internet]*. 2017 Feb 1 [cited 2019 Mar 25];61:120–5. Available from: https://www.sciencedirect.com/science/article/pii/S0740002016302027

49. Teh AHT, Lee SM, Dykes GA. The Influence of Prior Modes of Growth, Temperature, Medium, and Substrate Surface on Biofilm Formation by Antibiotic-Resistant *Campylobacter jejuni*. *Curr Microbiol [Internet]*. 2016 Dec 13 [cited 2017 Oct 18];73(6):859–66. Available from: http://link.springer.com/10.1007/s00284-016-1134-5

50. Moore JE. Comparison of basal broth media for the optimal laboratory recovery of *Campylobacter jejuni* and *Campylobacter coli*. *Ir J Med Sci [Internet]*. 2000 [cited 2020 Jun 23];169(3):187–9. Available from: https://link.springer.com/article/10.1007/BF03167693

51. Kim SH, Park C, Lee EJ, Bang WS, Kim YJ, Kim JS. Biofilm formation of *Campylobacter* strains isolated from raw chickens and its reduction with DNase I treatment. *Food Control*. 2017 Jan 1;71:94–100.

52. Dzianach PA, Dykes GA, Strachan NJC, Forbes KJ, Pérez-Reche FJ. Challenges of biofilm control and utilization: lessons from mathematical modelling. *J R Soc Interface [Internet]*. 2019 Jun 28 [cited 2019 Jul 11];16(155):20190042. Available from: https://royalsocietypublishing.org/doi/10.1098/rsif.2019.0042

53. Zweifel C, Stephan R. Microbial decontamination of poultry carcasses. In: Demirci A, Ngadi MO, editors. Microbial decontamination in the food industry : novel methods and applications. Zurich: Woodhead Publishing 2012; 2012. p. 804.

54. Zhang L, Garner LJ, McKee SR, Bilgili SF. Effectiveness of several antimicrobials used in a postchill decontamination tank against *Salmonella* and *Campylobacter* on broiler carcass parts. *J Food Prot*. 2018 Jul 1;81(7):1134–41.

55. Kure CF, Axelsson L, Carlehög M, Måge I, Jensen MR, Holck A. The effects of a pilot-scale steam decontamination system on the hygiene and sensory quality of

chicken carcasses. *Food Control*. 2020 Mar 1;109:106948.

56. Scientific Opinion on *Campylobacter* in broiler meat production: control options and performance objectives and/or targets at different stages of the food chain. *EFSA J [Internet]*. 2011 Apr 1 [cited 2020 Jul 14];9(4):2105. Available from: http://doi.wiley.com/10.2903/j.efsa.2011.2105

57. Yoon KS, Burnette CN, Oscar TP. Development of Predictive Models for the Survival of *Campylobacter jejuni* (ATCC 43051) on Cooked Chicken Breast Patties and in Broth as a Function of Temperature. *J Food Prot [Internet]*. 2004 [cited 2017 Dec 1];67(1):64–70. Available from: http://jfoodprotection.org/doi/pdf/10.4315/0362-028X-67.1.64

58. Davey KR. Applicability of the Davey (linear Arrhenius) predictive model to the lag phase of microbial growth. *J Appl Bacteriol [Internet]*. 1991 Mar 1 [cited 2019 May 28];70(3):253–7. Available from: http://doi.wiley.com/10.1111/j.1365-2672.1991.tb02933.x

59. Geeraerd AH, Valdramidis VP, Van Impe JF. GInaFiT, a freeware tool to assess non-log-linear microbial survivor curves. *Int J Food Microbiol*. 2005 Jun 25;102(1):95–105.

60. González M, Skandamis PN, Hänninen M-L. A modified Weibull model for describing the survival of *Campylobacter jejuni* in minced chicken meat. *Int J Food Microbiol [Internet]*. 2009 Nov 30 [cited 2019 May 27];136(1):52–8. Available from: https://www.sciencedirect.com/science/article/pii/S0168160509005054

61. Hong S, Kim H, Yoon K, Hong SH, Kim HS, Yoon KS. Survival and Risk Comparison of *Campylobacter jejuni* on Various Processed Meat Products. *Int J Environ Res Public Health [Internet]*. 2016 Jun 9 [cited 2019 May 29];13(6):580. Available from: http://www.mdpi.com/1660-4601/13/6/580

62. Blankenship LC, Craven SE. *Campylobacter jejuni* survival in chicken meat as a function of temperature. *Appl Environ Microbiol [Internet]*. 1982 Jul 1 [cited 2019 May 29];44(1):88–92. Available from: https://aem-asm-org.dbgw.lis.curtin.edu.au/content/44/1/88

63. Chan KF, Le Tran H, Kanenaka RY, Kathariou S. Survival of clinical and poultry-derived isolates of *Campylobacter jejuni* at a low temperature (4 degrees C). *Appl Environ Microbiol [Internet]*. 2001 Sep 1 [cited 2019 May 29];67(9):4186–91. Available from: http://www.ncbi.nlm.nih.gov/pubmed/11526022

64. Garénaux A, Jugiau F, Rama F, de Jonge R, Denis M, Federighi M, et al. Survival of *Campylobacter jejuni* Strains from Different Origins Under Oxidative Stress Conditions: Effect of Temperature. *Curr Microbiol [Internet]*. 2008 Apr 5 [cited 2019 May 29];56(4):293–7. Available from: http://link.springer.com/10.1007/s00284-007-9082-8

65. Umaraw P, Prajapati A, Verma AK, Pathak V, Singh VP. Control of *Campylobacter* in poultry industry from farm to poultry processing unit: A review. *Crit Rev Food Sci Nutr [Internet]*. 2017 Mar 4 [cited 2020 Jul 10];57(4):659–65. Available from: https://www.tandfonline.com/doi/abs/10.1080/10408398.2014.935847

66. Valtierra-Rodríguez D, Heredia NL, García S, Sánchez E. Reduction of *Campylobacter jejuni* and *Campylobacter coli* in poultry skin by fruit extracts. *J Food Prot [Internet]*. 2010 Mar 1 [cited 2020 Jul 14];73(3):477–82. Available from: http://meridian.allenpress.com/jfp/article-pdf/73/3/477/1680785/0362-028x-73_3_477.pdf

67. Oberhardt MA, Palsson BØ, Papin JA. Applications of genome-scale metabolic reconstructions. *Mol Syst Biol [Internet]*. 2009 Nov 3 [cited 2019 Jul 17];5(320):1–15. Available from: http://msb.embopress.org/cgi/doi/10.1038/msb.2009.77

68. Cole JA, Kohler L, Hedhli J, Luthey-Schulten Z. Spatially-resolved metabolic cooperativity within dense bacterial colonies. *BMC Syst Biol [Internet]*. 2015 Dec 18 [cited 2019 Apr 29];9(1):15. Available from: http://www.biomedcentral.com/1752-0509/9/15

69. Phalak P, Chen J, Carlson RP, Henson MA. Metabolic modeling of a chronic wound biofilm consortium predicts spatial partitioning of bacterial species. *BMC Syst Biol [Internet]*. 2016 Dec 7 [cited 2019 Mar 4];10(1):90. Available from: http://bmcsystbiol.biomedcentral.com/articles/10.1186/s12918-016-0334-8

70. Guccione E, del Rocio Leon-Kempis M, Pearson BM, Hitchin E, Mulholland F, van Diemen PM, et al. Amino acid-dependent growth of *Campylobacter jejuni* : key roles for aspartase (AspA) under microaerobic and oxygen-limited conditions and identification of AspB (Cj0762), essential for growth on glutamate. *Mol Microbiol [Internet]*. 2008 Jul 1 [cited 2019 Mar 15];69(1):77–93. Available from: http://doi.wiley.com/10.1111/j.1365-2958.2008.06263.x

71. Muraoka WT, Zhang Q. Phenotypic and genotypic evidence for L-fucose utilization by *Campylobacter jejuni*. *J Bacteriol [Internet]*. 2011 Mar 1 [cited 2019 Mar 15];193(5):1065–75. Available from:

http://www.ncbi.nlm.nih.gov/pubmed/21193610

72. Newell DG. Animal models of *Campylobacter jejuni* colonization and disease and the lessons to be learned from similar *Helicobacter pylori* models. *J Appl Microbiol [Internet]*. 2001 Jun 1 [cited 2019 Mar 20];90(S6):57S-67S. Available from: http://doi.wiley.com/10.1046/j.1365-2672.2001.01354.x

73. Black RE, Perlman D, Clements ML, Levine MM, Blaser MJ. Human Volunteer Studies with *Campylobacter jejuni*. In: Nachamkin I, Blaser MJ, Tompkins L., editors. Campylobacter jejuni: Current Status and Future Trends. Washington: ASM Press; 1993 [cited 2019 May 24]. p. 207–15. Available from: https://apps.dtic.mil/docs/citations/ADA271892

74. Bacon DJ, Alm RA, Burr DH, Hu L, Kopecko DJ, Ewing CP, et al. Involvement of a plasmid in virulence of *Campylobacter jejuni* 81-176. *Infect Immun [Internet]*. 2000 Aug 1 [cited 2019 May 24];68(8):4384–90. Available from: http://www.ncbi.nlm.nih.gov/pubmed/10899834

75. Bereswill S, Fischer A, Plickert R, Haag L-M, Otto B, Kühl AA, et al. Novel Murine Infection Models Provide Deep Insights into the "Ménage à Trois" of *Campylobacter jejuni*, Microbiota and Host Innate Immunity. Metzger DW, editor. *PLoS One [Internet]*. 2011 Jun 15 [cited 2019 May 22];6(6):e20953. Available from: https://dx.plos.org/10.1371/journal.pone.0020953

76. Champion OL, Wagley S, Titball RW. *Galleria mellonella* as a model host for microbiological and toxin research. *Virulence [Internet]*. 2016 Oct 2 [cited 2019 Jul 17];7(7):840–5. Available from: https://www.tandfonline.com/doi/full/10.1080/21505594.2016.1203486

77. Tsai CJ-Y, Loh JMS, Proft T. *Galleria mellonella* infection models for the study of bacterial diseases and for antimicrobial drug testing. *Virulence [Internet]*. 2016 Apr 2 [cited 2019 Jul 17];7(3):214–29. Available from: https://www.tandfonline.com/doi/full/10.1080/21505594.2015.1135289

78. Senior NJ, Bagnall MC, Champion OL, Reynolds SE, La Ragione RM, Woodward MJ, et al. *Galleria mellonella* as an infection model for *Campylobacter jejuni* virulence. *J Med Microbiol [Internet]*. 2011 May 1 [cited 2019 Jul 17];60(5):661–9. Available from: http://jmm.microbiologyresearch.org/content/journal/jmm/10.1099/jmm.0.026658-0

79. Teunis PFM, Nagelkerke NJD, Haas CN. Dose Response Models For Infectious

Gastroenteritis. *Risk Anal*. 1999 Dec 1;19(6):1251–60.

80. Teunis PFM, Kasuga F, Fazil A, Ogden ID, Rotariu O, Strachan NJC. Dose-response modeling of *Salmonella* using outbreak data. *Int J Food Microbiol*. 2010 Dec 15;144(2):243–9.

81. Strachan NJC, Doyle MP, Kasuga F, Rotariu O, Ogden ID. Dose response modelling of *Escherichia coli* O157 incorporating data from foodborne and environmental outbreaks. *Int J Food Microbiol*. 2005 Aug 15;103(1):35–47.

82. Teunis PFM, Ogden ID, Strachan NJC. Hierarchical dose response of *E. coli* O157:H7 from human outbreaks incorporating heterogeneity in exposure. *Epidemiol Infect*. 2008 Jun;136(6):761–70.

83. Xie G, Roiko A, Stratton H, Lemckert C, Dunn PK, Mengersen K. A Generalized QMRA Beta-Poisson Dose-Response Model. *Risk Anal [Internet]*. 2016 Oct 1 [cited 2020 Jun 4];36(10):1948–58. Available from: http://doi.wiley.com/10.1111/risa.12561

84. Langley GR, Adcock IM, Busquet F, Crofton KM, Csernok E, Giese C, et al. Towards a 21st-century roadmap for biomedical research and drug discovery: consensus report and recommendations. Vol. 22, Drug Discovery Today. Elsevier Ltd; 2017. p. 327–39.

85. Wagenaar JA, French NP, Havelaar AH. Preventing *Campylobacter* at the Source: Why Is It So Difficult? *Clin Infect Dis*. 2013;57(11):1600–1606.

86. Brauer F. Mathematical epidemiology: Past, present, and future. Vol. 2, Infectious Disease Modelling. KeAi Communications Co.; 2017. p. 113–27.

87. Yébenes JC, Ruiz-Rodriguez JC, Ferrer R, Clèries M, Bosch A, Lorencio C, et al. Epidemiology of sepsis in Catalonia: analysis of incidence and outcomes in a European setting. *Ann Intensive Care*. 2017 Dec 1;7(1):1–10.

88. Kleinbaum DG, Kupper LL, Chambless LE. Logistic regression analysis of epidemiologic data: Theory and practice. *Commun Stat - Theory Methods*. 1982 Jan 1;11(5):485–547.

89. Bavishi C, DuPont HL. Systematic review: The use of proton pump inhibitors and increased susceptibility to enteric infection [Internet]. Vol. 34, Alimentary Pharmacology and Therapeutics. John Wiley & Sons, Ltd; 2011 [cited 2020 Jul 28]. p. 1269–81. Available from: https://onlinelibrary.wiley.com/doi/full/10.1111/j.1365-2036.2011.04874.x

90. Futagami S, Itoh T, Sakamoto C. Systematic review with meta-analysis: post-infectious functional dyspepsia. *Aliment Pharmacol Ther [Internet]*. 2015 Jan 1 [cited 2019 Jul 18];41(2):177–88. Available from: http://doi.wiley.com/10.1111/apt.13006

91. Riddle MS, Murray JA, Cash BD, Pimentel M, Porter CK. Pathogen-Specific Risk of Celiac Disease Following Bacterial Causes of Foodborne Illness: A Retrospective Cohort Study. *Dig Dis Sci [Internet]*. 2013 Nov 29 [cited 2019 Jul 18];58(11):3242–5. Available from: http://link.springer.com/10.1007/s10620-013-2733-7

92. Sears A, Baker MG, Wilson N, Marshall J, Muellner P, Campbell DM, et al. Marked Campylobacteriosis Decline after Interventions Aimed at Poultry, New Zealand. *Emerg Infect Dis [Internet]*. 2011 Jun [cited 2019 Jul 18];17(6):1007–15. Available from: http://wwwnc.cdc.gov/eid/article/17/6/10-1272_article.htm

93. Yahara K, Méric G, Taylor AJ, de Vries SPW, Murray S, Pascoe B, et al. Genome-wide association of functional traits linked with *Campylobacter jejuni* survival from farm to fork. *Environ Microbiol [Internet]*. 2017 Jan 1 [cited 2021 Jan 29];19(1):361–80. Available from: https://pubmed.ncbi.nlm.nih.gov/27883255/

94. Bender R. Quantitative Risk Assessment in Epidemiological Studies Investigating Threshold Effects. *Biometrical J*. 1999 Jun 1;41(3):305–19.

95. Duffy LL, Blackall PJ, Cobbold RN, Fegan N. Quantitative effects of in-line operations on *Campylobacter* and *Escherichia coli* through two Australian broiler processing plants. *Int J Food Microbiol [Internet]*. 2014 Oct 1 [cited 2018 Aug 15];188:128–34. Available from: https://www.sciencedirect.com/science/article/pii/S0168160514003675

96. Muller CJ, Maclehose RF. Estimating predicted probabilities from logistic regression: different methods correspond to different target populations. [cited 2020 Jun 6]; Available from: https://academic.oup.com/ije/article-abstract/43/3/962/763470

97. Klapper I, Dockery J. Mathematical Descriptions of Microbial Biofilms. *SIAM Rev*. 2010;52(2):221–65.

98. Shekhar S, Sundaramanickam A, Balasubramanian T. Biosurfactant Producing Microbes and their Potential Applications: A Review. *Crit Rev Environ Sci Technol [Internet]*. 2015 Jul 18 [cited 2019 Feb 20];45(14):1522–54. Available from: http://www.tandfonline.com/doi/full/10.1080/10643389.2014.955631

99. Hasan F, Shah AA, Hameed A. Industrial applications of microbial lipases. *Enzyme Microb Technol [Internet]*. 2006 Jun 26 [cited 2019 Feb 20];39(2):235–51. Available

from: https://www.sciencedirect.com/science/article/pii/S0141022905004606

100. De Vuyst L, Leroy F. Bacteriocins from Lactic Acid Bacteria: Production, Purification, and Food Applications. *J Mol Microbiol Biotechnol [Internet]*. 2007 [cited 2019 Feb 20];13(4):194–9. Available from: http://www.ncbi.nlm.nih.gov/pubmed/17827969

101. Rabaey K, Verstraete W. Microbial fuel cells: novel biotechnology for energy generation. *Trends Biotechnol [Internet]*. 2005 Jun 1 [cited 2018 Jul 24];23(6):291–8. Available from: https://www.sciencedirect.com/science/article/pii/S0167779905000922

102. Gasner LL. Microorganisms for Waste Treatment. In: Peppler HJ, Perlman D, editors. Microbial Technology. New York City: Academic Press; 1979 [cited 2019 Feb 20]. p. 211–22. Available from: https://www.sciencedirect.com/science/article/pii/B9780125515023500153

103. Rawlings DE. Biomining : theory, microbes and industrial processes. Rawlings DE, editor. Rondebosh: Springer Science & Business Media; 2013. 302 p.

104. Karunakaran E, Mukherjee J, Ramalingam B, Biggs CA. "Biofilmology": a multidisciplinary review of the study of microbial biofilms. *Appl Microbiol Biotechnol [Internet]*. 2011 Jun 3 [cited 2019 Mar 4];90(6):1869–81. Available from: http://link.springer.com/10.1007/s00253-011-3293-4

105. Wong GCL, O'Toole GA. All together now: Integrating biofilm research across disciplines. *MRS Bull [Internet]*. 2011 May 18 [cited 2019 Mar 4];36(5):339–42. Available from: http://www.journals.cambridge.org/abstract_S0883769411000649

106. Jayasinghe N, Mahadevan R. Metabolic Modeling of Spatial Heterogeneity of Biofilms in Microbial Fuel Cells. *IFAC Proc Vol [Internet]*. 2010 Jan 1 [cited 2019 Mar 4];43(6):215–20. Available from: https://www.sciencedirect.com/science/article/pii/S1474667016303573

107. Davey ME, O'toole GA. Microbial biofilms: from ecology to molecular genetics. *Microbiol Mol Biol Rev [Internet]*. 2000 Dec 1 [cited 2018 Jun 13];64(4):847–67. Available from: http://www.ncbi.nlm.nih.gov/pubmed/11104821

108. Marshall KC. Planktonic Versus Sessile Life of Prokaryotes. In: Rosenberg E, DeLong EF, Lory S, Stackebrandt E, Thompson F, editors. The Prokaryotes. Berlin, Heidelberg: Springer Berlin Heidelberg; 2013 [cited 2019 Jan 7]. p. 191–201. Available from: http://link.springer.com/10.1007/978-3-642-30123-0_49

109. Costerton J, Lewandowski Z, Caldwell D, Lappin-Scott H. Microbial biofilms. *Annu Rev Microbiol [Internet]*. 1995;49:711–45. Available from: http://www.annualreviews.org/doi/pdf/10.1146/annurev.mi.49.100195.003431

110. Moore WEC, Moore LVH. The bacteria of periodontal diseases. *Periodontol 2000 [Internet]*. 1994 Jun 1 [cited 2019 Feb 21];5(1):66–77. Available from: http://doi.wiley.com/10.1111/j.1600-0757.1994.tb00019.x

111. Sutherland IW. The biofilm matrix - An immobilized but dynamic microbial environment. *Trends Microbiol*. 2001;9(5):222–7.

112. Young IM, Crawford JW. Interactions and self-organization in the soil-microbe complex. *Science (80-)*. 2004;304(5677):1634–7.

113. Pérez-Reche FJ, Taraskin SN, Otten W, Viana MP, Costa LDF, Gilligan C a., et al. Prominent Effect of Soil Network Heterogeneity on Microbial Invasion. *Phys Rev Lett [Internet]*. 2012 Aug 18 [cited 2018 Nov 27];109(9):098102. Available from: http://arxiv.org/abs/1209.3974

114. Donlan RM, Costerton JW. Biofilms: survival mechanisms of clinically relevant microorganisms. *Clin Microbiol Rev [Internet]*. 2002 Apr 1 [cited 2019 Feb 21];15(2):167–93. Available from: http://www.ncbi.nlm.nih.gov/pubmed/11932229

115. Costerton JW, Cheng KJ, Geesey GG, Ladd TI, Nickel JC, Dasgupta M, et al. Bacterial Biofilms in Nature and Disease. *Annu Rev Microbiol [Internet]*. 1987;41(1):435–64. Available from: http://www.annualreviews.org/doi/abs/10.1146/annurev.mi.41.100187.002251

116. Zelver N. Biofilm development and associated energy losses in water conduits [Internet]. Rice University; 1979 [cited 2019 Feb 21]. Available from: https://scholarship.rice.edu/handle/1911/104227

117. Characklis WG. Bioengineering report: Fouling biofilm development: A process analysis. *Biotechnol Bioeng [Internet]*. 1981 Sep [cited 2018 Jul 18];23(9):1923–60. Available from: http://doi.wiley.com/10.1002/bit.260230902

118. Lewus CB, Kaiser A, Montville TJ. Inhibition of food-borne bacterial pathogens by bacteriocins from lactic acid bacteria isolated from meat. *Appl Environ Microbiol [Internet]*. 1991 Jun 1 [cited 2018 Jul 20];57(6):1683–8. Available from: http://www.ncbi.nlm.nih.gov/pubmed/1908209

119. O'Sullivan L, Ross R., Hill C. Potential of bacteriocin-producing lactic acid bacteria for improvements in food safety and quality. *Biochimie [Internet]*. 2002 May 1 [cited

2018 Jul 20];84(5–6):593–604. Available from: https://www.sciencedirect.com/science/article/pii/S0300908402014578

120. Schnürer J, Magnusson J. Antifungal lactic acid bacteria as biopreservatives. *Trends Food Sci Technol [Internet]*. 2005 Jan 1 [cited 2018 Jul 20];16(1–3):70–8. Available from: https://www.sciencedirect.com/science/article/pii/S0924224404001943

121. Rubin EM. Genomics of cellulosic biofuels. *Nat Rev [Internet]*. 2008 Aug 14 [cited 2019 Feb 27];454(7206):841–5. Available from: http://www.nature.com/articles/nature07190

122. Dominik A, Zverlov V V, Schwarz WH. Biofuels from microbes. *Appl Microbiol Biotechnol [Internet]*. 2007 [cited 2019 Feb 26];77:23–35. Available from: http://www.ybiofuels.org/

123. Azari M, Le A V., Denecke M. Population Dynamic of Microbial Consortia in a Granular Activated Carbon-Assisted Biofilm Reactor: Lessons from Modelling. In: Mannina G, editor. Frontiers in Wastewater Treatment and Modelling, Lecture Notes in Civil Engineering 4. Springer, Cham; 2017 [cited 2018 Apr 6]. p. 588–95. Available from: http://link.springer.com/10.1007/978-3-319-58421-8_92

124. Ardern E, Lockett WT. Experiments on the oxidation of sewage without the aid of filters. *J Soc Chem Ind [Internet]*. 1914 May 30 [cited 2019 Mar 2];33(10):523–39. Available from: http://doi.wiley.com/10.1002/jctb.5000331005

125. Palanisamy N, Ramya J, Kumar S, Vasanthi N, Chandran P, Khan S. Diesel biodegradation capacities of indigenous bacterial species isolated from diesel contaminated soil. *J Environ Heal Sci Eng [Internet]*. 2014 [cited 2018 Jul 24];12(1):142. Available from: http://www.ncbi.nlm.nih.gov/pubmed/25530870

126. Kato Marcus A, Torres CI, Rittmann BE. Conduction-based modeling of the biofilm anode of a microbial fuel cell. *Biotechnol Bioeng [Internet]*. 2007 Dec 15 [cited 2018 Jul 24];98(6):1171–82. Available from: http://doi.wiley.com/10.1002/bit.21533

127. Bixler GD, Bhushan B. Biofouling: lessons from nature. *Philos Trans R Soc A [Internet]*. 2012 May 28 [cited 2019 Feb 21];370(1967):2381–417. Available from: http://rsta.royalsocietypublishing.org/cgi/doi/10.1098/rsta.2011.0502

128. O'hara AM, Shanahan F. The gut flora as a forgotten organ [Internet]. Vol. 7, EMBO reports. 2006 [cited 2019 Feb 21]. Available from: http://embor.embopress.org/content/embor/7/7/688.full.pdf

129. Vincent J-L. Microbial resistance: lessons from the EPIC study. *Intensive Care Med*

[Internet]. 2000 Feb 24 [cited 2019 Feb 21];26(0):S003–8. Available from: http://link.springer.com/10.1007/s001340051111

130. Rainey PB, Rainey K. Evolution of cooperation and conflict in experimental bacterial populations. *Nature [Internet]*. 2003 Sep 4 [cited 2019 Feb 21];425(6953):72–4. Available from: http://www.nature.com/articles/nature01906

131. Hartman G, Wise R. Quorum sensing: potential means of treating gram-negative infections? *Lancet [Internet]*. 1998 Mar 21 [cited 2019 Feb 21];351(9106):848–9. Available from: http://www.ncbi.nlm.nih.gov/pubmed/9525356

132. Høiby N, Bjarnsholt T, Givskov M, Molin S, Ciofu O. Antibiotic resistance of bacterial biofilms. *Int J Antimicrob Agents [Internet]*. 2010 Apr 1 [cited 2018 Jun 13];35(4):322–32. Available from: https://www.sciencedirect.com/science/article/pii/S0924857910000099

133. Dockery J, Keener JP. A Mathematical Model for Quorum Sensing in *Pseudomonas aeruginosa*. *Bull Math Biol [Internet]*. 2001 Jan [cited 2018 Jul 19];63(1):95–116. Available from: http://link.springer.com/10.1006/bulm.2000.0205

134. Anguige K, King JR, Ward JP, Williams P. Mathematical modelling of therapies targeted at bacterial quorum sensing. *Math Biosci [Internet]*. 2004 Nov 1 [cited 2018 Jul 19];192(1):39–83. Available from: https://www.sciencedirect.com/science/article/pii/S0025556404001336

135. Anguige K, King JR, Ward JP. Modelling antibiotic- and anti-quorum sensing treatment of a spatially-structured *Pseudomonas aeruginosa* population. *J Math Biol [Internet]*. 2005 Nov 13 [cited 2018 Jul 19];51(5):557–94. Available from: http://link.springer.com/10.1007/s00285-005-0316-8

136. Romero-Campero FJ, Pérez-Jiménez MJ. A Model of the Quorum Sensing System in *Vibrio fischeri* Using P Systems. *Artif Life [Internet]*. 2008 Jan 2 [cited 2018 Jul 19];14(1):95–109. Available from: http://www.mitpressjournals.org/doi/10.1162/artl.2008.14.1.95

137. Ward J. Mathematical Modeling of Quorum-Sensing Control in Biofilms. In: Balaban N, editor. Control of Biofilm Infections by Signal Manipulation. Berlin: Springer, Berlin, Heidelberg; 2008 [cited 2018 Jul 19]. p. 79–108. Available from: http://link.springer.com/10.1007/7142_2007_010

138. Frederick MR, Kuttler C, Hense BA, Eberl HJ. A mathematical model of quorum sensing regulated EPS production in biofilm communities. *Theor Biol Med Model*

[Internet]. 2011 [cited 2017 Oct 12];8:8. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3090360/pdf/1742-4682-8-8.pdf

139. Emerenini BO, Hense BA, Kuttler C, Eberl HJ. A Mathematical Model of Quorum Sensing Induced Biofilm Detachment. Forestier C, editor. *PLoS One [Internet]*. 2015 Jul 21 [cited 2018 Jul 19];10(7):e0132385. Available from: http://www.ncbi.nlm.nih.gov/pubmed/26197231

140. Né Dicte Martin B, Tamanai-Shacoori Z, Bronsard J, Ginguené F, Meuric V, Mahé F, et al. A new mathematical model of bacterial interactions in two-species oral biofilms. *PLoS One [Internet]*. 2017 [cited 2018 Apr 6];12(3):1–24. Available from: http://europepmc.org/backend/ptpmcrender.fcgi?accid=PMC5333920&blobtype=pdf

141. Wanner O, Gujer W. A multispecies biofilm model. *Biotechnol Bioeng [Internet]*. 1986 Mar [cited 2017 Apr 17];28(3):314–28. Available from: http://doi.wiley.com/10.1002/bit.260280304

142. Xavier JB, Foster KR. Cooperation and conflict in microbial biofilms. *Proc Natl Acad Sci USA [Internet]*. 2007 Jan 16 [cited 2017 Jun 5];104(3):876–81. Available from: http://www.ncbi.nlm.nih.gov/pubmed/17210916

143. Nichols WW, Evans MJ, Slack MPE, Walmsley HL. The Penetration of Antibiotics into Aggregates of Mucoid and Non-mucoid *Pseudomonas aeruginosa*. *Microbiology [Internet]*. 1989 May 1 [cited 2018 Jul 11];135(5):1291–303. Available from: http://mic.microbiologyresearch.org/content/journal/micro/10.1099/00221287-135-5-1291

144. Ehret AE, Böl M. Modelling mechanical characteristics of microbial biofilms by network theory. *J R Soc Interface [Internet]*. 2013 Jan 6 [cited 2018 Nov 14];10(78):20120676. Available from: http://www.ncbi.nlm.nih.gov/pubmed/23034354

145. Ferrier R, Hezard B, Lintz A, Stahl V, Augustin J-C. Combining individual-based modeling and food microenvironment descriptions to predict the growth of *Listeria monocytogenes* on smear soft cheese. *Appl Environ Microbiol [Internet]*. 2013 Oct 1 [cited 2018 May 30];79(19):5870–81. Available from: http://www.ncbi.nlm.nih.gov/pubmed/23872572

146. Wang Z-W, Lee S-H, Elkins JG, Morrell-Falvey JL. Spatial and temporal dynamics of cellulose degradation and biofilm formation by *Caldicellulosiruptor obsidiansis* and *Clostridium thermocellum*. *AMB Express [Internet]*. 2011 Oct 7 [cited 2018 May

1];1(1):30. Available from: http://amb-express.springeropen.com/articles/10.1186/2191-0855-1-30

147. Wang Q, Zhang T. Review of mathematical models for biofilms. *Solid State Commun [Internet]*. 2010 Jun 1 [cited 2018 Mar 29];150(21–22):1009–22. Available from: https://www-sciencedirect-com.dbgw.lis.curtin.edu.au/science/article/pii/S0038109810000281

148. Van Loosdrecht MCM, Heijnen JJ, Eberl H, Kreft J, Picioreanu C. Mathematical modelling of biofilm structures. *Antonie van Leeuwenhoek, Int J Gen Mol Microbiol*. 2002;81(1–4):245–56.

149. Ben-Jacob E, Schochet O, Tenenbaum a, Cohen I, Czirók a, Vicsek T. Generic modelling of cooperative growth patterns in bacterial colonies. *Nature*. 1994;368(6466):46–9.

150. Wanner O, Reichert P. Mathematical modeling of mixed-culture biofilms. *Biotechnol Bioeng*. 1995;49:172–84.

151. Stewart PS, Hamilton MA, Goldstein BR, Schneider BT. Modeling biocide action against biofilms. *Biotechnol Bioeng [Internet]*. 1996 Mar 26 [cited 2018 Nov 27];49(4):445–55. Available from: http://doi.wiley.com/10.1002/%28SICI%291097-0290%2819960220%2949%3A4%3C445%3A%3AAID-BIT12%3E3.0.CO%3B2-9

152. Picioreanu C, van Loosdrecht MC, Heijnen JJ. A theoretical study on the effect of surface roughness on mass transport and transformation in biofilms. *Biotechnol Bioeng [Internet]*. 2000;68(4):355–69. Available from: http://www.ncbi.nlm.nih.gov/pubmed/10745204%5Cnhttp://onlinelibrary.wiley.com/store/10.1002/(SICI)1097-0290(20000520)68:4%3C355::AID-BIT1%3E3.0.CO;2-A/asset/1_ftp.pdf?v=1&t=h164w50x&s=3d9bf2a8743663c4ca288835d25de0f6a612a637

153. Dodds MG, Grobe KJ, Stewart PS. Modeling biofilm antimicrobial resistance. *Biotechnol Bioeng [Internet]*. 2000 May 20 [cited 2018 Nov 27];68(4):456–65. Available from: http://doi.wiley.com/10.1002/%28SICI%291097-0290%2820000520%2968%3A4%3C456%3A%3AAID-BIT11%3E3.0.CO%3B2-Z

154. Chopp DL, Kirisits MJ, Moran B, Parsek MR. A mathematical model of quorum sensing in a growing bacterial biofilm. *J Ind Microbiol Biotechnol [Internet]*. 2002 Dec 1 [cited 2018 Jul 19];29(6):339–46. Available from: http://link.springer.com/10.1038/sj.jim.7000316

155. Chang I, Gilbert ES, Eliashberg N, Keasling JD. A three-dimensional, stochastic simulation of biofilm growth and transport-related factors that affect structure. *Microbiology [Internet]*. 2003 Oct 1 [cited 2018 May 1];149(10):2859–71. Available from: http://mic.microbiologyresearch.org/content/journal/micro/10.1099/mic.0.26211-0

156. Picioreanu C, Kreft J-U, Van Loosdrecht MCM. Particle-based multidimensional multispecies biofilm model. *Appl Environ Microbiol [Internet]*. 2004 May 1 [cited 2017 Oct 2];70(5):3024–40. Available from: http://www.ncbi.nlm.nih.gov/pubmed/15128564

157. Xavier JB, Picioreanu C, Van Loosdrecht MCM. Assessment of three-dimensional biofilm models through direct comparison with confocal microscopy imaging. *Water Sci Technol [Internet]*. 2004 [cited 2018 Apr 6];49(11):177–85. Available from: http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.532.3481&rep=rep1&type=pdf

158. Xavier JB, Picioreanu C, van Loosdrecht MCM. A framework for multidimensional modelling of activity and structure of multispecies biofilms. *Environ Microbiol [Internet]*. 2005 Aug 1 [cited 2018 Apr 12];7(8):1085–103. Available from: http://doi.wiley.com/10.1111/j.1462-2920.2005.00787.x

159. Xavier JB, Picioreanu C, Abdul Rani S, M van Loosdrecht MC, Stewart PS, Joao Xavier JXavier CB. Biofilm-control strategies based on enzymic disruption of the extracellular polymeric substance matrix – a modelling study. *Microbiology [Internet]*. 2005 [cited 2018 Apr 6];151:3817–32. Available from: http://www.microbiologyresearch.org/docserver/fulltext/micro/151/12/3817.pdf?expires=1523004696&id=id&accname=guest&checksum=D163A1E417C785EDECEC0D4697AA8004

160. Hunt SM, Hamilton MA, Stewart PS. A 3D model of antimicrobial action on biofilms. *Water Sci Technol [Internet]*. 2005 Oct 1 [cited 2018 Nov 27];52(7):143–8. Available from: https://iwaponline.com/wst/article/52/7/143/11441/A-3D-model-of-antimicrobial-action-on-biofilms

161. Chanbless JD, Hunt SM, S. SP. A Three-Dimensional Computer Model of Four Hypothetical Mechanisms Protecting Biofilms from Antimicrobials. *Appl Environ Microbiol [Internet]*. 2006 Oct 1 [cited 2018 Nov 27];72(3):2005–13. Available from: http://www.ncbi.nlm.nih.gov/pubmed/9758837

162. Kapellos GE, Alexiou TS, Payatakes AC. Hierarchical simulator of biofilm growth

and dynamics in granular porous materials. *Adv Water Resour [Internet]*. 2007 Jun 1 [cited 2018 Nov 2];30(6–7):1648–67. Available from: https://www.sciencedirect.com/science/article/pii/S0309170806001412

163. Wang Z-W, Hamilton-Brehm SD, Lochner A, Elkins JG, Morrell-Falvey JL. Mathematical modeling of hydrolysate diffusion and utilization in cellulolytic biofilms of the extreme thermophile *Caldicellulosiruptor obsidiansis*. *Bioresour Technol [Internet]*. 2011 Feb 1 [cited 2018 May 1];102(3):3155–62. Available from: https://www.sciencedirect.com/science/article/pii/S0960852410017608?via%3Dihub

164. Lardon LA, Merkey B V., Martins S, Dötsch A, Picioreanu C, Kreft JU, et al. iDynoMiCS: Next-generation individual-based modelling of biofilms. *Environ Microbiol*. 2011;13(9):2416–34.

165. Rodriguez D, Einarsson B, Carpio A. Biofilm growth on rugose surfaces. *Phys Rev E [Internet]*. 2012 [cited 2017 May 11];86(87). Available from: https://journals.aps.org/pre/pdf/10.1103/PhysRevE.86.061914

166. Asally M, Kittisopikul M, Rué P, Du Y, Hu Z, Çağatay T, et al. Localized cell death focuses mechanical forces during 3D patterning in a biofilm. *Proc Natl Acad Sci U S A [Internet]*. 2012 Nov 13 [cited 2017 Nov 7];109(46):18891–6. Available from: http://www.ncbi.nlm.nih.gov/pubmed/23012477

167. Ferrier R, Hezard B, Lintz A, Stahl V, Augustin J-C. Combining individual-based modeling and food microenvironment descriptions to predict the growth of *Listeria monocytogenes* on smear soft cheese. *Appl Environ Microbiol [Internet]*. 2013 Oct 1 [cited 2018 May 30];79(19):5870–81. Available from: http://www.ncbi.nlm.nih.gov/pubmed/23872572

168. Bottero S, Storck T, Heimovaara TJ, van Loosdrecht MCM, Enzien M V., Picioreanu C. Biofilm development and the dynamics of preferential flow paths in porous media. *Biofouling [Internet]*. 2013 Oct [cited 2018 Nov 2];29(9):1069–86. Available from: http://www.tandfonline.com/doi/abs/10.1080/08927014.2013.828284

169. Harcombe WR, Riehl WJ, Dukovski I, Granger BR, Betts A, Lang AH, et al. Metabolic Resource Allocation in Individual Microbes Determines Ecosystem Interactions and Spatial Dynamics. *Cell Rep [Internet]*. 2014 May 22 [cited 2019 Apr 29];7(4):1104–15. Available from: http://www.ncbi.nlm.nih.gov/pubmed/24794435

170. Jayasinghe N, Franks A, Nevin KP, Mahadevan R. Metabolic modeling of spatial heterogeneity of biofilms in microbial fuel cells reveals substrate limitations in

electrical current generation. *Biotechnol J [Internet]*. 2014 Oct 1 [cited 2019 Apr 29];9(10):1350–61. Available from: http://doi.wiley.com/10.1002/biot.201400068

171. Bennett RR, Lee CK, De Anda J, Nealson KH, Yildiz FH, O'Toole GA, et al. Species-dependent hydrodynamics of flagellum-tethered bacteria in early biofilm development. *J R Soc Interface [Internet]*. 2016 Feb 1 [cited 2018 Nov 15];13(115):20150966. Available from: http://www.ncbi.nlm.nih.gov/pubmed/26864892

172. Tack ILMM, Nimmegeers P, Akkermans S, Hashem I, Van Impe JFM. Simulation of *Escherichia coli* Dynamics in Biofilms and Submerged Colonies with an Individual-Based Model Including Metabolic Network Information. *Front Microbiol [Internet]*. 2017 [cited 2018 May 30];8(2509). Available from: http://www.ncbi.nlm.nih.gov/pubmed/29321772

173. Coyte KZ, Tabuteau H, Gaffney EA, Foster KR, Durham WM. Microbial competition in porous environments can select against rapid biofilm growth. *Proc Natl Acad Sci U S A [Internet]*. 2017 Jan 10 [cited 2018 Oct 26];114(2):E161–70. Available from: http://www.ncbi.nlm.nih.gov/pubmed/28007984

174. Stump SM, Johnson EC, Klausmeier CA. Local interactions and self-organized spatial patterns stabilize microbial cross-feeding against cheaters. *J R Soc Interface [Internet]*. 2018 Mar 1 [cited 2018 Nov 15];15(140):20170822. Available from: http://www.ncbi.nlm.nih.gov/pubmed/29563243

175. Rittmann BE, McCarty PL. Model of steady-state-biofilm kinetics. *Biotechnol Bioeng [Internet]*. 1980 Nov 1 [cited 2019 Feb 19];22(11):2343–57. Available from: http://doi.wiley.com/10.1002/bit.260221110

176. Chang I, Gilbert ES, Eliashberg N, Keasling JD. A three-dimensional, stochastic simulation of biofilm growth and transport-related factors that affect structure. *Microbiology*. 2003 Oct;149(10):2859–71.

177. Rodriguez D, Einarsson B, Carpio A. Biofilm growth on rugose surfaces. *Phys Rev E*. 2012;86(87).

178. Wanner O, Reichert P. Mathematical modeling of mixed-culture biofilms. *Biotechnol Bioeng*. 1995;49:172–84.

179. Emerenini BO, Hense BA, Kuttler C, Eberl HJ. A Mathematical Model of Quorum Sensing Induced Biofilm Detachment. *PLoS One*. 2015;10(7):e0132385.

180. Ward J. Mathematical Modeling of Quorum-Sensing Control in Biofilms. In: Control

of Biofilm Infections by Signal Manipulation. Springer, Berlin, Heidelberg; 2008. p. 79–108.

181. Frederick MR, Kuttler C, Hense BA, Eberl HJ. A mathematical model of quorum sensing regulated EPS production in biofilm communities. *Theor Biol Med Model*. 2011;8:8.

182. Asally M, Kittisopikul M, Rue P, Du Y, Hu Z, Cagatay T, et al. Localized cell death focuses mechanical forces during 3D patterning in a biofilm. *Proc Natl Acad Sci [Internet]*. 2012 Nov 13;109(46):18891–6. Available from: http://www.pnas.org/cgi/doi/10.1073/pnas.1212429109

183. Ben-Jacob E, Schochet O, Tenenbaum a, Cohen I, Czirók a, Vicsek T. Generic modelling of cooperative growth patterns in bacterial colonies. *Nature*. 1994;368(6466):46–9.

184. Picioreanu C, Van Loosdrecht MCM, Heijnen JJ. Effect of diffusive and convective substrate transport on biofilm structure formation: A two-dimensional modeling study. *Biotechnol Bioeng [Internet]*. 2000 Sep 5 [cited 2019 Feb 27];69(5):504–15. Available from: http://doi.wiley.com/10.1002/1097-0290%2820000905%2969%3A5%3C504%3A%3AAID-BIT5%3E3.0.CO%3B2-S

185. Battin TJ, Sloan WT, Kjelleberg S, Daims H, Head IM, Curtis TP, et al. Microbial landscapes: new paths to biofilm research. *Nat Rev Microbiol [Internet]*. 2007 Jan 1 [cited 2018 Oct 26];5(1):76–81. Available from: http://www.nature.com/articles/nrmicro1556

186. Bennett RR, Lee CK, De Anda J, Nealson KH, Yildiz FH, O'Toole GA, et al. Species-dependent hydrodynamics of flagellum-tethered bacteria in early biofilm development. *J R Soc Interface [Internet]*. 2016 Feb 1 [cited 2018 Nov 20];13(115):20150966. Available from: http://www.ncbi.nlm.nih.gov/pubmed/26864892

187. Picioreanu C, Kreft J-U, Klausen M, Haagensen JAJ, Tolker-Nielsen T, Molin S. Microbial motility involvement in biofilm structure formation-a 3D modelling study. *Water Sci Technol [Internet]*. 2007 [cited 2019 Mar 6];55(8):337–43. Available from: https://search-proquest-com.dbgw.lis.curtin.edu.au/docview/1943617335?OpenUrlRefId=info:xri/sid:primo&accountid=10382

188. Mabrouk N, Deffuant G, Tolker-Nielsen T, Lobry C. Bacteria can form

interconnected microcolonies when a self-excreted product reduces their surface motility: evidence from individual-based model simulations. *Theory Biosci [Internet]*. 2010 Jun 28 [cited 2019 Mar 6];129(1):1–13. Available from: http://link.springer.com/10.1007/s12064-009-0078-8

189. Flemming H-C, Wingender J. The biofilm matrix. *Nat Rev Microbiol [Internet]*. 2010 Sep 2 [cited 2019 Feb 22];8(9):623–33. Available from: http://www.nature.com/articles/nrmicro2415

190. Flemming H-C, Wingender J, Szewzyk U, Steinberg P, Rice SA, Kjelleberg S. Biofilms: an emergent form of bacterial life. *Nat Rev Microbiol [Internet]*. 2016 [cited 2017 Jun 5];14. Available from: https://www.nature.com/nrmicro/journal/v14/n9/pdf/nrmicro.2016.94.pdf

191. Rani SA, Pitts B, Stewart PS. Rapid diffusion of fluorescent tracers into *Staphylococcus epidermidis* biofilms visualized by time lapse microscopy. *Antimicrob Agents Chemother [Internet]*. 2005 Feb 1 [cited 2018 Apr 12];49(2):728–32. Available from: http://www.ncbi.nlm.nih.gov/pubmed/15673757

192. Waters CM, Bassler BL. QUORUM SENSING: Cell-to-Cell Communication in Bacteria. *Annu Rev Cell Dev Biol [Internet]*. 2005 Nov 7 [cited 2018 Jul 18];21(1):319–46. Available from: http://www.annualreviews.org/doi/10.1146/annurev.cellbio.21.012704.131001

193. Donlan RM. Biofilms: microbial life on surfaces. *Emerg Infect Dis [Internet]*. 2002 Sep [cited 2017 Apr 5];8(9):881–90. Available from: http://www.ncbi.nlm.nih.gov/pubmed/12194761

194. Schuster M, Lostroh CP, Ogi T, Greenberg EP. Identification, timing, and signal specificity of *Pseudomonas aeruginosa* quorum-controlled genes: a transcriptome analysis. *J Bacteriol [Internet]*. 2003 Apr 1 [cited 2018 Jul 19];185(7):2066–79. Available from: http://www.ncbi.nlm.nih.gov/pubmed/12644476

195. Mitri S, Foster KR. The Genotypic View of Social Interactions in Microbial Communities. *Annu Rev Genet [Internet]*. 2013 [cited 2017 Jun 5];47:247–73. Available from: http://www.ncbi.nlm.nih.gov/pubmed/24016192

196. Kirisits MJ, Margolis JJ, Purevdorj-Gage BL, Vaughan B, Chopp DL, Stoodley P, et al. Influence of the hydrodynamic environment on quorum sensing in *Pseudomonas aeruginosa* biofilms. *J Bacteriol [Internet]*. 2007 Nov 15 [cited 2018 Jul 19];189(22):8357–60. Available from:

http://www.ncbi.nlm.nih.gov/pubmed/17704224

197. Anguige K, King JR, Ward JP. Modelling antibiotic- and anti-quorum sensing treatment of a spatially-structured *Pseudomonas aeruginosa* population. *J Math Biol*. 2005 Nov;51(5):557–94.

198. Abdallah M, Benoliel C, Drider D, Dhulster P, Chihib N-E. Biofilm formation and persistence on abiotic surfaces in the context of food and medical environments. *Arch Microbiol [Internet]*. 2014 Jul 18 [cited 2019 Mar 2];196(7):453–72. Available from: http://link.springer.com/10.1007/s00203-014-0983-1

199. Høiby N, Bjarnsholt T, Givskov M, Molin S, Ciofu O. Antibiotic resistance of bacterial biofilms. *Int J Antimicrob Agents [Internet]*. 2010 Apr 1 [cited 2017 Jun 5];35(4):322–32. Available from: http://www.ncbi.nlm.nih.gov/pubmed/20149602

200. Cogan N, Cortez R, Fauci L. Modeling physiological resistance in bacterial biofilms. *Bull Math Biol [Internet]*. 2005 Jul [cited 2018 Nov 27];67(4):831–53. Available from: http://link.springer.com/10.1016/j.bulm.2004.11.001

201. Chanbless JD, Hunt SM, S. SP. A Three-Dimensional Computer Model of Four Hypothetical Mechanisms Protecting Biofilms from Antimicrobials. *Appl Environ Microbiol*. 2006 Oct;72(3):2005–13.

202. Nichols WW, Evans MJ, Slack MPE, Walmsley HL. The Penetration of Antibiotics into Aggregates of Mucoid and Non-mucoid *Pseudomonas aeruginosa*. *Microbiology*. 1989 May;135(5):1291–303.

203. Dodds MG, Grobe KJ, Stewart PS. Modeling biofilm antimicrobial resistance. *Biotechnol Bioeng*. 2000 May;68(4):456–65.

204. Hunt SM, Hamilton MA, Stewart PS. A 3D model of antimicrobial action on biofilms. *Water Sci Technol*. 2005 Oct;52(7):143–8.

205. Stewart PS, Hamilton MA, Goldstein BR, Schneider BT. Modeling biocide action against biofilms. *Biotechnol Bioeng*. 1996 Mar;49(4):445–55.

206. Bridier A, Dubois-Brissonnet F, Greub G, Thomas V, Briandet R. Dynamics of the action of biocides in *Pseudomonas aeruginosa* biofilms. *Antimicrob Agents Chemother [Internet]*. 2011 Jun 1 [cited 2018 May 29];55(6):2648–54. Available from: http://www.ncbi.nlm.nih.gov/pubmed/21422224

207. Bridier A, Dubois-Brissonnet F, Greub G, Thomas V, Briandet R. Dynamics of the action of biocides in *Pseudomonas aeruginosa* biofilms. *Antimicrob Agents*

Chemother [Internet]. 2011 Jun 1 [cited 2018 May 29];55(6):2648–54. Available from: http://www.ncbi.nlm.nih.gov/pubmed/21422224

208. Torres CE, Lenon G, Craperi D, Wilting R, Blanco Á. Enzymatic treatment for preventing biofilm formation in the paper industry. *Appl Microbiol Biotechnol [Internet]*. 2011 [cited 2018 Jul 23];92(1):95–103. Available from: https://link.springer.com/article/10.1007/s00253-011-3305-4

209. Alves D, Olívia Pereira M. Mini-review: Antimicrobial peptides and enzymes as promising candidates to functionalize biomaterial surfaces. *Biofouling [Internet]*. 2014 Apr 21 [cited 2018 May 30];30(4):483–99. Available from: http://www.tandfonline.com/doi/abs/10.1080/08927014.2014.889120

210. Whitman WB, Coleman DC, Wiebe WJ. Prokaryotes: the unseen majority. *Proc Natl Acad Sci U S A [Internet]*. 1998 Jun 9 [cited 2018 Nov 15];95(12):6578–83. Available from: http://www.ncbi.nlm.nih.gov/pubmed/9618454

211. O'Mahony C, Seman DL. Modeling the Microbiological Shelf Life of Foods and Beverages. In: Subramaniam P, editor. The Stability and Shelf Life of Food. Duxford, UK: Elsevier; 2016 [cited 2018 May 30]. p. 253–89. Available from: http://linkinghub.elsevier.com/retrieve/pii/B9780081004357000095

212. Karlsson FH, Nookaew I, Petranovic D, Nielsen J. Prospects for systems biology and modeling of the gut microbiome. *Trends Biotechnol [Internet]*. 2011 Jun [cited 2017 Jun 5];29(6):251–8. Available from: http://www.ncbi.nlm.nih.gov/pubmed/21392838

213. Pérez-Reche FJ, Taraskin SN, Otten W, Viana MP, Costa LDF, Gilligan C a. Prominent Effect of Soil Network Heterogeneity on Microbial Invasion. *Phys Rev Lett [Internet]*. 2012 Aug;109(9):098102. Available from: http://link.aps.org/doi/10.1103/PhysRevLett.109.098102

214. Carrel M, Morales VL, Beltran MA, Derlon N, Kaufmann R, Morgenroth E, et al. Biofilms in 3D porous media: Delineating the influence of the pore network geometry, flow and mass transfer on biofilm development. *Water Res*. 2018 May;134:280–91.

215. Nadell CD, Ricaurte D, Yan J, Drescher K, Bassler BL. Flow environment and matrix structure interact to determine spatial competition in *Pseudomonas aeruginosa* biofilms. *Elife [Internet]*. 2017 Jan 13 [cited 2018 Oct 26];6. Available from: https://elifesciences.org/articles/21855

216. Ammar Y, Swailes D, Bridgens B, Chen J. Influence of surface roughness on the

initial formation of biofilm. *Surf Coatings Technol [Internet]*. 2015 Dec 25 [cited 2019 Feb 25];284:410–6. Available from: https://www.sciencedirect.com/science/article/abs/pii/S0257897215004697

217. Siegismund D, Undisz A, Germerodt S, Schuster S, Rettenmayr M. Quantification of the interaction between biomaterial surfaces and bacteria by 3-D modeling. *Acta Biomater [Internet]*. 2014 Jan 1 [cited 2019 Feb 25];10(1):267–75. Available from: https://www.sciencedirect.com/science/article/pii/S1742706113004698

218. Alnnasouri M, Lemaitre C, Gentric C, Dagot C, Pons M-N. Influence of surface topography on biofilm development: Experiment and modeling. *Biochem Eng J [Internet]*. 2011 Nov 15 [cited 2019 Feb 26];57:38–45. Available from: https://www-sciencedirect-com.dbgw.lis.curtin.edu.au/science/article/pii/S1369703X11002063

219. Meiron TS, Saguy IS. Adhesion Modeling on Rough Low Linear Density Polyethylene. *J Food Sci [Internet]*. 2007 Nov 1 [cited 2019 Feb 26];72(9):E485–91. Available from: http://doi.wiley.com/10.1111/j.1750-3841.2007.00523.x

220. Pedersen K. Biofilm development on stainless steel and pvc surfaces in drinking water. *Water Res*. 1990;24(2):239–43.

221. Geesey GG, Gillis RJ, Avci R, Daly D, Hamilton M, Shope P, et al. The influence of surface features on bacterial colonization and subsequent substratum chemical changes of 316L stainless steel. *Corros Sci*. 1996;38(1):73–95.

222. Sorensen JA. A rationale for comparison of plaque-retaining properties of crown systems. *J Prosthet Dent [Internet]*. 1989 Sep 1 [cited 2018 Nov 27];62(3):264–9. Available from: https://www.sciencedirect.com/science/article/pii/0022391389903296

223. Ikeda M, Matin K, Nikaido T, Foxton RM, Tagami J. Effect of Surface Characteristics on Adherence of *S. mutans* Biofilms to Indirect Resin Composites. *Dent Mater J [Internet]*. 2007 [cited 2018 Nov 27];26(6):915–23. Available from: http://joi.jlc.jst.go.jp/JST.JSTAGE/dmj/26.915?from=CrossRef

224. Park J, Song C, Jung J, Ahn S, Ferracane J. The Effects of Surface Roughness of Composite Resin on Biofilm Formation of *Streptococcus mutans* in the Presence of Saliva. *Oper Dent [Internet]*. 2012 Sep 16 [cited 2018 Nov 27];37(5):532–9. Available from: http://www.jopdentonline.org/doi/10.2341/11-371-L

225. Singh AV, Vyas V, Patil R, Sharma V, Scopelliti PE, Bongiorno G, et al. Quantitative Characterization of the Influence of the Nanoscale Morphology of

Nanostructured Surfaces on Bacterial Adhesion and Biofilm Formation. Bansal V, editor. *PLoS One [Internet]*. 2011 Sep 26 [cited 2018 Nov 27];6(9):e25029. Available from: https://dx.plos.org/10.1371/journal.pone.0025029

226. Efimenko K, Finlay J, Callow ME, Callow JA, Genzer J. Development and Testing of Hierarchically Wrinkled Coatings for Marine Antifouling. *ACS Appl Mater Interfaces [Internet]*. 2009 May 27 [cited 2019 Feb 26];1(5):1031–40. Available from: http://pubs.acs.org/doi/10.1021/am9000562

227. Banerjee I, Pangule RC, Kane RS. Antifouling Coatings: Recent Developments in the Design of Surfaces That Prevent Fouling by Proteins, Bacteria, and Marine Organisms. *Adv Mater [Internet]*. 2011 Feb 8 [cited 2019 Feb 26];23(6):690–718. Available from: http://doi.wiley.com/10.1002/adma.201001215

228. Salta M, Wharton JA, Stoodley P, Dennington SP, Goodes LR, Werwinski S, et al. Designing biomimetic antifouling surfaces. *Philos Trans R Soc A Math Phys Eng Sci [Internet]*. 2010 Oct 28 [cited 2019 Feb 26];368(1929):4729–54. Available from: http://www.royalsocietypublishing.org/doi/10.1098/rsta.2010.0195

229. Xavier JB, Foster KR. Cooperation and conflict in microbial biofilms. *Proc Natl Acad Sci USA*. 2007;104(3):876–81.

230. Picioreanu C, Kreft J-U, Van Loosdrecht MCM. Particle-based multidimensional multispecies biofilm model. *Appl Environ Microbiol*. 2004 May;70(5):3024–40.

231. Simionato MR, Tucker CM, Kuboniwa M, Lamont G, Demuth DR, Tribble GD, et al. *Porphyromonas gingivalis* genes involved in community development with *Streptococcus gordonii*. *Infect Immun [Internet]*. 2006 Nov 1 [cited 2019 Mar 6];74(11):6419–28. Available from: http://www.ncbi.nlm.nih.gov/pubmed/16923784

232. Bridier A, Sanchez-Vizuete P, Guilbaud M, Piard J-C, Naïtali M, Briandet R. Biofilm-associated persistence of food-borne pathogens. *Food Microbiol [Internet]*. 2015 Feb 1 [cited 2018 May 29];45:167–78. Available from: https://www.sciencedirect.com/science/article/pii/S0740002014000902

233. Wang Z-W, Lee S-H, Elkins JG, Morrell-Falvey JL. Spatial and temporal dynamics of cellulose degradation and biofilm formation by *Caldicellulosiruptor obsidiansis* and *Clostridium thermocellum*. *AMB Express [Internet]*. 2011 Oct 7 [cited 2018 May 1];1(1):30. Available from: http://amb-express.springeropen.com/articles/10.1186/2191-0855-1-30

234. Wang Z-W, Hamilton-Brehm SD, Lochner A, Elkins JG, Morrell-Falvey JL.

Mathematical modeling of hydrolysate diffusion and utilization in cellulolytic biofilms of the extreme thermophile *Caldicellulosiruptor obsidiansis*. *Bioresour Technol [Internet]*. 2011 Feb 1 [cited 2018 May 1];102(3):3155–62. Available from: https://www.sciencedirect.com/science/article/pii/S0960852410017608?via%3Dihub

235. Arrhenius S. Über die Reaktionsgeschwindigkeit bei der Inversion von Rohrzucker durch Säuren. *Zeitschrift für Phys Chemie [Internet]*. 1889 Jan 1 [cited 2019 Feb 25];4U(1):226–48. Available from: http://www.degruyter.com/view/j/zpch.1889.4.issue-1/zpch-1889-0416/zpch-1889-0416.xml

236. Wilhelmy L. Ueber das Gesetz, nach welchem die Einwirkung der Säuren auf den Rohrzucker stattfindet. *Ann der Phys und Chemie [Internet]*. 1850 Jan 1 [cited 2019 Feb 25];157(12):499–526. Available from: http://doi.wiley.com/10.1002/andp.18501571203

237. O'Mahony C, Seman DL. Modeling the Microbiological Shelf Life of Foods and Beverages. In: The Stability and Shelf Life of Food. Elsevier; 2016 [cited 2018 May 30]. p. 253–89. Available from: http://linkinghub.elsevier.com/retrieve/pii/B9780081004357000095

238. Dalgaard P. Modelling of microbial activity and prediction of shelf life for packed fresh fish. *Int J Food Microbiol [Internet]*. 1995 Aug 1 [cited 2019 Feb 25];26(3):305–17. Available from: https://www.sciencedirect.com/science/article/pii/016816059400136T

239. Zwietering MH, de Koos JT, Hasenack BE, de Witt JC, van't Riet K. Modeling of bacterial growth as a function of temperature. *Appl Environ Microbiol [Internet]*. 1991 Apr 1 [cited 2019 Feb 25];57(4):1094–101. Available from: http://www.ncbi.nlm.nih.gov/pubmed/2059034

240. Ratkowsky DA, Olley J, McMeekin TA, Ball A. Relationship between temperature and growth rate of bacterial cultures. *J Bacteriol [Internet]*. 1982 Jan 1 [cited 2019 Feb 25];149(1):1–5. Available from: http://www.ncbi.nlm.nih.gov/pubmed/7054139

241. Laidler KJ. The development of the Arrhenius equation. *J Chem Educ [Internet]*. 1984 Jun [cited 2019 Feb 25];61(6):494. Available from: http://pubs.acs.org/doi/abs/10.1021/ed061p494

242. Sheen S, Hwang C-A. Mathematical modeling the cross-contamination of *Escherichia coli* O157:H7 on the surface of ready-to-eat meat product while slicing.

Food Microbiol [Internet]. 2010 Feb 1 [cited 2019 Feb 25];27(1):37–43. Available from: https://www-sciencedirect-com.dbgw.lis.curtin.edu.au/science/article/pii/S0740002009001786

243. Sheen S. Modeling Surface Transfer of *Listeria monocytogenes* on Salami during Slicing. *J Food Sci [Internet]*. 2008 Aug 1 [cited 2019 Feb 25];73(6):E304–11. Available from: http://doi.wiley.com/10.1111/j.1750-3841.2008.00833.x

244. Daigger GT. Ardern and Lockett Remembrance. In: Jenkins D, Wanner J, editors. Activated sludge : 100 years and counting. London, UK: IWA Publishing; 2014. p. 1–16.

245. Gujer W, Henze M, Mino T, Loosdrecht M van. Activated sludge model No. 3. *Water Sci Technol [Internet]*. 1999 Jan 1 [cited 2019 Mar 2];39(1):183–93. Available from: https://www-sciencedirect-com.dbgw.lis.curtin.edu.au/science/article/pii/S0273122398007859

246. Henze M, Grady CPL, Gujer W, Marais G, Matsuo T. Activated Sludge Model No. 1. London; 1986.

247. Hauduc H, Gillot S, Rieger L, Ohtsuki T, Shaw A, Takács I, et al. Activated sludge modelling in practice: an international survey. *Water Sci Technol [Internet]*. 2009 Oct [cited 2018 Apr 30];60(8):1943. Available from: http://wst.iwaponline.com/cgi/doi/10.2166/wst.2009.223

248. Ekama GE, Takács I. Modelling. In: Jenkins D, Wanner J, editors. Activated sludge : 100 years and counting. London, UK: IWA Publishing; 2014. p. 271–93.

249. Wu S-Y, Hung C-H, Lin C-N, Chen H-W, Lee A-S, Chang J-S. Fermentative hydrogen production and bacterial community structure in high-rate anaerobic bioreactors containing silicone-immobilized and self-flocculated sludge. *Biotechnol Bioeng [Internet]*. 2006 Apr 5 [cited 2019 Feb 26];93(5):934–46. Available from: http://doi.wiley.com/10.1002/bit.20800

250. Schenk PM, Thomas-Hall SR, Stephens E, Marx UC, Mussgnug JH, Posten C, et al. Second Generation Biofuels: High-Efficiency Microalgae for Biodiesel Production. *BioEnergy Res [Internet]*. 2008 Mar 4 [cited 2019 Feb 26];1(1):20–43. Available from: http://link.springer.com/10.1007/s12155-008-9008-8

251. Alonso DM, Bond JQ, Dumesic JA. Catalytic conversion of biomass to biofuels. *Green Chem [Internet]*. 2010 Sep 3 [cited 2019 Feb 26];12(9):1493. Available from: http://xlink.rsc.org/?DOI=c004654j

252. Wijffels RH, Barbosa MJ. An outlook on microalgal biofuels. *Science [Internet]*. 2010 Aug 13 [cited 2019 Feb 27];329(5993):796–9. Available from: http://www.ncbi.nlm.nih.gov/pubmed/20705853

253. de Jong B, Siewers V. Systems biology of yeast: enabling technology for development of cell factories for production of advanced biofuels. *Curr Opin Biotechnol [Internet]*. 2012 Aug 1 [cited 2019 Feb 27];23(4):624–30. Available from: https://www-sciencedirect-com.dbgw.lis.curtin.edu.au/science/article/pii/S0958166911007324

254. Other Organisms [Internet]. Systems Biology Research Group. 2017 [cited 2019 Apr 29]. Available from: http://systemsbiology.ucsd.edu/InSilicoOrganisms/OtherOrganisms

255. Raes J, Bork P. Molecular eco-systems biology: towards an understanding of community function. *Nat Rev Microbiol [Internet]*. 2008 [cited 2019 Feb 22];6(9):693–9. Available from: www.nature.com/reviews/micro

256. Curran KA, Alper HS. Expanding the chemical palate of cells by combining systems biology and metabolic engineering. *Metab Eng [Internet]*. 2012 Jul 1 [cited 2019 Feb 27];14(4):289–97. Available from: https://www-sciencedirect-com.dbgw.lis.curtin.edu.au/science/article/pii/S1096717612000493

257. Kitano H. Systems biology: a brief overview. *Science (80-) [Internet]*. 2002 Mar 1 [cited 2019 Feb 22];295(5560):1662–4. Available from: http://www.ncbi.nlm.nih.gov/pubmed/11872829

258. Lewenza S. Extracellular DNA-induced antimicrobial peptide resistance mechanisms in *Pseudomonas aeruginosa*. *Front Microbiol [Internet]*. 2013 Feb 14 [cited 2019 Jul 9];4:21. Available from: http://journal.frontiersin.org/article/10.3389/fmicb.2013.00021/abstract

259. Wilkinson DJ. Stochastic modelling for quantitative description of heterogeneous biological systems. *Nat Rev Genet*. 2009;10:122–33.

260. O'Brien SJ. The consequences of *Campylobacter* infection. *Curr Opin Gastroenterol [Internet]*. 2017 Jan [cited 2019 Feb 5];33(1):14–20. Available from: http://insights.ovid.com/crossref?an=00001574-201701000-00004

261. Sulaeman S, Hernould M, Schaumann A, Coquet L, Bolla J-M, Dé E, et al. Enhanced Adhesion of *Campylobacter jejuni* to Abiotic Surfaces Is Mediated by Membrane Proteins in Oxygen-Enriched Conditions. Heimesaat MM, editor. *PLoS One*

[Internet]. 2012 Sep 28 [cited 2017 Sep 22];7(9):e46402. Available from: http://dx.plos.org/10.1371/journal.pone.0046402

262. Bronowski C, James CE, Winstanley C. Role of environmental survival in transmission of *Campylobacter jejuni*. *FEMS Microbiol Lett*. 2014;356(1):8–19.

263. Teh AHT, Lee SM, Dykes GA. Association of some *Campylobacter jejuni* with *Pseudomonas aeruginosa* biofilms increases attachment under conditions mimicking those in the environment. Wells JE, editor. *PLoS One [Internet]*. 2019 Apr 10 [cited 2020 Jan 11];14(4). Available from: http://dx.plos.org/10.1371/journal.pone.0215275

264. Brown HL, Reuter M, Hanman K, Betts RP, Van Vliet AHM. Prevention of Biofilm Formation and Removal of Existing Biofilms by Extracellular DNases of *Campylobacter jejuni*. 2015;

265. Pascoe B, Méric G, Murray S, Yahara K, Mageiros L, Bowen R, et al. Enhanced biofilm formation and multi-host transmission evolve from divergent genetic backgrounds in *Campylobacter jejuni*. *Environ Microbiol [Internet]*. 2015 Nov 1 [cited 2020 Feb 25];17(11):4779–89. Available from: http://doi.wiley.com/10.1111/1462-2920.13051

266. Hermanowicz SW. A simple 2D biofilm model yields a variety of morphological features. *Math Biosci [Internet]*. 2001 Jan [cited 2017 May 4];169(1):1–14. Available from: http://www.sciencedirect.com.dbgw.lis.curtin.edu.au/science/article/pii/S0025556400000493

267. Wimpenny JWT, Colasanti R. A unifying hypothesis for the structure of microbial biofilms based on cellular automaton models. *FEMS Microbiol Ecol [Internet]*. 1997 [cited 2017 May 2];22(1):1–16. Available from: http://www.sciencedirect.com/science/article/pii/S0168649696000785

268. Picioreanu C, Van Loosdrecht MCM, Heijnen JJ. Mathematical Modeling of Biofilm Structure with a Hybrid Differential-Discrete Cellular Automaton Approach. *Biotechnol Bioeng [Internet]*. 1998 [cited 2017 Apr 17];58(1):101–16. Available from: https://pdfs.semanticscholar.org/1410/36be7931ae67f7def7d90865c52c29c6b3d9.pdf

269. Xavier JB, Picioreanu C, van Loosdrecht MCM. A modelling study of the activity and structure of biofilms in biological reactors. *Biofilms [Internet]*. 2004 Oct [cited 2018 Apr 6];1(4):377–91. Available from:

http://www.journals.cambridge.org/abstract_S1479050505001560

270. Pérez-Reche FJ, Ludlam JJ, Taraskin SN, Gilligan CA. Synergy in spreading processes: From exploitative to explorative foraging strategies. *Phys Rev Lett [Internet]*. 2011 May 24 [cited 2020 Jul 15];106(21):218701. Available from: https://journals.aps.org/prl/abstract/10.1103/PhysRevLett.106.218701

271. Gilligan CA, van den Bosch F. Epidemiological Models for Invasion and Persistence of Pathogens. *Annu Rev Phytopathol [Internet]*. 2008 Sep 4 [cited 2020 Jul 15];46(1):385–418. Available from: www.annualreviews.org

272. Jansen APJ. Kinetic Monte Carlo Algorithms. In: Beiglböck W, Ehlers J, Hepp K, Weidenmüller H, editors. An Introduction to Kinetic Monte Carlo Simulations of Surface Reactions. Berlin, Heidelberg: Springer-Verlag Berlin Heidelberg; 2012. p. 37–71. Available from: https://www.springer.com/gp/book/9783642294877

273. Yan J, Sharo AG, Stone HA, Wingreen NS, Bassler BL. *Vibrio cholerae* biofilm growth program and architecture revealed by single-cell live imaging. *Proc Natl Acad Sci [Internet]*. 2016;1–7. Available from: http://www.pnas.org/content/early/2016/08/22/1611494113.abstract

274. Eberl HJ, Parker DF, Vanloosdrecht MCM. A new deterministic spatio-temporal continuum model for biofilm development. *J Theor Med*. 2001;3(3):161–75.

275. Sen PN. Time-dependent diffusion coefficient as a probe of geometry. *Concepts Magn Reson Part A Bridg Educ Res [Internet]*. 2004 Sep 1 [cited 2020 Dec 4];23(1):1–21. Available from: https://onlinelibrary.wiley.com/doi/full/10.1002/cmr.a.20017

276. Mohammed KAS, Miles RJ, Halablab MA. The pattern and kinetics of substrate metabolism of *Campylobacter jejuni* and *Campylobacter coli*. *Lett Appl Microbiol [Internet]*. 2004 Sep 1 [cited 2019 Jun 12];39(3):261–6. Available from: http://doi.wiley.com/10.1111/j.1472-765X.2004.01574.x

277. Hazeleger WC, Wouters JA, Rombouts FM, Abee T. Physiological Activity of *Campylobacter jejuni* Far below the Minimal Growth Temperature. *Appl Environ Microbiol [Internet]*. 1998 Oct 1 [cited 2019 Mar 14];64(10):3917–22. Available from: https://aem.asm.org/content/64/10/3917.short

278. Schulze KL, Lipe RS. Relationship between substrate concentration, growth rate, and respiration rate of *Escherichia coli* in continuous culture. *Arch Mikrobiol [Internet]*. 1964 [cited 2019 Jun 12];48(1):1–20. Available from:

http://link.springer.com/10.1007/BF00406595

279. Rollins DM, Colwell RR. Viable but Nonculturable Stage of *Campylobacter jejuni* and Its Role in Survival in the Natural Aquatic Environment. *Appl Environ Microbiol*. 1986;52(3):531–8.

280. Pérez-Reche FJ, Taraskin SN, Costa LDF, Neri FM, Gilligan CA. Complexity and anisotropy in host morphology make populations less susceptible to epidemic outbreaks. *J R Soc Interface*. 2010 Jul 6;7(48):1083–92.

281. Grassberger P. On the critical behavior of the general epidemic process and dynamical percolation. *Math Biosci*. 1983 Apr 1;63(2):157–72.

282. Duffy LL, Dykes GA. The Ability of *Campylobacter jejuni* Cells to Attach to Stainless Steel Does Not Change as They Become Nonculturable. *Foodborne Pathog Dis [Internet]*. 2009 [cited 2017 Aug 31];6(5). Available from: http://online.liebertpub.com/doi/pdf/10.1089/fpd.2008.0250

283. Valdés A, Ramos M, Beltrán A, Jiménez A, Garrigós M. State of the Art of Antimicrobial Edible Coatings for Food Packaging Applications. *Coatings [Internet]*. 2017 Apr 19 [cited 2020 Jul 16];7(4):56. Available from: http://www.mdpi.com/2079-6412/7/4/56

284. Mehlman IJ, Romero A. Improved growth medium for *Campylobacter* species. *Appl Environ Microbiol*. 1982;43(3).

285. Stahl M, Butcher J, Stintzi A. Nutrient acquisition and metabolism by *Campylobacter jejuni*. *Front Cell Infect Microbiol [Internet]*. 2012 [cited 2019 May 22];2:5. Available from: http://www.ncbi.nlm.nih.gov/pubmed/22919597

286. Warepam M, Singh LR. Osmolyte mixtures have different effects than individual osmolytes on protein folding and functional activity. *Arch Biochem Biophys*. 2015 May 1;573:77–83.

287. Sleator RD, Hill C. Bacterial osmoadaptation: the role of osmolytes in bacterial stress and virulence. *FEMS Microbiol Rev [Internet]*. 2002 Mar 1 [cited 2020 Jul 16];26(1):49–71. Available from: www.fems-microbiology.org

288. Zhang T, Dong J, Cheng Y, Lu Q, Luo Q, Wen G, et al. Genotypic diversity, antimicrobial resistance and biofilm-forming abilities of *Campylobacter* isolated from chicken in Central China. *Gut Pathog [Internet]*. 2017 Nov 9 [cited 2020 Feb 25];9(1):62. Available from: https://gutpathogens.biomedcentral.com/articles/10.1186/s13099-017-0209-6

289. Grove J, Ripke S, Als TD, Mattheisen M, Walters RK, Won H, et al. Identification of common genetic risk variants for autism spectrum disorder. *Nat Genet*. 2019 Mar 1;51(3):431–44.

290. Laabei M, Recker M, Rudkin JK, Aldeljawi M, Gulay Z, Sloan TJ, et al. Predicting the virulence of MRSA from its genome sequence. *Genome Res*. 2014 May 1;24(5):839–49.

291. Falush D. Bacterial genomics: Microbial GWAS coming of age. Vol. 1, Nature Microbiology. Nature Publishing Group; 2016. p. 1–2.

292. Saber MM, Jesse Shapiro B. Benchmarking bacterial genome-wide association study methods using simulated genomes and phenotypes. *Microb Genomics*. 2020;6(3).

293. Fritsch L, Felten A, Palma F, Mariet JF, Radomski N, Mistou MY, et al. Insights from genome-wide approaches to identify variants associated to phenotypes at pan-genome scale: Application to *L. monocytogenes*' ability to grow in cold conditions. *Int J Food Microbiol*. 2019 Feb 16;291:181–8.

294. Parks DH, Imelfort M, Skennerton CT, Hugenholtz P, Tyson GW. CheckM: Assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. *Genome Res*. 2015 Jul 1;25(7):1043–55.

295. Page AJ, Cummins CA, Hunt M, Wong VK, Reuter S, Holden MTG, et al. Roary: rapid large-scale prokaryote pan genome analysis. *Bioinformatics*. 2015;31(22):3691–3.

296. Kozlov AM, Darriba D, Flouri T, Morel B, Stamatakis A. RAxML-NG: a fast, scalable and user-friendly tool for maximum likelihood phylogenetic inference. *Bioinformatics [Internet]*. 2019 [cited 2020 May 21];35(21):4453–5. Available from: https://academic.oup.com/bioinformatics/article/35/21/4453/5487384

297. Brynildsrud O, Bohlin J, Scheffer L, Eldholm V. Rapid scoring of genes in microbial pan-genome-wide association studies with Scoary. *Genome Biol [Internet]*. 2016 Nov 25 [cited 2020 Mar 20];17(1):238. Available from: http://genomebiology.biomedcentral.com/articles/10.1186/s13059-016-1108-8

298. Lees JA, Galardini M, Bentley SD, Weiser JN, Corander J. pyseer: a comprehensive tool for microbial pangenome-wide association studies. *Bioinformatics [Internet]*. 2018 [cited 2020 Apr 8];34(24):4310–2. Available from: https://academic.oup.com/bioinformatics/article/34/24/4310/5047751

299. Saber MM, Shapiro J. Benchmarking bacterial genome-wide association study

(GWAS) methods using simulated genomes and phenotypes. *bioRxiv*. 2019 Oct 7;795492.

300. Treangen TJ, Ondov BD, Koren S, Phillippy AM. The harvest suite for rapid core-genome alignment and visualization of thousands of intraspecific microbial genomes. *Genome Biol [Internet]*. 2014 Nov 19 [cited 2020 May 29];15(11):524. Available from: http://genomebiology.biomedcentral.com/articles/10.1186/s13059-014-0524-x

301. Reumerman RA, Tucker NP, Herron PR, Hoskisson PA, Sangal V. Tool for rapid annotation of microbial SNPs (TRAMS): A simple program for rapid annotation of genomic variation in prokaryotes. *Antonie van Leeuwenhoek, Int J Gen Mol Microbiol [Internet]*. 2013 Sep [cited 2020 Jul 20];104(3):431–4. Available from: https://pubmed.ncbi.nlm.nih.gov/23828175/

302. Cody AJ, Bray JE, Jolley KA, McCarthy ND, Maidena MCJ. Core genome multilocus sequence typing scheme for stable, comparative analyses of *Campylobacter jejuni* and *C. coli* human disease isolates. *J Clin Microbiol [Internet]*. 2017 Jul 1 [cited 2020 Feb 25];55(7):2086–97. Available from: https://doi.org/10

303. Murkin AS, Chou WK, Wakarchuk WW, Tanner ME. Identification and mechanism of a bacterial hydrolyzing UDP-N-acetylglucosamine 2-epimerase. *Biochemistry [Internet]*. 2004 Nov 9 [cited 2021 Jan 30];43(44):14290–8. Available from: https://pubs.acs.org/sharingguidelines

304. Guerry P, Ewing CP, Hickey TE, Prendergast MM, Moran AP. Sialylation of lipooligosaccharide cores affects immunogenicity and serum resistance of *Campylobacter jejuni*. *Infect Immun [Internet]*. 2000 [cited 2021 Jan 30];68(12):6656–62. Available from: /pmc/articles/PMC97763/?report=abstract

305. Zebian N, Merkx-Jacques A, Pittock PP, Houle S, Dozois CM, Lajoie GA, et al. Comprehensive analysis of flagellin glycosylation in *Campylobacter jejuni* NCTC 11168 reveals incorporation of legionaminic acid and its importance for host colonization. *Glycobiology [Internet]*. 2016 Apr 1 [cited 2021 Jan 30];26(4):386–97. Available from: https://academic.oup.com/glycob/article-lookup/doi/10.1093/glycob/cwv104

306. Greiner LL, Watanabe H, Phillips NJ, Shao J, Morgan A, Zaleski A, et al. Nontypeable *Haemophilus influenzae* strain 2019 produces a biofilm containing N-acetylneuraminic acid that may mimic sialylated O-linked glycans. *Infect Immun*. 2004 Jul;72(7):4249–60.

307. Sharma Y, Miladi M, Dukare S, Boulay K, Caudron-Herger M, Groß M, et al. A pan-cancer analysis of synonymous mutations. *Nat Commun [Internet]*. 2019 Dec 1 [cited 2020 Jul 21];10(1):1–14. Available from: https://doi.org/10.1038/s41467-019-10489-2

308. Abd El-Tawab AA, Ammar AM, Ahmed HA, Hefny AA. Efflux Pump Inhibitors, Alpha-Tocopherol and Aspirin: Role in *Campylobacter jejuni* and *Campylobacter coli* Fluoroquinolone Resistance. *Microb Drug Resist [Internet]*. 2019 Mar 1 [cited 2021 Jan 30];25(2):203–11. Available from: https://www.liebertpub.com/doi/10.1089/mdr.2018.0086

309. Guo W, Cai LL, Zou HS, Ma WX, Liu XL, Zou LF, et al. Ketoglutarate transport protein KgtP is secreted through the type III secretion system and contributes to virulence in *Xanthomonas oryzae pv. oryzae*. *Appl Environ Microbiol [Internet]*. 2012 Aug 15 [cited 2020 Jul 21];78(16):5672–81. Available from: http://www.ncbi

310. Freeman ZN, Dorus S, Waterfield NR. The KdpD/KdpE Two-Component System: Integrating K+ Homeostasis and Virulence. Chitnis CE, editor. *PLoS Pathog [Internet]*. 2013 Mar 28 [cited 2020 Jul 21];9(3):e1003201. Available from: https://dx.plos.org/10.1371/journal.ppat.1003201

311. Hoo R, Lam JH, Huot L, Pant A, Li R, Hot D, et al. Evidence for a Role of the Polysaccharide Capsule Transport Proteins in Pertussis Pathogenesis. Hozbor DF, editor. *PLoS One [Internet]*. 2014 Dec 12 [cited 2020 Jul 23];9(12):e115243. Available from: https://dx.plos.org/10.1371/journal.pone.0115243

312. Bachtiar BM, Coloe PJ, Fry BN. Knockout mutagenesis of the *kpsE* gene of *Campylobacter jejuni* 81116 and its involvement in bacterium–host interactions. *FEMS Immunol Med Microbiol [Internet]*. 2007 Feb 1 [cited 2020 Jul 23];49(1):149–54. Available from: https://academic.oup.com/femspd/article-lookup/doi/10.1111/j.1574-695X.2006.00182.x

313. Nguyen VT, Barlow RS, Fegan N, Turner MS, Dykes GA. Role of capsular polysaccharides and lipooligosaccharides in *Campylobacter* surface properties, autoagglutination, and attachment to abiotic surfaces. *Foodborne Pathog Dis [Internet]*. 2013 Jun 1 [cited 2020 Jul 23];10(6):506–13. Available from: http://www.liebertpub.com/doi/10.1089/fpd.2012.1365

314. Keo T, Collins J, Kunwar P, Blaser MJ, Iovine NM. *Campylobacter* capsule and lipooligosaccharide confer resistance to serum and cationic antimicrobials. *Virulence [Internet]*. 2011 Jan 27 [cited 2020 Jul 23];2(1):30–40. Available from:

http://www.tandfonline.com/doi/abs/10.4161/viru.2.1.14752

315. Zeng S, Constant P, Yang D, Baulard A, Lefèvre P, Daffé M, et al. Cpn60.1 (GroEL1) Contributes to Mycobacterial Crabtree Effect: Implications for Biofilm Formation. *Front Microbiol [Internet]*. 2019 Jun 11 [cited 2020 Jul 20];10(JUN):1149. Available from: https://www.frontiersin.org/article/10.3389/fmicb.2019.01149/full

316. Hassanov T, Karunker I, Steinberg N, Erez A, Kolodkin-Gal I. Novel antibiofilm chemotherapies target nitrogen from glutamate and glutamine. *Sci Rep [Internet]*. 2018 Dec 1 [cited 2020 Jul 20];8(1):1–12. Available from: www.nature.com/scientificreports

317. Sampathkumar B, Napper S, Carillo CD, Willson P, Taboada E, Nash JHE, et al. Transcriptional and translational expression patterns associated with immobilized growth of *Campylobacter jejuni*. *Microbiology [Internet]*. 2006 Feb 1 [cited 2020 Jul 23];152(2):567–77. Available from: https://www.microbiologyresearch.org/content/journal/micro/10.1099/mic.0.28405-0

318. Oh E, Jeon B. Role of alkyl hydroperoxide reductase (AhpC) in the biofilm formation of *Campylobacter jejuni*. *PLoS One*. 2014 Jan 31;9(1).

319. Waters RC, O'Toole PW, Ryan KA. The FliK protein and flagellar hook-length control. *Protein Sci [Internet]*. 2007 May 1 [cited 2020 Jul 23];16(5):769–80. Available from: http://doi.wiley.com/10.1110/ps.072785407

320. Mao M-Y, Yang Y-M, Li K-Z, Lei L, Li M, Yang Y, et al. The rnc Gene Promotes Exopolysaccharide Synthesis and Represses the vicRKX Gene Expressions via MicroRNA-Size Small RNAs in *Streptococcus mutans*. *Front Microbiol [Internet]*. 2016 May 10 [cited 2020 Jul 24];7(May):687. Available from: http://journal.frontiersin.org/Article/10.3389/fmicb.2016.00687/abstract

321. Saund K, Lapp Z, Thiede SN, Pirani A, Snitkin ES. Prewas: Data pre-processing for more informative bacterial gwas. *Microb Genomics [Internet]*. 2020 [cited 2021 Jan 30];6(5):1–8. Available from: /pmc/articles/PMC7371116/?report=abstract

322. Ma KC, Mortimer TD, Duckett MA, Hicks AL, Wheeler NE, Sánchez-Busó L, et al. Increased power from conditional bacterial genome-wide association identifies macrolide resistance mutations in *Neisseria gonorrhoeae*. *Nat Commun [Internet]*. 2020 Dec 1 [cited 2021 Jan 30];11(1):1–8. Available from: https://doi.org/10.1038/s41467-020-19250-6

Every reasonable effort has been made to acknowledge the owners of copyright material. I would be pleased to hear from any copyright owner who has been omitted or incorrectly acknowledged.