Ecology and population genetic structure of the genus *Hafnia* (Enterobacteriaceae)

Shoko Okada

September 2004

A thesis submitted for the degree of Doctor of Philosophy of the Australian National University

Declaration

Results presented in this thesis are my original work except where reference has been provided. Chapters 2, 3, 4, 5, and 6 had co-authors but, in all cases, I am the primary contributor to the work. None of the work in this thesis has been submitted for the award of any other degree.

Shoko Okada

September 2004

Acknowledgments

My first and biggest gratitude goes to my supervisor David Gordon. Starting out as 'Dr Gordon' six years ago when I first stepped onto Australian soil, you some how saw a spec of 'research skin' (as we Japanese like to call it) in me and kindly introduced me into the world of microbial population biology. It has been a wonderful and memorable five years, being inspired by your enthusiasm towards poo-ology, and your being fully supportive financially as well as mentally. I appreciate your unlimited patience and understanding towards whatever I tried (or not tried). The last months of write up had one of the steepest learning curves I've ever experienced, without your knowledge and expertise I wouldn't have understood most of what I've gone through. Thank you David.

I am indebted to Chris Hayes, who thoroughly taught me the sequencing procedure. I would also like to thank the people in the workshop; in particular Alan, who marvellously maintained the dish washer and autoclave for us heavy users in the microbiology lab, Garry and Karen, who delivered the goodies, Jim and Bob with their computational support, and Ginny with the maintenance of lab equipment and pilates classes.

A special thanks goes to Dr John Trueman, who went through my trees and taught me some aspects of cladistics.

Friends in Botany and Zoology made my PhD life quite interesting. Dr Jess, Dave, Anke Maria, Dirk, Dr DD Donuts, you all gave me the opportunity to step out into the wilderness and see some funky skinks, birds, shrubs and plants. Bob and Louise kept on reminding me how nice Japanese food was. I had one attempt with Stu to isolate bacteria from termites, didn't quite succeed but was fun nonetheless. Andy, a big thanks and hug to you for always being there when I needed you.

I must thank Jess in particular for introducing me to the next round of people I am about to mention. The frisbee scene in Canberra has given me an oasis outside my research time. Helen, Leon and Thea, thanks for your kindness and providing laughter. Sue, your encouragement throughout the last months of writing kept me going hard. Minh, thanks for feeding me with tasty food and music and going through knee rehab with me. Thanks to Lorna for the ppc coffee breaks and sprints. Joy, we should really resurrect the Twin Peaks sessions. Kelly, I am ready to borrow those books now. Team ICE, thanks for letting me be part of a great bunch of girls. Max, your 3/4 strength coffee has been (and will always be) a great way to start the Saturday mornings. Pete, I do inspire towards your dedication to ANUDE, I really do. The Pirates Pleasure team was good fun at EUGs and AUGs in 2003, thanks to all of you team members for an everlasting memory. I should mention

Leanne for letting us stay over at her place and giving me a bit extra on the field during AUGs 2003. Thanks to the Canberra frisbee society and ACTUA for keeping me stay fit and healthy.

Second to last round of thanks goes to my close ones. Marco, I really appreciate your support and understanding towards my passion, thankyou. Happy birthday Dad. Mum, thanks for always spending lots of time with me over the phone.

Finally I would like to thank all of the researchers around Australia for your kind donations of faecal samples, many of which enabled my research to happen.

This PhD was supported by the International Postgraduate Research Scholarship and the ANU PhD scholarship. Research was funded by the Australian Research Council through David.

"The definition widely adopted in recent decades-'Evolution is the change of gene frequencies in populations' - refers only to the transformational component. It tells us nothing about the multiplication of species nor, more broadly, about the origin of organic diversity. A broader definition is needed which would include both transformation and diversification."

Mayr E.
The Growth of Biological Thought: Diversity, Evolution, and Inheritance

Abstract

It is well known that many bacterial species consist of multiple sequence clusters that are ecologically distinct. The existence of these ecologically distinct populations raises the question concerning the species definition of bacteria. How does one define a species boundary, and how does a species maintain their genetic integrity? Among the many eukaryotic species concepts that have been applied to prokaryotic speciation in attempts to answer these questions, a combination of the ecological species concept and biological species concept has succeeded in establishing a foundation that explains the evolutionary processes and consequences of bacterial speciation. This thesis has attempted to elucidate the ecological and genetic structure of an enteric bacteria, *Hafnia alvei*, based on these species concepts.

The aim of the initial project was to investigate the overall genetic and phenotypic structure of the species. Multilocus enzyme electrophoresis was used to characterise 161 strains that were biochemically identified as *H. alvei*. Phylogenetic analysis of the allozyme variation revealed that this species consists of two electrophoretically distinct clusters. Two biochemical traits – utilisation of malonate as a carbon source, and acid production during fermentation of glucose – were distinguishing features of the two *Hafnia* groups. Host distribution of the two *Hafnia* groups was also considerably different. Whilst both groups were equally frequent in reptiles, group 2 contained isolates from fish, and mammals of any size, whereas the group 1 were mostly isolated from small mammals. Ten to 20 % of the electrophoretic and biochemical variation within each *Hafnia* group was explained by the taxonomic class of the host from which the strains were isolated.

The next two projects investigated the difference in the phenotypic variation between the two *Hafnia* groups and the extent to which the host influenced the observed variation. The first trait that was examined was bacteriocin production and resistance. Approximately 20 % of the strains produced bacteriocinogenic agents, the majority of which were *Hafnia* group 2 strains. *Hafnia* group 2 isolates from reptiles produced bacteriocins with the strongest and broadest killing pattern. Overall there were at least 14 different bacteriocin phenotypes detected based on the killing patterns of the bacteriocins. Resistance against the bacteriocins was high, with an average of 61 % of the strains being resistant to a particular bacteriocin. The second trait that was examined was the variation in the growth rate – temperature relationship among strains. There was a significant difference in the growth rate parameters of the two *Hafnia* groups. Variation in the growth rate parameters was significantly influenced by host taxonomic class in group 2 strains, but not in those of group 1. After accounting for genetic group and host class, there still was residual variation in the parameters that were partially explained by several biochemical characteristics of the strains.

The final project was to identify the extent of sharing of genetic information within and between the two *Hafnia* groups. Multilocus sequence typing of partial regions of six housekeeping genes confirmed lack of allele sharing between the two *Hafnia* groups. This evidence is interpreted as the two groups representing two separate species. The two biochemical traits, when mapped onto the maximum likelihood tree, indicated that *Hafnia* group 2 evolved from group 1.

There are three major conclusions based on these results; that the two *Hafnia* groups clearly represent two species within the genus *Hafnia*, that *Hafnia* group 2 evolved from group 1, and that the genetic and ecological structure of the two *Hafnia* groups are shaped

differently, where *Hafnia* group 2 is under stronger selection pressure by the host in comparison to group 1..

CHAPTER 1 .... **1**

GENERAL INTRODUCTION **1**

THE EXISTENCE OF ECOLOGICAL NICHES IN BACTERIA 1
THE HISTORY OF BACTERIAL SPECIES CLASSIFICATION 1
INCORPORATION OF EUKARYOTIC SPECIATION THEORY TO BACTERIAL SPECIATION 3
THE CASE OF GENOTYPE DIFFERENTIATION IN *HAFNIA* 5
REFERENCES 7

CHAPTER 2 .... **11**

GENETIC AND ECOLOGICAL STRUCTURE OF *HAFNIA ALVEI* IN AUSTRALIA **11**

INTRODUCTION 11
MATERIAL AND METHODS 12
Isolation of bacteria *12*
Multilocus enzyme electrophoresis *13*
Statistical analysis *18*
RESULTS 19
Genetic structure of Hafnia alvei *19*
Biochemical profile of H. alvei *23*
Ecological factors *25*
DISCUSSION 30
REFERENCES 32

CHAPTER 3 .... **38**

THE FREQUENCY AND DIVERSITY OF BACTERIOCINS IN *HAFNIA ALVEI* **38**

INTRODUCTION 38
MATERIAL AND METHODS 39
Strains *39*
Detection of bacteriocin production *40*
Bacteriocin typing and resistance profile *41*
RESULTS 41
Bacteriocin production in Hafnia alvei *41*
Bacteriocin typing in H. alvei *42*
Resistance profile of H. alvei to bacteriocin *45*
DISCUSSION 46
REFERENCES 48

CHAPTER 4 .... **51**

ADAPTIVE NATURE OF THE GROWTH RATE VARIATION IN THE GENUS *HAFNIA* **51**

INTRODUCTION 51
MATERIAL AND METHODS 53
Bacterial strains *53*
Growth characteristics *53*

Calculation of the growth rate parameters ... 54
RESULTS ... 55
Effect of genetic group ... 55
Effect of host taxonomy ... 55
Effect of biochemical characteristics ... 59
DISCUSSION ... 61
Adaptation of H. alvei group 2 strains to its host but not in group 1 ... 61
Correlation of some growth parameters to biochemical characteristics ... 62
Future directions ... 63
REFERENCES ... 64

CHAPTER 5 ... 67

GENETIC AND PHENOTYPIC CHARACTERISATION OF THE GENUS *HAFNIA* AND PROPOSAL OF *HAFNIA TEMPUS* SP. NOV. ... 67

INTRODUCTION ... 67
MATERIAL AND METHODS ... 69
Strain selection ... 69
Gene selection ... 69
Sequencing procedure ... 69
Phylogenetic analysis ... 70
RESULTS ... 74
Genetic differentiation between the two genetic groups of Hafnia ... 74
Biochemical differentiation between the two genetic groups of Hafnia ... 81
DISCUSSION ... 83
Application of the biological species concept to bacterial speciation ... 83
Subspecific versus species structure in bacterial populations ... 84
The extent of genetic differentiation between the two Hafnia groups ... 85
The ancestor group of Hafnia may be genetic group 1 ... 85
Description of Hafnia tempus sp. nov. ... 90
REFERENCES ... 90

CHAPTER 6 ... 97

COMPARATIVE ANALYSES ON THE POPULATION GENETICS OF THE TWO *HAFNIA* GROUPS ... 97

INTRODUCTION ... 97
MATERIAL AND METHODS ... 99
DNA sequences ... 99
Phylogenetic analyses ... 99
RESULTS ... 103
Allelic variation in Hafnia group 1 ... 103
Allelic variation in Hafnia group 2 ... 104
Intragenic recombination estimates of the six genes in the two Hafnia groups ... 105
Intergenic recombination estimates of the two genetic groups of Hafnia ... 106
Effect of host taxonomic class on the genetic structure of the two Hafnia groups ... 112
DISCUSSION ... 115

Contributions of recombination and mutation in the two genetic groups of the genus Hafnia............115
Conclusion............117
REFERENCES............118

CHAPTER 7............**122**

GENERAL DISCUSSION............**122**

LIFE HISTORY OF THE TWO *HAFNIA* SPECIES............122
FUTURE DIRECTIONS............124
REFERENCES............125

Chapter 1

General introduction

The existence of ecological niches in bacteria

A given ecosystem contains an extensive number of organisms. The number of species in an environment depends on the number of ecological niches, each defined by the resources and conditions that they provide. Bacteria are not an exception to the classic observation of niche specialisation that is observed in higher eukaryotes. Two characteristics of bacteria provide a rich source of biodiversity. Indeed, Dykhuizen (1998) estimated that there are approximately 500,000 species of bacteria in 30 g of soil. First, bacteria are small. Their sizes range from 0.25×1.2 μm for *Haemophilus influenzae* to 8×50 μm for *Oscillatoria* (cyanobacterium) (Madigan 1997). As a consequence of their small size, their habitats are also small. Secondly, bacteria reproduce asexually by binary fission, and mutations will have a direct effect on the fitness of a strain. By contrast, in diploid organisms, the dominant phenotype will mask the mutational effect of the other gene copy. As proven by laboratory studies, adaptation, which is the initial step of speciation, is an easy and rapid process due to the mode of reproduction in bacteria (Bennett and Lenski 1993; Rainey and Travisano 1998).

The history of bacterial species classification

In order to estimate the number of bacterial species in any given environment, there needs to be a consensus definition of what constitutes a species. Until the 1970s, phenotypic classification (based mainly on metabolic characteristics) was used for identifying and categorising bacteria (Sneath 1984). With the advent of genetic techniques in the early 1970s, DNA-DNA hybridisation was used to genotypically distinguish species, and this

method dominated bacterial systematics for over a decade (Wayne *et al.* 1987). 16rRNA-based classification was subsequently introduced following the advent of modern sequencing technology with the aim to develop a global database for bacterial species identification (Stackebrandt and Goebel 1994). Although all of these techniques are still regarded as prerequisites for describing a new species, there are issues concerning the validity of their application and interpretation. The utility of phenotypic (metabolic) characterisation breaks down when mutation or recombination results in a strain acquiring a novel trait or losing the signature phenotypic characteristics of the species. Species delineation using 16S rRNA or DNA-DNA hybridisation approaches are based on arbitrary cut-offs of 3 % divergence and 70 % annealing at 60 °C, respectively (Wayne *et al.* 1987). However there are populations assigned to different genera that show more than 70 % annealing or less than 3 % divergence. A classic example is the genera *Shigella* and *Escherichia* (Lan and Reeves 2002). *Shigella* established its genus status in the 1940s, far before genetic techniques were available for bacterial classification. It was due to the fact that it had clinical significance, and had a specific host - humans. However, genetic classification methods now show that *Shigella* are no further apart from *E. coli* than other pathogenic forms of *E. coli* (e.g. EIEC, *E. coli* O157:H7). *Shigella* strains are quite clearly members of the species *Escherichia coli*, yet are, phenotypically, very atypical *E. coli*. *Shigella* has acquired genes for an intracellular lifestyle and lost genes that code for functions of *E. coli*. Consequently the phenotypic characteristics of *Shigella* resembles other intracellular bacteria such as *Yersinia* sp. and *Providencia* sp. (Dodd and Jones 1982). The advent of molecular genetic techniques has shown that the classification of bacteria based on their clinical significance (Lan and Reeves 2001) has lead to many cases of species misclassification.

By contrast, there are different populations, related well below these arbitrary cut-off values, that are categorised as one species such as *Enterobacter cloacae* (Hoffmann and Roggenkamp 2003).

Incorporation of eukaryotic speciation theory to bacterial speciation

The problems with current methods of bacterial species classification arise because numerical taxonomy approaches and methods involving arbitrary cut-offs to classify species ignores the evolutionary history of bacterial populations. One of the goals of workers in the field of eukaryotic systematics has been to develop a 'universal species concept'. Despite the fact that none of the many attempts to formulate such a 'universal species concept' has been widely accepted, eukaryotic systematists recognise the need to interpret the results of their studies using some form of a species concept. However, it has only been relatively recently that prokaryotic systematists have recognised that real progress in the field will only be achieved using a sound evolutionary framework. At present, an amalgamation of two species concepts has gained fairly broad acceptance by prokaryotic systematists; the ecological species concept and a modification of the biological species concept (Rodriguez-Valera 2002). Studies based on DNA sequence data or other genetic analysis techniques typically show the occurrence of 'sequence clusters', that is the existence of a population of genetically similar taxa that are different from other groups of closely related taxa. Cohan (1994) argues that each of the sequence clusters represents an ecotype. An ecotype begins with a clone with a beneficial mutation or recombination event that enables the clone to exploit a novel niche that the ancestral population can not. Ecological isolation between the new and ancestral population will limit the extent of recombination and the effect of periodic selection (a purging effect of accumulated genetic variation caused by an adaptive mutation or recombination) to within the ecological niche. Therefore, an ecotype may be defined as the boundary for

homologous recombination and periodic selection. The divergence of these ecological populations over time ultimately results in two distinct populations that no longer share genetic information. The consequence of ecological isolation on the genetic relationship of the two populations can be explained with the biological species concept. Dykhuizen and Green (1991) proposed a bacterial species boundary as the limit of homologous recombination. It is known that homologous recombination is less likely to occur as the two populations become more distinct (Vulic *et al.* 1997). This is due to barriers of recombination, such as mismatch repair and SOS systems that limit recombination to within genetically close populations (Matic *et al.* 1996). Therefore, strains share alleles among other members within a species, but not with members of other species because they are genetically too distant.

There are two ways that recombination may influence the ecological structure of bacterial populations. First, it can function to diversify populations by introducing novel adaptive genes. But recombination can also function to maintain the species integrity by shuffling genes within the population. Studies examining the electrophoretic variation of metabolic allozyme loci of different bacterial species suggested that the importance of recombination varies among species (Maynard Smith *et al.* 1993). *Neisseria gonorrhoeae* is known for its natural transformation capability, and is highly sexual. As a result a phylogenetic 'tree' can not be constructed, and the neutral genetic relationship between strains is better described as a 'network'. At the other end of the spectrum, the genus *Salmonella* has a highly clonal population structure, with minimal evidence for homologous intra-species recombination (Beltran *et al.* 1991). This barrier of recombination forms the fundamental basis of the argument for the application of the ecological and biological species concepts to bacteria. The nature of the differences between species in the frequency of recombination is currently unclear, although there are some hypotheses that recombination evolves in species

that are subjected to stress from fluctuating environmental conditions (Claverys *et al.* 2000; Otto and Michalakis 1998).

The case of genotype differentiation in *Hafnia*

The history of *H. alvei* systematics begins from the initial description of the species in 1954 (Moeller 1954). The species name *alvei* (from alveus, meaning beehive) originates from a strain isolated from bees and suspected to be a bee pathogen (Moeller 1954). It was assigned to the genus *Enterobacter* during the 1960s as *Ent. alvei* (Sakazaki 1961) and *Ent. hafniae* (Ewing 1968), owing to its biochemical similarities to other *Enterobacter* species. In 1975, Steigerwalt *et al.* (1976) conducted a wide-scale DNA-DNA hybridisation study on species in the genera *Enterobacter* and *Serratia*. They made two significant observations regarding *Hafnia*: 1) It exhibited only 20 % annealing to *Enterobacter*, and 2) there were two distinct groups within *Hafnia* that had an annealing rate of 50 % to each other. Therefore the genus status of *Hafnia* was re-erected, however the existence of two groups was interpreted as 'indicative of extensive diversity' (Steigerwalt *et al.* 1976), hence they remained as one species, *H. alvei*. Strain characterisation by biochemical profiles (esculin, arbutin, salicin and D-(-)-arabinose utilisation, called Barbe biotypes) failed to show any consistent differences between the two genetic groups, as the most frequent biotype (biotype 1) was present in both *Hafnia* groups (Janda *et al.* 2002). There is an extensive amount of work published on the lipopolysaccharide variation of this genus (Romanowska 2000). The latest serotyping scheme describes 57 serotypes representing 39 O- and 36 H- antigen types (Baturo 1978). Unfortunately there is no additional genetic or phenotypic information on these strains, consequently the extent to which the two groups are serologically different is not known.

Despite the fact that these two groups were perfectly distinguishable using DNA-DNA hybridisation, *Hafnia* did not receive much attention because the species was not considered to be of any clinical significance. Its life history has not been investigated in detail, which has lead to relatively few strains being available for any study on this species. It was not until the 1990s when pathogenic isolates now known as *Escherichia albertii* were thought to be members of the genus *Hafnia*, that *Hafnia* regained the attention of clinical microbiologists (Albert *et al.* 1992; Huys *et al.* 2003; Ridell *et al.* 1995). It is now known that members of this genus can cause extraintestinal infection among patients with an underlying illness and in infants. It also has long been recognised as a significant species in the veterinary and food science fields.

Gordon and FitzGibbon (Gordon and FitzGibbon 1999) initiated an extensive collection of enteric bacteria from a wide range of Australian native mammals, which continues to grow in size. With the generous assistance from researchers around Australia, the collection currently contains over 30 described enteric species isolated from five host taxonomic classes (Teleostei, Amphibia, Reptilia, Mammalia, Aves) and one class of invertebrates (Insecta) as well as from soil, water and sediments. This collection of enteric bacteria has allowed many studies to be conducted, whose broad aims have been to elucidate aspects of the ecology and evolution of these microbes (Gordon and Cowling 2003; Gordon and Lee 1999; Okada and Gordon 2001; Wertz *et al.* 2003; Wertz and Riley 2004). This thesis took advantage of the existence of over 180 isolates biochemically identified as *H. alvei* to investigate the genetic and phenotypic properties of the species. The basis of the study was established by electrophoretically typing 161 isolates (which was the maximum number of isolates at the start of the project) using multilocus enzyme electrophoresis. The biochemical profiles of all isolates were also used as phenotypic information. The next two studies investigated other aspects of the phenotypic diversity of the genus: The intrinsic rate

of growth in minimal glucose media as well as bacteriocin production and resistance. The last project utilised multilocus sequence typing to observe the neutral genetic structure of the two genetic groups.

The aims of this thesis is five fold: 1) to assess the amount of genetic and phenotypic diversity in the genus *Hafnia*, 2) to detect the presence of ecological structure created by biotic / abiotic aspects of the environments from where the isolates were collected, 3) to detect evidence (if any) for sharing of alleles among the two *Hafnia* groups, 4) to investigate the evolutionary mechanisms of the population structure of the two genetic groups, and 5) to identify the ecological characteristics of *Hafnia*.

References

Albert MJ, Faruque SM, Ansaruzzaman M, Islam MM, Haider K, Alam K, Kabir I, Robinsbrowne R (1992) Sharing of virulence-associated properties at the phenotypic and genetic levels between enteropathogenic *Escherichia coli* and *Hafnia alvei*. Journal of Medical Microbiology 37:310-314.

Baturo AP, Raginskaya VP (1978) Antigenic scheme for the *Hafniae*. International Journal of Systematic Bacteriology 28:126-127.

Beltran P, Plock SA, Smith NH, Whittam TS, Old DC, Selander RK (1991) Reference collection of strains of the *Salmonella typhimurium* complex from natural populations. Journal of General Microbiology 137:601-606.

Bennett AF, Lenski RE (1993) Evolutionary adaptation to temperature .2. Thermal niches of experimental lines of *Escherichia coli*. Evolution 47:1-12.

Claverys JP, Prudhomme M, Mortier-Barriere I, Martin B (2000) Adaptation to the environment: *Streptococcus pneumoniae*, a paradigm for recombination-mediated genetic plasticity? Molecular Microbiology 35:251-259.

Cohan FM (1994) The effects of rare but promiscuous genetic exchange on evolutionary divergence in prokaryotes. American Naturalist 143:965-986.

Dodd CER, Jones D (1982) A numerical taxonomic study of the genus *Shigella*. Journal of General Microbiology 128:1933-1957.

Dykhuizen DE (1998) Santa Rosalia revisited - Why are there so many species of bacteria? Antonie van Leeuwenhoek International Journal of General and Molecular Microbiology 73:25-33.

Dykhuizen DE, Green L (1991) Recombination in *Escherichia coli* and the definition of biological species. Journal of Bacteriology 173:7257-7268.

Ewing WH, Fife MA (1968) *Enterobacter hafniae* (the *Hafnia* group). International Journal of Systematic Bacteriology 18:263-271.

Gordon DM, Cowling A (2003) The distribution and genetic structure of *Escherichia coli* in Australian vertebrates: host and geographic effects. Microbiology 149:3575-3586.

Gordon DM, FitzGibbon F (1999) The distribution of enteric bacteria from Australian mammals: host and geographical effects. Microbiology 145:2663-2671.

Gordon DM, Lee J (1999) The genetic structure of enteric bacteria from Australian mammals. Microbiology 145:2673-2682.

Hoffmann H, Roggenkamp A (2003) Population genetics of the nomenspecies *Enterobacter cloacae*. Applied and Environmental Microbiology 69:5306-5318.

Huys G, Cnockaert M, Janda JM, Swings J (2003) *Escherichia albertii* sp nov., a diarrhoeagenic species isolated from stool specimens of Bangladeshi children. International Journal of Systematic and Evolutionary Microbiology 53:807-810.

Janda JM, Abbott SL, Khashe S, Probert W (2002) Phenotypic and genotypic properties of the genus *Hafnia*. Journal of Medical Microbiology 51:575-580.

Lan R, Reeves PR (2001) When does a clone deserve a name? A perspective on bacterial species based on population genetics. Trends in Microbiology 9:419-424.

Lan RT, Reeves PR (2002) *Escherichia coli* in disguise: molecular origins of *Shigella.* Microbes and Infection 4:1125-1132.

Madigan MT, Martinko JM, Parker J (1997) *Brock Biology of Microorganisms*, 8th edition. Prentice Hall, Inc. Upper Saddle River, NJ.

Matic I, Taddei F, Radman M (1996) Genetic barriers among bacteria. Trends in Microbiology 4:69-73.

Maynard Smith J, Smith NH, Orourke M, Spratt BG (1993) How clonal are bacteria? Proceedings of the National Academy of Sciences of the United States of America 90:4384-4388.

Moeller V (1954) Distribution of amino acid decarboxylases in Enterobacteriaceae. Acta Pathologica Microbiologica Scandinavica 35:259-277.

Okada S, Gordon DM (2001) Host and geographical factors influence the thermal niche of enteric bacteria isolated from native Australian mammals. Molecular Ecology 10:2499-2513.

Otto SP, Michalakis Y (1998) The evolution of recombination in changing environments. Trends in Ecology and Evolution 13:145-151.

Rainey PB, Travisano M (1998) Adaptive radiation in a heterogeneous environment. Nature 394:69-72.

Ridell J, Siitonen A, Paulin L, Lindroos O, Korkeala H, Albert MJ (1995) Characterization of *Hafnia alevi* by biochemical tests, random amplified polymorphic DNA PCR, and partial sequencing of 16S rRNA gene. Journal of Clinical Microbiology 33:2372-2376.

Rodriguez-Valera F (2002) Approaches to prokaryotic biodiversity: a population genetics perspective. Environmental Microbiology 4:628-633.

Romanowska E (2000) Immunochemical aspects of *Hafnia alvei* O antigens. FEMS Immunology and Medical Microbiology 27:219-225.

Sakazaki R (1961) Studies on the *Hafnia* group of Enterobacteriaceae. Japan Journal of Medical Science and Biology 14:223-241.

Sneath PHA (1984) Numerical Taxonomy. In: Krieg NR, Holt JG (ed) *Bergey's Manual of Systematic Bacteriology*. Williams and Wilkins Co., Baltimore, pp. 111-118.

Stackebrandt E, Goebel BM (1994) A place for DNA-DNA reassociation and 16S rRNA sequence analysis in the present species definition in bacteriology. International Journal of Systematic Bacteriology 44:846-849.

Steigerwalt AG, Fanning GR, Fife-Asbury MA, Brenner, DJ (1976) DNA relatedness among species of *Enterobacter* and *Serratia*. Canadian Journal of Microbiology 22:121-137.

Vulic M, Dionisio F, Taddei F, Radman M (1997) Molecular keys to speciation - DNA polymorphism and the control of genetic exchange in Enterobacteria. Proceedings of the National Academy of Sciences of the United States of America 94:9763-9767.

Wayne LG, Brenner DJ, Colwell RR, Grimont PAD, Kandler O, Krichevsky MI, Moore LH, Moore WEC, Murray RGE, Stackebrandt E, Starr MP, Truper HG (1987) Report of the ad-hoc-committee on reconciliation of approaches to bacterial systematics. International Journal of Systematic Bacteriology 37:463-464.

Wertz JE, Goldstone C, Gordon DM, Riley MA (2003) A molecular phylogeny of enteric bacteria and implications for a bacterial species concept. Journal of Evolutionary Biology 16:1236-1248.

Wertz JE, Riley MA (2004) Chimeric nature of two plasmids of *Hafnia alvei* encoding the bacteriocins alveicins A and B. Journal of Bacteriology 186:1598-1605.

Chapter 2

Genetic and ecological structure of *Hafnia alvei* in Australia

Published in *Systematics and Applied Microbiology 26, 585-594 (2003)*

Introduction

Hafnia alvei (Moeller 1954) is a commensal species of the family Enterobacteriaceae. It was originally assigned to the genus *Enterobacter* (*Ent. hafniae*) (Ewing 1968) due to similarities in their biochemical profiles. *H. alvei* was later reassigned to its own genus based on a DNA hybridisation study revealing that it had a binding ratio of only 20 % with *Ent. cloacae* (Steigerwalt *et al.* 1976). It has clinical significance as an opportunistic pathogen of humans, having been implicated as the cause of gastrointestinal tract disease, as well as extraintestinal infections, many of which were associated with nosocomial multiple infections (Gunthard and Pennekamp 1996; Klapholz *et al.* 1994; Ramos and Damaso 2000). *H. alvei* is found in water columns (Allen *et al.* 1983; Shirey and Bissonnette 1992), and is also thought to be an extraintestinal pathogen of freshwater fish (Gelev *et al.* 1990). It is frequently isolated from dairy, meat and fish products as one of the major bacterial food contaminants (Gonzalez-Rodriguez *et al.* 2001; Lindberg *et al.* 1998; Tornadijo *et al.* 2001). Although it is currently included in the category of the 'coliform' group, there is some question as to whether *H. alvei* truly serves a role as an indicator species for faecal contamination because of its ubiquity in the environment (Leclerc *et al.* 2001).

Steigerwalt *et al.* (1976) examined the DNA relatedness of four *H. alvei* strains using a hybridisation technique and found that these strains could be classified into two groups. Since then, three other studies have confirmed the existence of two genetic groups within the species for both clinical and non-clinical isolates (Brenner 1981; Janda *et al.* 2002;

Ridell *et al.* 1995). *H. alvei* from wild Australian mammals were also found to split into two genetic groups that have distinct phenotypic characteristics (Okada and Gordon 2001). The two groups differed in their preferred temperature for growth (Okada and Gordon 2001), and their antibiotic resistance profiles (Sherley *et al.* 2004).

In a study of six species of enteric bacteria Gordon and Lee (1999) demonstrated that ecological factors such as the taxonomic family and geographic locality of the host accounted for a significant fraction of the observed genetic variation in four of the five species examined. Although *H. alvei* is reported to be found in a diverse range of host species (Brenner 1981), past studies examining its genetic structure have been restricted to isolates from a single host group, usually mammals. Therefore these previous studies may be underestimating the genetic diversity of *H. alvei* at the species level. This study reports the results of genotypic and phenotypic characterisation of *H. alvei* strains isolated from over 150 host species representing six taxonomic classes and examines the extent to which variation in these characteristics is influenced by the host and its geographic location.

Material and Methods

Isolation of bacteria

One hundred and fifty eight *H. alvei* strains were isolated from faecal or intestinal swabs of 1488 hosts representing 97 bird, 78 mammal, 54 reptile, 8 frog, 6 freshwater fish and 11 arthropod species from over 190 localities throughout Australia between 1993 and 2001. Three strains from water column samples in New South Wales were also included. Three reference strains (type strain ATCC 13337, ATCC 29926, and ATCC 29927) were included in this study in order to determine genetic grouping of the strains.

Techniques used to isolate the strains are described by Gordon and FitzGibbon (1999). Biochemical characteristics of isolates were determined by the BBL Crystal Enteric / Nonfermenter Identification kit (Becton Dickinson). Methyl red and Voges – Proskauer tests were conducted on all strains according to the method of Ewing (1986). Strain designations, host of origin and geographic locality of the strains examined in this study are provided in Table 1. Geographic localities were assigned to one of the major climate zones occurring in Australia based on maps presented by the Bureau of Meteorology (www.bom.gov.au).

Multilocus enzyme electrophoresis

H. alvei strains were genetically characterised according to the methods described by Gordon and Lee (1999). In brief, overnight cultures of *H. alvei* in Luria Bertani broth were collected and sonicated in lysis buffer containing NADP (100 mg/l) and 2-mercapto-ethanol (0.5 ml/l). Supernatants were stored in -70 °C until further use. Cellulose acetate gel electrophoresis was used to detect protein variation in 11 enzyme systems: adenylate kinase (EC 2.7.4.3), glucose-6-dehydrogenase (EC 1.1.1.49), isocitrate dehydrogenase (EC 1.1.1.42), malic enzyme (EC 1.1.1.40), malate dehydrogenase (EC 1.1.1.37), menadione reductase (EC 1.6.99.2), 6-phosphogluconate dehydrogenase (EC 1.1.1.44), phosphoglucose isomerase (EC 5.3.1.9), phosphoglucomutase (EC 5.4.2.2), shikimic acid dehydrogenase (EC 1.1.1.25), and superoxide dismutase (EC 1.15.1.1). Menadione reductase yielded two bands, therefore the total number of loci that was used for analysis was 12. All strains were run at least twice for every locus to confirm their genotype.

Table 1. Source (taxonomic class, species, and state of the host) of isolates of *Hafnia alvei* used in this study, and the MLEE results. Assignment to either genetic group 1 or 2 was based on multilocus enzyme electrophoresis. NSW, New South Wales; QLD, Queensland; TAS, Tasmania; SA, South Australia; WA, Western Australia; NT, Northern Territory; AK, adenylate kinase; G6PDH, glucose-6-dehydrogenase; IDH, isocitrate dehydrogenase; MDH, malate dehydrogenase; ME, malic enzyme; MR, menadione reductase; PGD, 6-phosphogluconate dehydrogenase; PGI, phosphoglucose isomerase; PGM, phosphoglucomutase; SDH, shikimic acid dehydrogenase; SOD, superoxide dismutase.

Strain ID	Host Class	Host Species	State	Group	AK	G6PDH	IDH	MDH	ME	MR-1	MR-2	PGD	PGI	PGM	SDH	SOD
B348	Aves	*Acanthiza katherina*	QLD	1	1	2	8	3	7	6	2	5	1	6	8	4
B429	Aves	*Sericornis frontalis*	NSW	1	1	2	8	3	1	6	2	5	4	6	8	4
B459	Aves	*Malurus cyaneus*	TAS	1	2	2	8	3	7	6	2	5	3	6	13	4
B642	Aves	*Gymnorhina tibicen*	NSW	1	1	2	8	3	6	6	2	4	3	3	8	4
E072	Water		NSW	1	1	2	8	3	6	6	1	5	2	6	12	2
E113																
E095	Water		NSW	1	1	2	8	3	6	4	2	7	2	6	10	4
F138	Actinopterygii	*Nematalosa erebi*	SA	1	1	2	8	3	4	6	0	4	2	4	8	4
F156	Actinopterygii	*Nematalosa erebi*	SA	1	1	2	8	3	6	6	4	5	2	5	10	4
F169	Actinopterygii	*Leiopotheropon unicolor*	SA	1	1	2	8	3	6	6	2	7	3	5	10	4
F182	Actinopterygii	*Macquaria ambigua*	SA	1	1	2	8	3	6	6	2	7	3	6	4	2
F187	Actinopterygii	*Macquaria ambigua*	SA	1	1	2	8	3	4	6	2	5	3	6	4	4
H610	Mammalia	*Homo sapiens*	NSW	1	1	2	8	3	6	2	2	5	3	6	12	4
I031	Insecta	not identified	NSW	1	1	2	8	3	72	6	1	5	3	6	9	4
I040	Insecta	not identified	NSW	1	1	2	6	3	6	6	1	5	3	6	12	4
I044	Insecta	not identified	NSW	1	1	2	8	3	6	6	1	5	3	6	12	4
I097	Insecta	not identified	NSW	1	1	2	8	3	2	6	1	5	6	5	9	4
I100	Insecta	not identified	NSW	1	1	2	8	3	6	6	1	4	3	6	12	2
I101	Insecta	not identified	NSW	1	1	2	8	3	6	6	1	5	3	4	9	4
I128	Insecta	not identified	NSW	1	21	2	8	3	2	6	1	6	6	5	9	4
L009	Amphibia	*Litoria adelaidensis*	WA	1	1	2	8	3	2	6	4	5	2	6	12	4
L069	Amphibia	*Litoria ewingii*	NSW	1	1	2	8	3	2	6	1	5	4	6	9	4
M039	Mammalia	*Rattus rattus*	NSW	1	1	2	8	3	1	6	2	5	5	5	9	4
M044	Mammalia	*Antechinus flavipes*	NSW	1	1	2	8	3	6	6	2	5	2	6	14	4
M051	Mammalia	*Antechinus flavipes*	NSW	1	1	2	8	3	1	6	2	5	4	6	14	4
M061	Mammalia	*Sminthopsis murina*	NSW	1	1	2	8	3	1	6	2	5	2	6	12	4
M082	Mammalia	*Rattus fuscipes*	NSW	1	1	2	8	3	6	6	2	10	2	6	9	4
M089	Mammalia	*Rattus fuscipes*	NSW	1	1	2	8	3	6	6	2	5	2	6	9	4
M165	Mammalia	*Rattus fuscipes*	NSW	1	1	2	8	3	1	6	1	5	4	6	14	4
M172	Mammalia	*Rattus fuscipes*	NSW	1	1	2	8	3	1	64	4	5	2	6	7	4
MI519	Mammalia	*Rattus fuscipes*	NSW	1	1	2	8	3	6	6	2	5	2	6	7	4
MI565	Mammalia	*Rattus fuscipes*	NSW	1	1	2	8	3	1	6	2	5	2	6	7	4
MI577	Mammalia	*Rattus fuscipes*	NSW	1	1	2	8	3	1	6	1	6	1	3	7	4
MI581	Mammalia	*Rattus fuscipes*	NSW	1	1	2	8	3	6	5	2	8	2	6	7	4
MI789	Mammalia	*Rattus fuscipes*	NSW	1	1	2	8	3	6	6	2	5	2	6	7	1

Table 1. Continued.

Strain ID	Host Class	Host Species	State	Group	AK	G6PDH	IDH	MDH	ME	MR-1	MR-2	PGD	PGI	PGM	SDH	SOD
M129	Mammalia	*Burramys parvus*	NSW	1	1	2	8	3	1	6	2	5	4	6	9	4
M153	Mammalia	*Mastacomys fuscus*	NSW	1	1	2	8	3	6	6	2	5	2	6	12	4
MI414	Mammalia	*Dasyurus viverrinus*	TAS	1	1	2	8	3	6	6	2	10	2	6	8	4
MI472	Mammalia	*Dasyurus viverrinus*	TAS	1	1	2	8	3	6	6	2	5	1	5	7	4
MI490 MI494	Mammalia	*Mastacomys fuscus*	NSW	1	1	2	8	3	6	6	2	10	2	6	8	4
MI574	Mammalia	*Mastacomys fuscus*	NSW	1	1	2	8	3	6	6	2	5	2	6	7	4
MI590	Mammalia	*Mastacomys fuscus*	NSW	1	1	2	8	3	6	6	2	7	2	6	7	4
M275	Mammalia	*Antechinus stuartii*	NSW	1	1	2	8	3	6	64	4	5	2	6	12	4
M276	Mammalia	*Antechinus stuartii*	NSW	1	1	2	6	3	6	6	2	5	2	6	12	4
M306	Mammalia	*Antechinus stuartii*	NSW	1	1	2	8	3	1	6	2	5	5	6	7	4
MI782	Mammalia	*Antechinus stuartii*	NSW	1	1	2	8	3	4	6	2	5	6	6	6	4
MI785	Mammalia	*Antechinus stuartii*	NSW	1	1	2	8	3	1	6	2	5	1	5	7	4
MI130	Mammalia	*Nyctophilus geoffroyi*	SA	1	1	2	8	3	6	6	2	5	2	6	7	4
MI147	Mammalia	*Antechinus flavipes*	SA	1	1	2	8	3	1	6	1	6	3	5	6	4
MI442	Mammalia	*Sminthopsis macroura*	SA	1	1	2	8	3	4	6	2	6	3	6	7	4
MI455	Mammalia	*Tachyglossus aculeatus*	SA	1	1	2	6	3	6	6	2	5	4	6	12	4
MI524	Mammalia	*Antechinus swainsonii*	NSW	1	1	2	8	3	6	6	2	5	2	6	7	4
MI640	Mammalia	*Antechinus swainsonii*	VIC	1	1	2	6	3	6	6	2	7	3	6	7	4
MI642	Mammalia	*Pseudomys shortridgii*	VIC	1	1	2	8	3	1	5	2	5	1	3	6	4
MI645	Mammalia	*Nyctophilus geoffroyi*	NSW	1	1	2	8	3	6	6	2	5	2	6	6	4
MI777	Mammalia	*Pseudomys fumeus*	NSW	1	1	2	8	3	6	6	2	5	2	6	7	4
MI779	Mammalia	*Antechinus swainsonii*	NSW	1	1	2	8	3	1	6	1	5	2	5	7	4
R012	Reptilia	*Eulamprus heatwolei*	NSW	1	1	2	8	3	6	6	1	10	2	5	11	4
R015	Reptilia	*Eulamprus heatwolei*	NSW	1	1	2	8	3	6	6	1	5	4	5	14	4
R021	Reptilia	*Eulamprus heatwolei*	NSW	1	1	2	8	3	6	6	2	5	2	5	9	4
R044	Reptilia	*Eulamprus heatwolei*	NSW	1	1	2	8	3	1	6	1	5	2	5	11	4
R055	Reptilia	*Eulamprus heatwolei*	NSW	1	1	2	8	3	6	6	1	5	2	5	11	4
R067	Reptilia	*Eulamprus heatwolei*	NSW	1	1	2	8	3	1	6	2	4	2	5	12	4
R022	Reptilia	*Diplodactylus byrnei*	NSW	1	1	2	8	3	6	6	2	8	4	5	4	4
R023	Reptilia	*Tiliqua rugosa*	SA	1	1	2	8	3	3	6	1	5	2	5	11	4
R027	Reptilia	*Tiliqua rugosa*	SA	1	1	2	8	3	1	8	8	5	2	3	11	4
R089	Reptilia	*Tiliqua rugosa*	SA	1	1	2	8	3	1	6	1	5	4	5	11	4
R114 R124	Reptilia	*Niveoscincus microlepidotus*	TAS	1	1	2	6	3	6	6	1	5	2	5	4	4
R126	Reptilia	*Niveoscincus mettalicum*	TAS	1	1	2	8	3	6	6	1	5	2	5	14	4
R178	Reptilia	*Eulamprus tympanum*	NSW	1	1	2	8	3	1	6	1	5	4	3	11	4
R180	Reptilia	*Eulamprus tympanum*	NSW	1	1	2	8	3	6	6	2	5	2	5	11	4
R269	Reptilia	*Eulamprus tympanum*	NSW	1	1	2	8	3	4	6	2	5	4	5	4	4
R202	Reptilia	*Egernia saxatilis*	NSW	1	1	2	8	3	1	6	2	5	2	5	12	4
R204	Reptilia	*Egernia saxatilis*	NSW	1	1	2	8	3	6	6	2	5	2	5	12	4
R211	Reptilia	*Egernia saxatilis*	NSW	1	1	2	8	3	1	6	2	5	2	5	11	4
R285	Reptilia	*Pseudemoia entrecasteauxii*	WA	1	1	2	8	3	1	6	1	5	6	3	11	4
R373	Reptilia	*Pseudechis porphyriacus*	NSW	1	1	2	8	3	6	6	2	8	6	5	4	4

Table 1. Continued.

Strain ID	Host Class	Host Species	State	Group	AK	G6PDH	IDH	MDH	ME	MR-1	MR-2	PGD	PGI	PGM	SDH	SOD
R384	Reptilia	*Rhinoplocephalus nigrescens*	NSW	1	1	2	8	3	6	6	4	5	3	5	12	4
R394	Reptilia	*Rhinoplocephalus nigrescens*	NSW	1	1	2	8	3	3	6	2	5	3	5	11	4
B255	Aves	*Gymnorhina tibicen*	NSW	2	1	2	8	3	4	4	6	6	8	3	3	4
F118 F120 F127	Actinopterygii	*Leiopotheropon unicolor*	SA	2	1	2	8	4	4	4	6	6	3	5	4	4
F121	Actinopterygii	*Leiopotheropon unicolor*	SA	2	1	2	8	3	4	4	6	6	6	5	4	4
F135 F152	Actinopterygii	*Nematalosa erebi*	SA	2	1	2	8	4	4	4	6	6	6	5	8	4
F171 F292 F383 F404	Actinopterygii	*Macquaria ambigua*	SA	2	1	2	8	3	4	4	6	6	6	5	8	4
F184 F385	Actinopterygii	*Macquaria ambigua*	SA	2	1	2	8	3	4	4	6	6	6	5	4	4
F258	Actinopterygii	*Macquaria ambigua*	SA	2	1	2	8	4	4	4	6	6	6	5	8	4
F231	Actinopterygii	*Bidyanus welchi*	SA	2	3	2	8	3	4	4	6	6	3	5	4	4
F286 F298 F300 F348 F364	Actinopterygii	*Macquaria ambigua*	SA	2	1	2	8	3	2	7	7	6	6	5	4	4
F296 F314	Actinopterygii	*Macquaria ambigua*	SA	2	1	2	8	3	2	7	7	6	6	5	8	4
F311	Actinopterygii	*Macquaria ambigua*	SA	2	1	2	8	3	4	4	6	6	8	5	8	4
F321	Actinopterygii	*Macquaria ambigua*	SA	2	1	2	9	3	4	4	6	10	6	5	4	4
F397	Actinopterygii	*Macquaria ambigua*	SA	2	1	2	8	3	4	4	6	6	6	5	5	4
L013	Amphibia	*Litoria moorie*	WA	2	1	2	8	4	4	4	6	6	4	6	2	4
H698	Mammalia	*Homo sapiens*	NSW	2	1	2	8	3	4	4	6	5	6	5	4	4
M132	Mammalia	*Burramys parvus*	NSW	2	1	2	8	3	4	4	6	6	6	6	3	4
MI163	Mammalia	*Phascogale tapoatafa*	WA	2	1	2	7	3	4	4	6	10	6	5	6	4
MI164	Mammalia	*Dasyurus geoffroii*	WA	2	1	2	8	3	4	4	6	6	6	5	6	4
MI177	Mammalia	*Notomys fuscus*	QLD	2	1	2	8	4	4	3	5	6	3	5	6	4
MI230	Mammalia	*Antechinus bellus*	NT	2	1	2	8	3	4	4	6	6	6	5	4	4
MI231 MI259 MI261	Mammalia	*Dasyurus hallucatus*	NT	2	1	2	8	3	4	4	6	6	6	5	6	4
MI387	Mammalia	*Dasyurus hallucatus*	NT	2	1	2	8	3	4	4	6	5	5	6	2	4
MI389	Mammalia	*Dasyurus hallucatus*	NT	2	1	2	8	3	4	4	6	5	5	6	3	4
MI240	Mammalia	*Antechinus stuartii*	NSW	2	1	2	8	4	4	3	5	6	3	5	6	4
MI242	Mammalia	*Antechinus stuartii*	NSW	2	1	2	8	3	4	4	6	6	6	5	6	4
MI293	Mammalia	*Rattus rattus*	NSW	2	1	2	8	7	4	4	6	6	3	6	4	4
MI310	Mammalia	*Rattus fuscipes*	NSW	2	1	2	8	2	1	7	6	6	6	5	3	4
MI319	Mammalia	*Rattus fuscipes*	NSW	2	1	2	8	7	4	4	6	6	3	5	4	4
MI332	Mammalia	*Dasyurus viverrinus*	TAS	2	1	2	8	3	4	4	6	6	6	5	3	4
MI355	Mammalia	*Antechinus flavipes*	SA	2	1	2	8	3	4	4	6	6	6	5	6	4
MI398	Mammalia	*Dasyurus viverrinus*	TAS	2	1	2	8	3	4	4	6	6	8	5	4	4

Table 1. Continued.

Strain ID	Host Class	Host Species	State	Group	AK	G6PDH	IDH	MDH	ME	MR-1	MR-2	PGD	PGI	PGM	SDH	SOD
MI400 MI408	Mammalia	*Dasyurus viverrinus*	TAS	2	1	2	8	3	4	4	6	6	3	3	4	4
MI463 MI464	Mammalia	*Dasyurus viverrinus*	TAS	2	1	2	8	3	1	6	6	6	6	5	4	4
MI468	Mammalia	*Dasyurus viverrinus*	TAS	2	1	2	8	3	4	4	6	6	6	6	6	4
MI470	Mammalia	*Dasyurus viverrinus*	TAS	2	1	2	8	3	4	4	6	6	6	5	4	4
MI454	Mammalia	*Tachyglossus aculeatus*	SA	2	1	2	8	3	1	6	6	6	6	6	6	4
MI476	Mammalia	*Ornithorhynchus anatinus*	TAS	2	1	2	8	3	4	4	6	6	6	5	1	4
MI690 MI692	Mammalia	*Homo sapiens*	WA	2	1	2	8	3	4	4	6	6	8	5	4	4
MI695	Mammalia	*Lasiorhinus latifrons*	SA	2	1	2	8	4	4	4	6	6	6	5	6	4
MI761	Mammalia	*Trichosurus caninus*	NSW	2	1	2	8	3	4	4	6	6	3	3	4	4
MI763	Mammalia	*Lagorchestes hirsutus*	NT	2	1	2	8	3	4	4	6	6	3	3	4	4
MI765	Mammalia	*Potorous tridactylus*	NSW	2	1	2	8	4	4	3	5	6	2	5	6	4
R008 R069	Reptilia	*Eulamprus heatwolei*	NSW	2	1	2	6	3	4	4	6	6	4	5	4	4
R018	Reptilia	*Eulamprus heatwolei*	NSW	2	1	2	6	3	4	4	6	3	4	5	4	4
R042 R048 R053 R054 R057 R061	Reptilia	*Eulamprus heatwolei*	NSW	2	1	2	8	3	4	4	6	6	4	5	4	4
R117	Reptilia	*Niveoscincus microlepidotus*	TAS	2	1	2	8	3	4	3	5	6	6	5	11	4
R119	Reptilia	*Niveoscincus microlepidotus*	TAS	2	1	2	8	3	4	4	6	9	7	5	11	4
R122	Reptilia	*Niveoscincus microlepidotus*	TAS	2	1	2	8	4	4	4	6	6	6	5	11	4
R130	Reptilia	*Tiliqua rugosa*	SA	2	1	2	8	3	4	4	6	6	4	5	2	4
R161	Reptilia	*Pseudemoia entrecasteauxii*	NSW	2	1	2	8	3	4	4	6	6	6	5	11	4
R175	Reptilia	*Rhinocephalus bicolour*	WA	2	1	2	8	4	4	4	6	6	3	5	11	4
R184	Reptilia	*Eulamprus tympanum*	NSW	2	1	2	8	3	4	4	6	6	4	3	2	4
R227	Reptilia	*Lerista distinguenda*	WA	2	1	2	8	3	5	3	5	9	4	3	11	4
R231	Reptilia	*Ramphotyphlops* sp.	WA	2	1	3	8	4	4	4	6	6	6	3	11	4
R235	Reptilia	*Ramphotyphlops australis*	WA	2	1	2	8	3	1	7	7	6	6	3	13	4
R252	Reptilia	*Echopsis curta*	WA	2	1	2	8	3	4	3	5	9	6	3	11	4
R257	Reptilia	*Parasuta gouldii*	WA	2	1	2	8	3	4	4	6	8	6	3	11	4

Statistical analysis

Distance matrices were constructed by determining the number of loci at which two strains differed for all pairwise strain combinations. Phenograms were constructed from the distance matrices using the UPGMA.

Mean genome diversity of all strains (H) was estimated as $H = \sum h_j/m$, where m is the number of loci and $h_j = (1-\sum p_i^2)[n/n-1]$, where p_i is the frequency of allele *i* at locus *j*, and n is the number of isolates. Mean genome diversity based only on electrophoretic types (H_{ET}) was also calculated (Maynard Smith *et al.* 1993). Haplotype diversity (G) was estimated as $G = (1-\sum g_i^2)[n/n-1]$, where g_i is the frequency of the *i*th electrophoretic type for n isolates.

Linkage disequilibrium was estimated using the index developed by Brown *et al.* (Brown *et al.* 1980). The average number of loci at which two isolates differ is $K = \sum h_j$, where there are n(n-1)/2 such pairs. If the alleles of different loci are independent of each other, the expected variance of K is $V_e = \sum h_j(1-h_j)$. If the observed variance of K (V_o) is equal to V_e, then this indicates the absence of linkage disequilibrium, and V_o/V_e will equal 1. To test if V_o/V_e is significantly different from one, a randomisation approach was taken. Samples were generated by randomly selecting alleles at a locus, without replacement, and this procedure was repeated for each locus to eliminate any linkage disequilibrium (Souza *et al.* 1992). The observed value of V_o/V_e was compared with 1000 randomly generated samples in order to determine the probability with which the observed V_o/V_e could have arisen by chance (Souza *et al.* 1992).

AMOVA was used to assess the extent to which host taxonomic group or climate zone explained the observed genetic and biochemical variation (Peakall and Smouse 2001). Phenograms were used to visualise the genetic similarity among strains from different

populations (i.e. host taxonomic class). Phenograms were constructed using the pairwise population differentiation estimates produced by the AMOVA analyses.

Results

Genetic structure of *Hafnia alvei*

The phenogram constructed from MLEE analysis of the 161 isolates and three reference strains showed that *H. alvei* consists of two major groups (70 representative strains shown in Figure 1), which were assigned to genetic group 1 (n = 83, including ATCC 13337 and ATCC 29926) and 2 (n = 81, including ATCC 29927). Allelic assignments for all 161 strains are listed in Table 1. Genetic group accounted for 43 % (AMOVA: $P < 0.001$) of the observed genetic variation in the 161 isolates. In genetic group 1, 73 electrophoretic types out of 81 isolates were identified, whereas 46 electrophoretic types were identified out of 80 group 2 isolates (Table 2). Haplotype diversities of the two genetic clusters were 0.991 for genetic group 1 and 0.903 for group 2 (Table 2).

Table 2. Genetic diversity estimates of *Hafnia alvei*. n, number of isolates; H, mean genome diversity based on all isolates; H_{ET}, mean genome diversity based on electrophoretic types; G, haplotype diversity; V_o/V_e, multilocus linkage disequilibrium estimate; $P > V_o/V_e$, probability of alleles being randomly associated; V_o/V_e ETs, multilocus linkage disequilibrium estimate based on electrophoretic types; $P > V_o/V_e$ ETs, probability of a random association of alleles based on electrophoretic types.

Group	n	H	H_{ET}	G	Vo/Ve	P>Vo/Ve	Vo/Ve ETs	P>Vo/Ve ETs
all	161	0.442	0.454	0.982	2.148	<0.001	1.798	<0.001
1	81	0.348	0.358	0.991	1.082	0.14	1.029	0.31
2	80	0.292	0.338	0.903	1.486	<0.001	1.271	<0.001

Figure 1. Genetic relationship of *Hafnia alvei* isolates based on MLEE. Thirty five isolates were randomly chosen from each of the two genetic clusters. ATCC13337T, ATCC29926, and ATCC29927 are also included. The phenogram was created using the unweighted pair group method with arethmetic mean (UPGMA). The genetic distance is expressed as the proportion of loci at which two isolates differ.

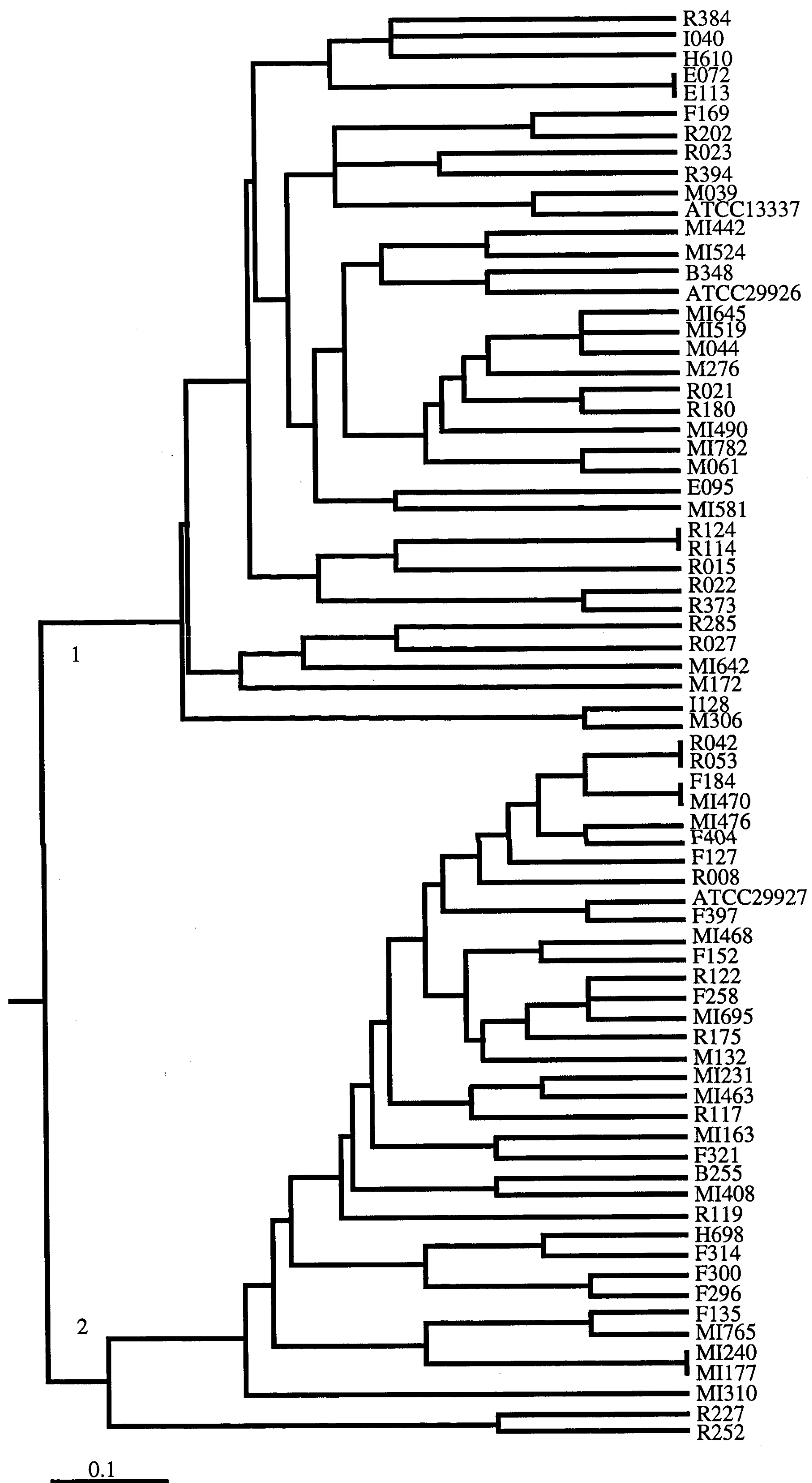


Figure 1.

There was a trend for isolates from group 1 to show greater genetic diversity compared to group 2 isolates. The average number of alleles and range of allelic diversity for individual loci in genetic group 1 were 4.4 and 0 - 0.87, respectively, whereas those for group 2 were 4.2 and 0 - 0.75, respectively, with nine out of 12 loci having higher allelic diversity in group 1 (Table 3).

Table 3. Allelic diversity estimates of the 12 loci for *Hafnia alvei.* AK, adenylate kinase; G6PDH, glucose-6-dehydrogenase; IDH, isocitrate dehydrogenase; MDH, malate dehydrogenase; ME, malic enzyme; MR, menadione reductase; PGD, 6-phosphogluconate dehydrogenase; PGI, phosphoglucose isomerase; PGM, phosphoglucomutase; SDH, shikimic acid dehydrogenase; SOD, superoxide dismutase.

locus	all strains		1		2	
	No. of alleles	h_i	No. of alleles	h_i	No. of alleles	h_i
AK	4	0.037	3	0.049	2	0.025
G6PDH	2	0.012	1	0.000	2	0.025
IDH	4	0.130	2	0.139	4	0.121
MDH	4	0.193	1	0.000	4	0.353
ME	8	0.693	7	0.618	4	0.291
MR-1	8	0.617	6	0.166	4	0.384
MR-2	7	0.711	4	0.530	3	0.308
PGD	8	0.635	6	0.443	6	0.234
PGI	8	0.780	6	0.674	7	0.634
PGM	4	0.572	4	0.578	3	0.378
SDH	14	0.868	10	0.872	9	0.745
SOD	3	0.060	3	0.119	1	0.000

The genetic groups also significantly differed in linkage disequilibrium estimates (V_o/V_e) (Table 2). Group 1 had a V_o/V_e that was not significantly different from one (1.038, $P > 0.3$), suggesting the random association of alleles. On the other hand, the V_o/V_e of group 2 was 1.563, indicating a significant level of linkage disequilibrium (Table 2). Estimates based only on the electrophoretic types showed a similar pattern (Table 2).

Biochemical profile of *H. alvei*

There was significant variation in the biochemical and physiological profile of the isolates between the two genetic groups (AMOVA, øPT = 0.29, $P < 0.01$). The frequency of 11 of 34 biochemical and physiological characteristics differed significantly between the two groups (Table 4). The largest difference in the proportion positive for a biochemical trait was for the utilisation of malonate; 97.5 % of the isolates were positive in genetic group 1 whereas only 22.5 % of isolates were positive in group 2. Acid production from fermentation of glucose (methyl red test) was also different between the two genetic groups, with 100 % positive for group 1 strains as opposed to 29 % positive for group 2 strains.

The variation in biochemical profiles between these two genetic clusters of *H. alvei* was sufficiently great that the profiles could be used to assign an isolate to one of the two genetic clusters. Assignment tests implemented by Cornuet *et al.* (1999) www.ensam.inra.fr/URLB/geneclass/geneclass.html] using the procedure whereby the isolate to be assigned was not included in the data set revealed that 95 % of the isolates were correctly assigned to their genetic group based on the biochemical profile variation alone. An independent sample consisting of 31 isolates from birds, reptiles and mammals were assigned to genetic groups 1 or 2 based only on their biochemical profiles. These isolates were then characterised using MLEE. MLEE analysis revealed that 95 % of the isolates had been correctly assigned.

Table 4. Biochemical profile of the two genetic clusters of *Hafnia alvei*. Numbers indicate the percentage of strains showing a positive reaction for each substrate. Bold indicates characteristics that were significantly different between genetic groups 1 and 2. NS, non-significantly different between the two genetic clusters in their biochemical profile.

Substrate or test	all	Group 1	Group 2	$p > \chi^2$
Arabinose	**94.4**	**90.1**	**98.8**	**0.01**
Mannose	99.4	100.0	98.8	NS
Sucrose	0.0	0.0	0.0	NS
Melibiose	**16.2**	**9.9**	**22.5**	**0.03**
Rhamnose	**78.9**	**60.5**	**97.5**	**0.01**
Sorbitol	1.2	0.0	2.5	NS
Mannitol	99.4	100.0	98.8	NS
Adonitol	0.0	0.0	0.0	NS
Galactose	99.4	100.0	98.8	NS
Inositol	0.0	0.0	0.0	NS
p-n-p-phosphate	**8.1**	**3.7**	**12.5**	**0.04**
p-n-p-α-β-glucoside	24.2	21.0	27.5	NS
p-n-p-β-galactoside	**70.2**	**53.1**	**87.5**	**0.01**
proline-p-nitroanilide	100.0	100.0	100.0	NS
p-n-p-bis-phosphate	99.4	100.0	98.8	NS
p-n-p-xyloside	0.0	0.0	0.0	NS
p-n-p-α-arabinoside	10.6	11.1	10.0	NS
p-n-p-phosphorylcoline	3.1	2.5	3.8	NS
p-n-p-β-glucuronide	0.6	0.0	1.3	NS
p-n-p-N-acetyl-glucosaminide	**80.1**	**63.0**	**97.5**	**0.01**
γ-L-glutamyl-p-nitroanilide	95.7	93.8	97.5	NS
Esculin	**9.9**	**19.8**	**0.0**	**0.01**
p-nitro-DL-phenylalanine	0.6	1.2	0.0	NS
Urea	74.5	72.8	76.3	NS
Glycine	5.6	8.6	2.5	NS
Citrate	70.2	59.3	81.3	0.03
Malonate	**60.3**	**97.5**	**22.5**	**0.01**
Tetrazolium	**66.5**	**74.1**	**58.8**	**0.04**
Arginine	**78.3**	**67.9**	**88.8**	**0.01**
Lysine	100.0	100.0	100.0	NS
Methyl red test	**64.6**	**100.0**	**28.8**	**0.01**
Motility	**47.2**	**33.3**	**61.3**	**0.01**
Indole	**1.9**	**3.7**	**0.0**	**0.05**
Voges-Proskauer test	**90.7**	**81.5**	**100.0**	**0.01**

Correlation between distance matrices of the genetic and biochemical profiles was assessed using a Mantel test (Peakall and Smouse 2001). Distance matrices derived using all isolates revealed that the genetic and biochemical results were significantly correlated ($r = 0.265$, $P < 0.01$). However, there was no correlation between the genetic profile and biochemical characteristics of strains within a genetic group (group 1: $r = 0.06$, $P > 0.1$; group 2: $r = 0.08$, $P > 0.1$).

Ecological factors

Isolates from the two genetic clusters were non-randomly distributed with respect to the taxonomic class of the host. Strains from freshwater fish were predominantly in genetic group 2 (83 %, $n = 29$). Strains from birds, frogs, invertebrates and water isolates were most often members of genetic group 1 (89 %, $n = 18$). Isolates from reptiles ($n = 44$) and mammals ($n = 70$) were equally likely to be members of either group 1 or 2. For strains isolated from mammals, 90 % of genetic group 1 strains were from animals that weighed less than 200 g, whereas genetic group 2 strains were isolated from animals of any body weight (Figure 2). No body size effect could be detected in reptiles. Most reptilian species in this study weighed less than 200 g and there were insufficient samples from large reptiles.

Within each genetic cluster, taxonomic class of the host explained a significant amount of the genetic variation. Analysis of molecular variance (AMOVA) revealed that host taxonomic class explained 21 % ($P < 0.01$) and 12 % ($P < 0.01$) of the variation in genetic groups 1 and 2, respectively. The relationships among strains from different host taxonomic classes are depicted using a phenogram (Figure 3) constructed with the pairwise population differentiation estimates derived from AMOVA analyses (Table 5).

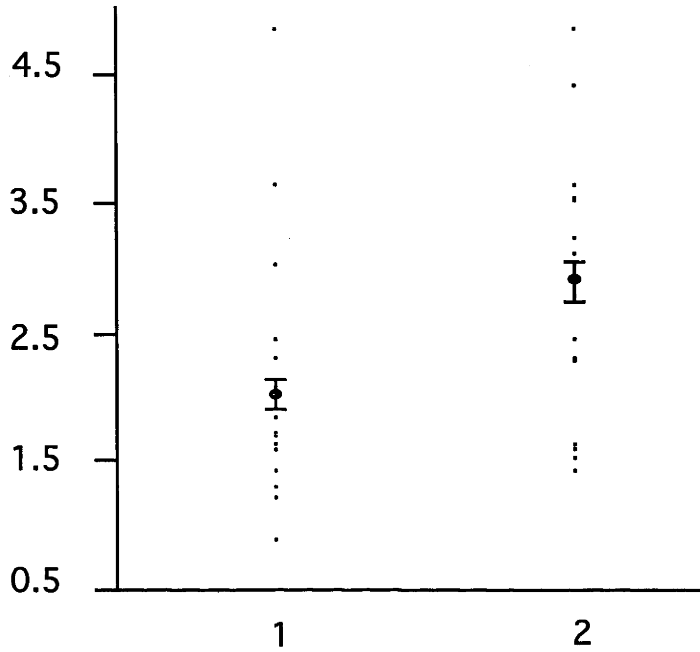


Figure 2. Body mass of mammalian host species from which *Hafnia alvei* genetic clusters A and B strains were isolated. Solid circles and error bars depict the mean and 95 % confidence intervals (genetic group 1, n = 37; group 2, n = 33).

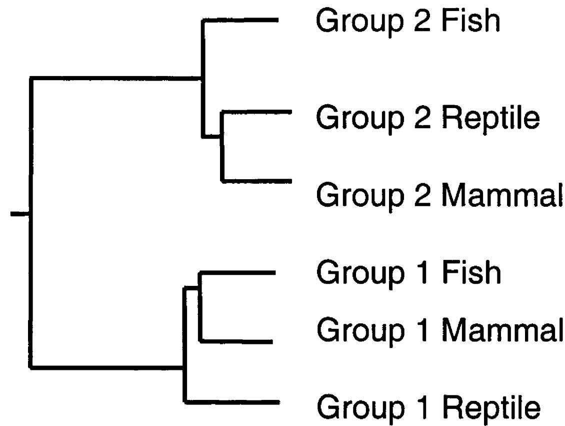


Figure 3. Genetic relationship of *Hafnia alvei* strains isolated from the three major host taxonomic classes. The phenogram was constructed using the pairwise differentiation estimates derived from the MLEE data (Table 5).

Table 5. Pairwise differentiation estimates (Analysis of Molecular Variance) for the genetic profile of *Hafnia alvei* isolated from hosts of different taxonomic classes. The values above the diagonal are for genetic group 1 strains and those below for group 2 strains.

	Fish	Mammal	Reptile
Fish	-	0.139*	0.174
Mammal	0.109	-	0.215
Reptile	0.165	0.096	-

* all non-0 estimates are significant, P < 0.01.

Among group 1 strains those from fish and mammals are most alike, whilst strains from reptiles are the most divergent. For group 2 strains, those from mammals and reptiles are similar, whilst fish are the most divergent.

Taxonomic class of the host also accounted for a significant amount of the variation in the biochemical properties within both genetic groups (20 % for group 1, AMOVA: $P < 0.01$; 24 % for group 2, AMOVA: $P < 0.01$). Biochemical profile data revealed a similar relationship among isolates from the three taxonomic classes for genetic group 2 strains as was observed for the genetic data (Figure 4, Table 6). However, for genetic group 1, isolates from mammals had very divergent biochemical profiles compared to those from fish and reptiles.

After accounting for host taxonomic class, neither state nor climate zone were found to explain any of the observed genetic or biochemical variation.

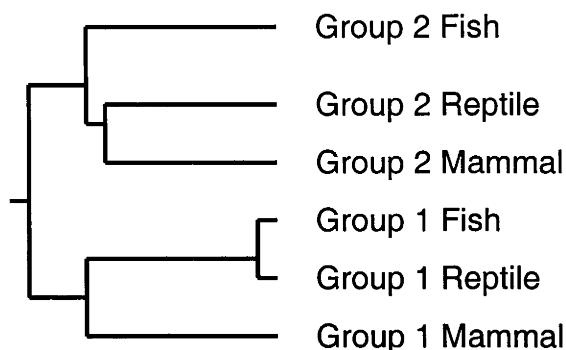


Figure 4. Biochemical relationship of *Hafnia alvei* strains isolated from the three major host taxonomic classes. The phenogram was constructed using the pairwise differentiation estimates derived from the biochemical profile data (Table 6).

Table 6. Pairwise differentiation estimates (Analysis of Molecular Variance) for the biochemical profile of *Hafnia alvei* isolated from hosts of different taxonomic classes. The values above the diagonal are for genetic group 1 strains and those below for group 2 strains.

	Fish	Mammal	Reptile
Fish	-	0.281*	0.0
Mammal	0.232	-	0.224
Reptile	0.261	0.229	-

* all non-0 estimates are significant, P < 0.01.

Discussion

According to DNA hybridisation criteria, genetic groups 1 and 2 have the potential of being regarded as two species. This conclusion is based on the observation that there was only 51 – 55 % DNA annealing between strains of the two clusters (Steigerwalt *et al.* 1976). By convention, for two strains to be regarded as the same species 70 % annealing should be achieved at 60 °C (Johnson 1986; Wayne *et al.* 1987). *Salmonella enterica* is an example of a species that was once divided into over 50 species by serotyping (Kaufmann 1966). However, DNA hybridisation and a later MLEE study revealed that these serovar groups should be considered as subspecies within a single species (Beltran *et al.* 1988; Crosa *et al.* 1973). Another example of genetic clustering within a species is *Escherichia coli*, which consists of four well-recognised genetic clusters (Herzer *et al.* 1990). Variation among these clusters explains 36 % of the electrophoretic variation observed in the *E. coli* reference collection [AMOVA result using MLEE data from Ochman (1984)], which is comparable to the 43 % for *H. alvei* observed in this study. While DNA hybridisation results suggest that the *H. alvei* clusters are more distinct than are the subgroups of *S. enterica*, the extent of MLEE differentiation is no greater than that observed in *E. coli*.

Currently there is no defined criterion that must be met to designate a bacterial species, but a polyphasic approach investigating both phenotypic and genotypic properties seems to be most reliable in terms of bacterial taxonomy (Lan and Reeves 2001; Vandamme *et al.* 1996). There have been several species concepts of eukaryotes that have been applied to bacteria including the ecological species concept (Cohan 2001; Vanvalen 1976), and biological species concept (Dykhuizen and Green 1991; Mayr 1982). The ecological species concept assumes that each bacterial species occupies a unique ecological niche. As a consequence, a beneficial mutation arising within a species will result in a periodic selection event in that species, but not in other closely related species (Cohan 1994). If a

bacterial species is defined by the ability of a 'population' to be able to maintain itself and not be subjected to periodic selection events by other genetically closely related populations, then the 'population' would be eligible for species status. The non-random distribution of strains from the two clusters among different host groups and their different biochemical (i.e. phenotypic) characteristics suggest that strains from the two clusters occupy different ecological niches.

The biological species concept for bacteria assumes that a species boundary is defined by the presence of lateral gene transfer (Dykhuizen and Green 1991). Lateral gene transfer is known to be a function of sequence divergence, whereby the frequency of recombination decreases exponentially with increasing DNA sequence divergence (Roberts and Cohan 1993; Vulic *et al.* 1997). As a consequence, a bacterial cell will incorporate genes that are from a donor of the same species more frequently than from a donor of a different species. For *E. coli*, there is extensive evidence of recombination in the housekeeping genes *trp*, *gnd*, *phoA* (Bisercic *et al.* 1991; Dykhuizen and Green 1991; Nelson and Selander 1994). Because of lateral gene transfer, *E. coli* maintains its species integrity, although specific adaptive traits that seem to transfer among strains within the genetic clusters of *E. coli* will tend to keep these clusters distinct. Further genetic analysis such as multilocus sequence typing (MLST) (Maiden *et al.* 1998) will provide information on the difference in lateral gene exchange and mutation rates among- and within- clusters of *H. alvei*. Genome subtraction studies (Pradel *et al.* 2002) will also provide information on cluster-specific genes that maybe responsible for the diversification within the genus *Hafnia*.

In summary, this study has confirmed the existence of two genetic clusters of *H. alvei*. The two clusters differed in their host distribution, degree of genetic diversity and biochemical characteristics. Within each genetic cluster the taxonomic class of the host, but not

geography, explained a significant amount of both genetic and biochemical variation. Similarly, the results indicate that the strains within a cluster are adapted to different host environments. Further work is needed to determine the nature of the traits responsible for this adaptation.

References

Allen DA, Austin B, Colwell RR (1983) Numerical taxonomy of bacterial isolates associated with a fresh water fishery. Journal of General Microbiology 129:2043-2062.

Beltran P, Musser JM, Helmuth R, Farmer JJ, Frerichs WM, Wachsmuth IK, Ferris K, McWhorter AC, Wells JG, Cravioto A, Selander RK (1988) Toward a population genetic analysis of *Salmonella* - genetic diversity and relationships among strains of serotypes *S. choleraesuis*, *S. derby*, *S. dublin*, *S. enteritidis*, *S. heidelberg*, *S. infantis*, *S. newport*, and *S. typhimurium*. Proceedings of the National Academy of Sciences of the United States of America 85:7753-7757.

Bisercic M, Feutrier JY, Reeves PR (1991) Nucleotide sequences of the *gnd* genes from nine natural isolates of *Escherichia coli* - evidence of intragenic recombination as a contributing factor in the evolution of the polymorphic *gnd* locus. Journal of Bacteriology 173:3894-3900.

Brenner DJ (1981) Chapter 88. Introduction to the family Enterobacteriaceae. In: Starr MP, Stolp, H, Trueper HG, Balows A, Schlegel HG (ed) *The Prokaryotes*. Springer-Verlag, Berlin, pp. 1105-1127.

Brown AHD, Feldman MW, Nevo E (1980) Multilocus structure of natural populations of *Hordeum spontaneum*. Genetics 96:523-536.

Cohan FM (1994) Genetic exchange and evolutionary divergence in prokaryotes. Trends in Ecology and Evolution 9:175-180.

Cohan FM (2001) Bacterial species and speciation. Systematic Biology 50:513-524.

Cornuet JM, Piry S, Luikart G, Estoup A, Solignac M (1999) New methods employing multilocus genotypes to select or exclude populations as origins of individuals. Genetics 153:1989-2000.

Crosa JH, Brenner DJ, Ewing WH, Falkow S (1973) Molecular relationships among *Salmonelleae*. Journal of Bacteriology 115:307-315.

Dykhuizen DE, Green L (1991) Recombination in *Escherichia coli* and the definition of biological species. Journal of Bacteriology 173:7257-7268.

Ewing WH (1986) *Edwards and Ewing's Identification of Enterobacteriaceae*. Elsevier, New York.

Ewing WH, Fife MA (1968) *Enterobacter hafniae* (the *Hafnia* group). International Journal of Systematic Bacteriology 18:263-271.

Gelev I, Gelev E, Steigerwalt AG, Carter GP, Brenner DJ (1990) Identification of the bacterium associated with haemorrhagic septicaemia in rainbow trout as *Hafnia alvei*. Research in Microbiology 141:573-576.

Gonzalez-Rodriguez MN, Sanz JJ, Santos JA, Otero A, Garcia-Lopez ML (2001) Bacteriological quality of aquacultured freshwater fish portions in prepackaged trays stored at 3 degrees C. Journal of Food Protection 64:1399-1404.

Gordon DM, FitzGibbon F (1999) The distribution of enteric bacteria from Australian mammals: host and geographical effects. Microbiology 145:2663-2671.

Gordon DM, Lee J (1999) The genetic structure of enteric bacteria from Australian mammals. Microbiology 145:2673-2682.

Gunthard H, Pennekamp A (1996) Clinical significance of extraintestinal *Hafnia alvei* isolates from 61 patients and review of the literature. Clinical Infectious Diseases 22:1040-1045.

Herzer PJ, Inouye S, Inouye M, Whittam TS (1990) Phylogenetic distribution of branched RNA-linked multicopy single-stranded DNA among natural isolates of *Escherichia coli.* Journal of Bacteriology 172:6175-6181.

Janda JM, Abbott SL, Khashe S, Probert W (2002) Phenotypic and genotypic properties of the genus *Hafnia.* Journal of Medical Microbiology 51:575-580.

Johnson JL (1986) Nucleic acids in bacterial classification. In: Sneath PHA, Mair NS, Sharp NE, Holts JG (ed) *Bergey's Manual of Systematic Bacteriology.* Williams and Wilkins, Baltimore, p 972-975.

Kaufmann F (1966) *The Bacteriology of Enterobacteriaceae.* Williams and Wilkins, Baltimore.

Klapholz A, Lessnau KD, Huang B, Talavera W, Boyle JF (1994) *Hafnia alvei* - respiratory tract isolates in a community hospital over a three-year period and a literature review. Chest 105:1098-1100.

Lan R, Reeves PR (2001) When does a clone deserve a name? A perspective on bacterial species based on population genetics. Trends in Microbiology 9:419-424.

Leclerc H, Mossel DAA, Edberg SC, Struijk CB (2001) Advances in the bacteriology of the coliform group: Their suitability as markers of microbial water safety. Annual Review of Microbiology 55:201-234.

Lindberg AM, Ljungh A, Ahrne S, Lofdahl S, Molin G (1998) Enterobacteriaceae found in high numbers in fish, minced meat and pasteurised milk or cream and the presence of toxin encoding genes. International Journal of Food Microbiology 39:11-17.

Maiden MCJ, Bygraves JA, Feil E, Morelli G, Russell JE, Urwin R, Zhang Q, Zhou JJ, Zurth K, Caugant DA, Feavers IM, Achtman M, Spratt BG (1998) Multilocus sequence typing - a portable approach to the identification of clones within

populations of pathogenic microorganisms. Proceedings of the National Academy of Sciences of the United States of America 95:3140-3145.

Maynard Smith J, Smith NH, Orourke M, Spratt BG (1993) How Clonal Are Bacteria? Proceedings of the National Academy of Sciences of the United States of America 90:4384-4388.

Mayr E (1982) *The Growth of Biological Thought: Diversity, Evolution and Inheritance.* Harvard University Press, Cambridge, Massachusetts.

Moeller V (1954) Distribution of amino acid decarboxylases in Enterobacteriaceae. Acta Pathologica Microbiologica Scandinavica 35:259-277.

Nelson K, Selander RK (1994) Intergeneric transfer and recombination of the 6-phosphogluconate dehydrogenase gene (*gnd*) in enteric bacteria. Proceedings of the National Academy of Sciences of the United States of America 91:10227-10231.

Ochman H, Selander RK (1984) Standard reference strains of *Escherichia coli* from natural populations. Journal of Bacteriology 157:690-693.

Okada S, Gordon DM (2001) Host and geographical factors influence the thermal niche of enteric bacteria isolated from native Australian mammals. Molecular Ecology 10:2499-2513.

Peakall R, Smouse PE (2001) GenAlEx V5: Genetic Analysis in Excel. Population genetic software for teaching and research.

Pradel N, Leroy-Setrin S, Joly B, Livrelli V (2002) Genomic subtraction to identify and characterize sequences of Shiga toxin-producing *Escherichia coli* O91 : H21. Applied and Environmental Microbiology 68:2316-2325.

Ramos A, Damaso D (2000) Extraintestinal infection due to *Hafnia alvei.* European Journal of Clinical Microbiology and Infectious Diseases 19:708-710.

Ridell J, Siitonen A, Paulin L, Lindroos O, Korkeala H, Albert MJ (1995) Characterization of *Hafnia alvei* by biochemical tests, random amplified polymorphic DNA PCR, and partial sequencing of 16S rRNA gene. Journal of Clinical Microbiology 33:2372-2376.

Roberts MS, Cohan FM (1993) The Effect of DNA sequence divergence on sexual isolation in *Bacillus*. Genetics 134:402-408.

Sherley M, Gordon DM, Collignon PJ (2004) Evolution of multi-resistance plasmids in Australian clinical isolates of *Escherichia coli*. Microbiology 150:1539-1546.

Shirey JJ, Bissonnette GK (1992) Sheen formation and growth-response of groundwater bacteria to reduced oxygen concentrations during incubation of m-endo medium. Canadian Journal of Microbiology 38:261-266.

Souza V, Nguyen TT, Hudson RR, Pinero D, Lenski RE (1992) Hierarchical analysis of linkage disequilibrium in rhizobium populations - evidence for sex. Proceedings of the National Academy of Sciences of the United States of America 89:8389-8393.

Steigerwalt AG, Fanning GR, Fife-Asbury MA, Brenner DJ (1976) DNA relatedness among species of *Enterobacter* and *Serratia*. Canadian Journal of Microbiology 22:121-137.

Tornadijo ME, Garcia MC, Fresno JM, Carballo J (2001) Study of Enterobacteriaceae during the manufacture and ripening of San Simon cheese. Food Microbiology 18:499-509.

Vandamme P, Pot B, Gillis M, Devos P, Kersters K, Swings J (1996) Polyphasic taxonomy, a consensus approach to bacterial systematics. Microbiological Reviews 60:407-438.

Vanvalen L (1976) Ecological species, multispecies, and oaks. Taxon 25:233-239.

Vulic M, Dionisio F, Taddei F, Radman M (1997) Molecular keys to speciation - DNA polymorphism and the control of genetic exchange in enterobacteria. Proceedings

of the National Academy of Sciences of the United States of America 94:9763-9767.

Wayne LG, Brenner DJ, Colwell RR, Grimont PAD, Kandler O, Krichevsky MI, Moore LH, Moore WEC, Murray RGE, Stackebrandt E, Starr MP, Truper HG (1987) Report of the ad hoc committee on reconciliation of approaches to bacterial systematics. International Journal of Systematic Bacteriology 37:463-464.

Chapter 3

The frequency and diversity of bacteriocins in *Hafnia alvei*

Introduction

Bacteria have a wide range of defence mechanisms. Antimicrobial agents such as microcins and antibiotics kill or suppress the growth of other microorganisms so that the producing cell can gain a competitive advantage in the environment where it finds itself. Bacteriocins are a class of antimicrobial compounds produced by almost all known bacterial species. They have a typical phenotypic characteristic in that their toxicity range is very narrow and is mostly restricted to within-species (Reeves 1972; Riley *et al.* 2003). The most well studied of bacteriocins are colicins produced by *Escherichia coli.* Colicins are proteins encoded on plasmids (Pugsley 1987). The gene cluster typically consists of the colicin, immunity, and lysis gene (Pugsley 1984). The immunity gene codes for a protein that binds specifically to the colicin that is being produced in the cell and inactivates the colicin in order to protect the cell. Upon stress, a small fraction of the cells in a population will produce and then release colicin by lysing (Mulec *et al.* 2003).

There is ample evidence for colicins playing an important role in intra-species population dynamics. First, there is a high frequency of colicin production in a given population; on average 30 % of strains produce one or more colicins (Pugsley 1984; Riley and Gordon 1996). Second, the diversity of colicins is considerable. Over 20 types of colicin gene clusters have been characterised from various *E. coli* populations (Pugsley 1984; Pugsley 1987; Riley *et al.* 2000). Third, resistance of *E. coli* to colicins is strikingly high. Most *E. coli* in the natural environment are resistant to most co-occurring colicins (Feldgarden and Riley 1998; Riley and Gordon 1992). In addition, there is theoretical and experimental evidence to support the ecological significance in the levels of colicin production in natural populations of *E. coli* (Czaran *et al.* 2002;

Frank 1994; Gordon *et al.* 1998; Kerr *et al.* 2002; Kirkup and Riley 2004; Riley and Gordon 1996, 1999). However, we do not know if the conceptual framework developed for colicins is a paradigm for the other numerous bacteriocins found in nature. There is growing interest in using bacteriocins as alternatives for antibiotics and food preservatives. Further knowledge of other bacteriocin systems is required to establish safe guidelines for exploiting these antimicrobial compounds (Cleveland *et al.* 2001; Kuipers *et al.* 2000; Riley and Wertz 2002).

Hafnia alvei is a member of the *Enterobacteriaceae*. *H. alvei* is known to produce bacteriocins, however there is limited information on any other aspect of these bacteriocins (Reeves 1972; Wertz and Riley 2004). *H. alvei* is reported frequently as a food contaminant in dairy, fish and meat products, and the effective use of bacteriocins produced by this species may decrease contamination levels of *H. alvei* in food (Gonzalez-Rodreguez *et al.* 2001; Lindberg *et al.* 1998; Tornadijo *et al.* 2001). The purpose of this study was to investigate the frequency and diversity of bacteriocin production and resistance in *H. alvei* and to contrast the results to those found for colicin production in *E. coli.*

Material and Methods

Strains

One hundred and eighty eight *H. alvei* strains were isolated from faecal swabs taken from a variety of vertebrate hosts as well as from water column samples. Techniques used to isolate the strains are described by Gordon and FitzGibbon (1999). Biochemical and multilocus enzyme electrophoresis profiles of the isolates used in this study are presented by Okada and Gordon (2003).

Detection of bacteriocin production

All 188 *H. alvei* strains were assayed for bacteriocin production using the method described by Gordon *et al.* (1998). A 1 ml aliquot of an overnight Luria Bertani (LB) culture of a *H. alvei* strain was transferred to 9 ml of fresh LB broth and incubated at 37 °C with shaking for 30 minutes. Subsequently, mitomycin C was added to the culture to produce a final concentration of 0.1 μg ml^{-1}. The culture was incubated for a further 4 hours, and chloroform was added to lyse bacterial cells, from which 1.5 ml of culture was removed and centrifuged at 14260 *g* for 5 minutes. The supernatant was then transferred to a microfuge tube containing 50 μl chloroform, vortexed vigorously and centrifuged (14260 *g*, 2 minutes). The crude bacteriocin extract was stored at 4 °C for up to a week until use. Bacterial lawns were prepared by overlaying a LB agar plate with 3 ml of LB soft agar containing 100 μl of an overnight culture of the indicator strain. Crude bacteriocin extracts were spotted onto the bacterial lawn, dried and incubated overnight at 37 °C, after which killing was detected by a clear zone on the lawn where the extract was applied. The degree of killing was scored on a five-point scale between 0 and 4, 0 representing no killing, and 4 corresponding to a completely clear zone with no bacterial growth. Most extracts were checked for bacteriocin activity against indicator strains that came from the same host group, i.e. all extracts of *H. alvei* isolated from mammals were tested against all mammalian isolates. This procedure was repeated at least twice for all strains that had crude extracts with an antibacterial effect to confirm the presence of a substance that inhibits growth of lawn strains.

To test whether the killing of the crude extract was due to bacteriocin or bacteriophage activity, subsamples of the crude extracts were digested with trypsin overnight at 37 °C or stored at - 70 °C overnight. Trypsin digests most bacteriocins and – 70 °C kills most bacteriophage. These treatments had no effect on some extracts, and these extracts were then passed through a Microcon 100 000 Da microconcentrator (Amicon). This procedure eliminates most bacteriophage from the extract, and therefore if the filtered

extract produces a clear zone on a sensitive lawn the activity is bacteriocinogenic (Gordon *et al.* 1998).

Bacteriocin typing and resistance profile

The efficacy of all bacteriocin-producing strains was assessed by spotting crude extracts from all producer strains onto different lawns, each containing one of the bacteriocin-producing strains. Identical bacteriocin genotypes have the following two physiological properties: (1) similar killing patterns when tested against a set of lawn strains, and (2) do not kill strains that produce the same bacteriocin type due to the presence of the immunity gene. A killing matrix was then organised to group bacteriocins that had similar killing patterns.

Resistance against each bacteriocin was calculated as the proportion of the 38 lawns that were not killed by a particular bacteriocin.

Results

Bacteriocin production in *Hafnia alvei*

Of the188 *H. alvei* strains, 38 produced bacteriocin. *H. alvei* has been shown to consist of two genetic groups, which are arbitrarily designated as genetic groups 1 and 2 (Janda *et al.* 2002; Okada and Gordon 2003). Isolates belonging to genetic group 2 had a significantly higher fraction of bacteriocin production than those of group 1 (Table 1, nominal logistic regression, $\chi^2_{(1)} = 39.1$, $p < 0.01$). Isolates from reptiles had the highest percentage of bacteriocin production, whilst isolates from fish, despite the fact that the majority belonged to genetic group 2, had the lowest frequency of production (Table 1, nominal logistic regression, $\chi^2_{(3)} = 22.6$, $p < 0.01$).

Table 1. Percentage of bacteriocin production in *Hafnia alvei*. Numbers in brackets indicate the number of strains in each host by genetic group.

Host Group	Genetic Group	
	1	2
Overall	4.5 (88)	34 (100)
Fish	0.0 (5)	4.0 (24)
Reptiles	0.0 (27)	64.0 (28)
Mammals	5.1 (39)	29.0 (42)
Birds	14.0 (7)	50.0 (6)
Others	10.0 (10)	-

Bacteriocin typing in *H. alvei*

Patterns of the killing scores resulting from testing the 38 bacteriocins against lawns of the 38 bacteriocin-producing strains revealed that the bacteriocins cluster into three major groups based on the combination of lawn strains that a particular bacteriocin would kill (Figure 1). The bacteriocin group that killed the majority of lawns consisted of genetic group 2 strains isolated from reptiles (Figure 1, phenotypes 1 to 7). This group of bacteriocins was further split into at least seven different bacteriocin types, as there were bacteriocins that killed other bacteriocinogenic lawns within the cluster. The next cluster of bacteriocins consisted of genetic group 2 strains isolated from mammals and birds. With the exception of the two strains isolated from birds (Figure 1, phenotype 8), all other bacteriocins were likely to be of a single type (Figure 1, phenotype 9). This result is supported by the fact that five of the bacteriocin plasmids from this cluster have been sequenced and shown to be genetically identical (Wertz and Riley, 2004). There were four unique bacteriocin producers that were isolated from fish (Figure 1, phenotype 10) and mammals (Figure 1, phenotypes 11, 12, 13). We could not further characterise the remainder of bacteriocins as they killed only a minority of the lawns they were tested against (Figure 1, phenotype 14). This analysis suggests that, in total, there are a least 14 different types of bacteriocins being produced in this collection of strains.

Figure 1. The killing matrix of *Hafnia alvei* bacteriocins. Killing intensity is in the scale of 1 (+) to 4 (++++). Cells with empty scores denote absence of killing.

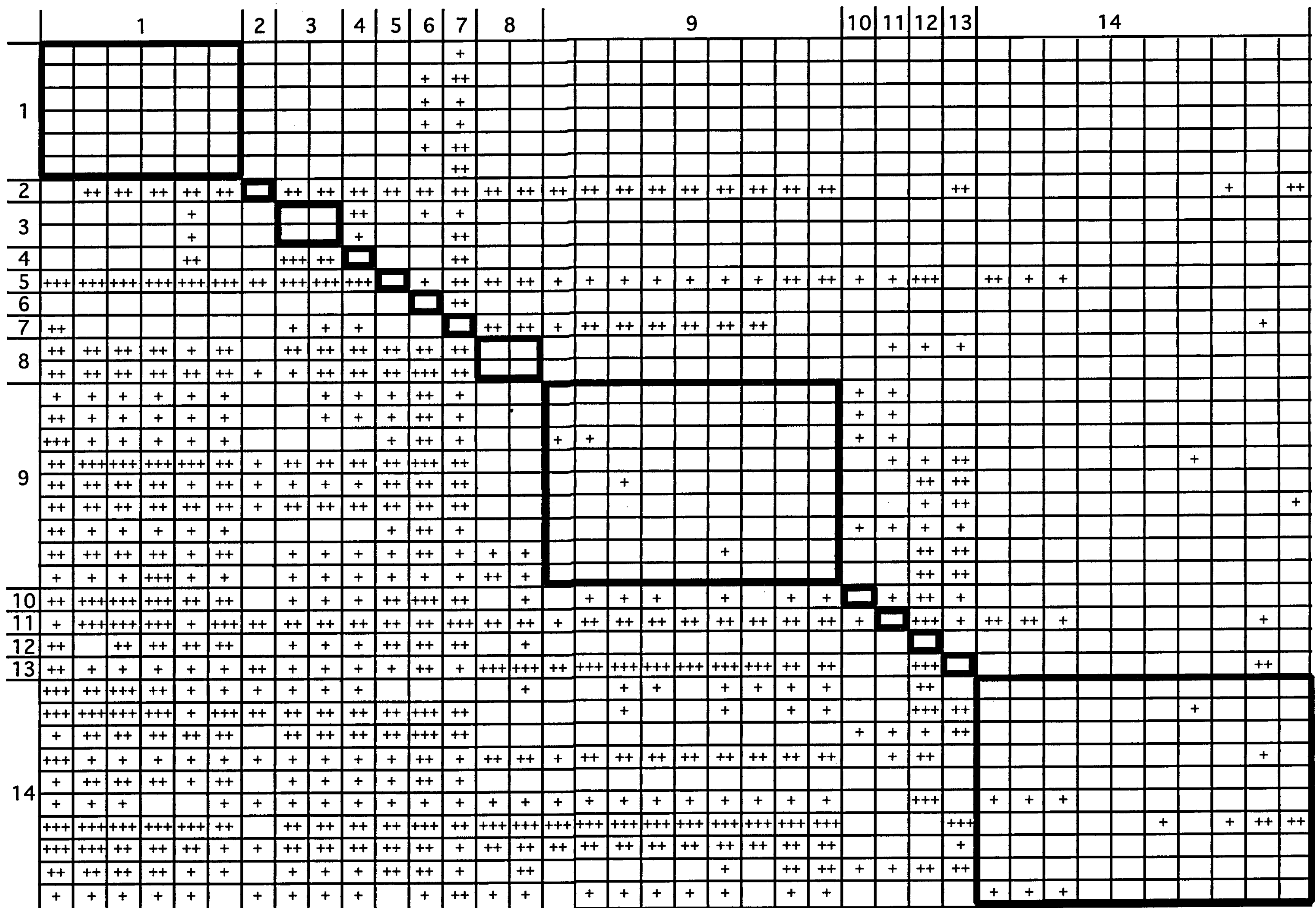


Figure 1.

Resistance profile of *H. alvei* to bacteriocin

The proportion of the 38 strains that were resistant to the 38 bacteriocins ranged from 5.5 % to 100 % (Figure 2). On average 61 % of the strains were resistant to a particular bacteriocin. A significantly greater fraction of strains (88 %) were resistant to the bacteriocins produced by genetic group 1 isolates compared to the fraction (58 %) of strains resistant to the bacteriocins produced by genetic group 2 strains (Wilcoxon test, $\chi^2_{(1)} = 4.3$, $p < 0.05$).

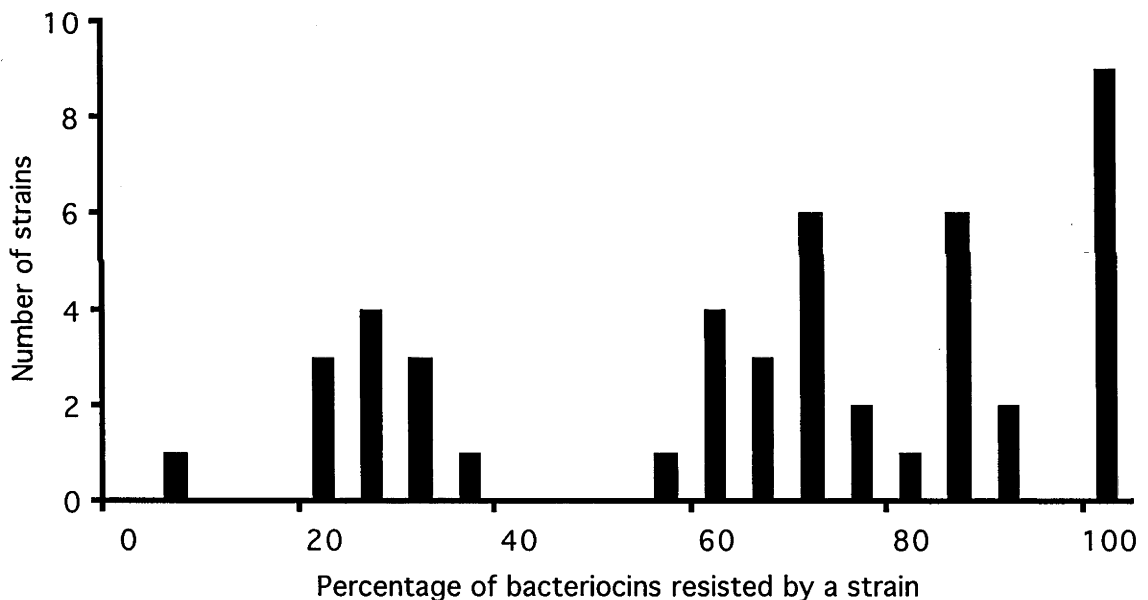


Figure 2. Resistance of *Hafnia alvei* strains to the bacteriocins produced by *H. alvei.* The X axis represents the fraction (%) of the 46 bacteriocin producing strains that the bacteriocin extracts failed to kill.

Discussion

The aim of this study was to observe phenotypic variation in bacteriocin production of *H. alvei* and assess its ecological significance in terms of the intra-specific competition theory by comparing the data to what is known from studies looking at colicin biology and dynamics.

The frequency of colicin production in different *E. coli* populations ranges between 16 and 60 %, with an overall average of 30 % (Gordon *et al.* 1998; Pugsley 1984; Riley and Gordon 1992). The frequency of bacteriocin production in *H. alvei* found in this study (20 %) is comparable to that found for colicin production.

This study found at least 14 different bacteriocin phenotypes in *H. alvei.* While the total number of colicin types described exceeds 20, the number of different colicin types found in a given *E. coli* population is far less, ranging between three and nine (Riley and Gordon 1996). Most of these studies have only examined isolates from a single host species (usually from humans), whilst this study included isolates from four different vertebrate groups. Therefore it appears that the diversity of bacteriocins in *H. alvei* is comparable to that found for colicins in *E. coli.*

The high resistance of *H. alvei* to its bacteriocins is very similar to observations of colicin resistance in different populations of *E. coli* (Feldgarden and Riley 1998; Gordon *et al.* 1998). The higher level of resistance to bacteriocins produced by genetic group 1 isolates compared to those of genetic group 2 isolates may be due to the different titres of bacteriocin being produced by each strain in these groups. Gordon *et al.* (1998) found that titre explained a significant amount of the variation in resistance of *E. coli* to colicin.

The prevalence, diversity and resistance patterns of the bacteriocins in this study were all influenced to an extent by the genetic background of the strains and the host from which they were recovered. The ecological significance of these patterns is unknown. Growth rate studies have shown that there is a fitness cost associated with colicin production (Chao and Levin 1981). Theoretical (Frank 1994; Kerr *et al.* 2002) and empirical (Chao and Levin 1981) studies have shown that colicin-producing cells have a competitive advantage over colicin-sensitive cells in structured, nutrient-rich environments, whilst sensitive cells are favoured in unstructured, nutrient-poor habitats. These outcomes indicate that the frequency of bacteriocin production in a community of strains should vary depending on the nature of the environment. However, the exact environmental factors that contribute to the seven-fold difference in bacteriocin production between genetic group 2 *H. alvei* isolates from reptiles and fish are unknown.

In summary, the prevalence of bacteriocinogeny in *H. alvei*, the diversity of bacteriocins produced, and the levels of resistance to these bacteriocins are similar to the patterns found for colicins in *E. coli*. These results support the idea that our understanding of the evolution and ecology of colicins will serve as a paradigm for the phenomenon of bacteriocinogeny in the *Enterobacteriaceae*. However, more empirical and theoretical work is required in order to understand the reasons for the frequency of bacteriocin production depending on the type of hosts from which the strains originate. Further work is also required to determine the molecular mechanisms responsible for the different bacteriocin types and resistance patterns found in *H. alvei* strains of the two genetic groups.

References

Chao L, Levin BR (1981) Structured habitats and the evolution of anticompetitor toxins in bacteria. Proceedings of the National Academy of Sciences of the United States of America 78:6324-6328.

Cleveland J, Montville TJ, Nes IF, Chikindas ML (2001) Bacteriocins: safe, natural antimicrobials for food preservation. International Journal of Food Microbiology 71:1-20.

Czaran TL, Hoekstra RF, Pagie L (2002) Chemical warfare between microbes promotes biodiversity. Proceedings of the National Academy of Sciences of the United States of America 99:786-790.

Feldgarden M, Riley MA (1998) High levels of colicin resistance in *Escherichia coli*. Evolution 52:1270-1276.

Frank SA (1994) Spatial polymorphism of bacteriocins and other allelopathic traits. Evolutionary Ecology 8:369-386.

Gonzalez-Rodriguez MN, Sanz JJ, Santos JA, Otero A, Garcia-Lopez ML (2001) Bacteriological quality of aquacultured freshwater fish portions in prepackaged trays stored at 3 degrees C. Journal of Food Protection 64:1399-1404.

Gordon DM, FitzGibbon F (1999) The distribution of enteric bacteria from Australian mammals: host and geographical effects. Microbiology 145:2663-2671.

Gordon DM, Riley MA, Pinou T (1998) Temporal changes in the frequency of colicinogeny in *Escherichia coli* from house mice. Microbiology-UK 144:2233-2240.

Janda JM, Abbott SL, Khashe S, Probert W (2002) Phenotypic and genotypic properties of the genus *Hafnia*. Journal of Medical Microbiology 51:575-580.

Kerr B, Riley MA, Feldman MW, Bohannan BJM (2002) Local dispersal promotes biodiversity in a real-life game of rock-paper-scissors. Nature 418:171-174.

Kirkup BC, Riley MA (2004) Antibiotic-mediated antagonism leads to a bacterial game of rock-paper-scissors *in vivo*. Nature 428:412-414.

Kuipers OP, Buist G, Kok J (2000) Current strategies for improving food bacteria. Research in Microbiology 151:815-822.

Lindberg AM, Ljungh A, Ahrne S, Lofdahl S, Molin G (1998) Enterobacteriaceae found in high numbers in fish, minced meat and pasteurised milk or cream and the presence of toxin encoding genes. International Journal of Food Microbiology 39:11-17.

Mulec J, Podlesek Z, Mrak P, Kopitar A, Ihan A, Zgur-Bertok D (2003) A *cka-gfp* transcriptional fusion reveals that the colicin K activity gene is induced in only 3 percent of the population. Journal of Bacteriology 185:654-659.

Okada S, Gordon DM (2003) Genetic and ecological structure of *Hafnia alvei* in Australia. Systematic and Applied Microbiology 26:585-594.

Pugsley AP (1984) The ins and outs of colicins .1. Production, and translocation across membranes. Microbiological Sciences 1:168-175.

Pugsley AP (1987) Nucleotide sequencing of the structural gene for colicin N reveals homology between the catalytic, C-terminal domains of colicin A and colicin N. Molecular Microbiology 1:317-325.

Reeves P (1972) *The Bacteriocins*. Springer-Verlag, New York.

Riley MA, Cadavid L, Collett MS, Neely MN, Adams MD, Phillips CM, Neel JV, Friedman D (2000) The newly characterized colicin Y provides evidence of positive selection in pore-former colicin diversification. Microbiology-UK 146:1671-1677.

Riley MA, Goldstone CM, Wertz JE, Gordon DM (2003) A phylogenetic approach to assessing the targets of microbial warfare. Journal of Evolutionary Biology 16:690-697.

Riley MA, Gordon DM (1992) A survey of col plasmids in natural isolates of *Escherichia coli* and an investigation into the stability of col plasmid lineages. Journal of General Microbiology 138:1345-1352.

Riley MA, Gordon DM (1996) The ecology and evolution of bacteriocins. Journal of Industrial Microbiology 17:151-158.

Riley MA, Gordon DM (1999) The ecological role of bacteriocins in bacterial competition. Trends in Microbiology 7:129-133.

Riley MA, Wertz JE (2002) Bacteriocins: Evolution, ecology, and application. Annual Review of Microbiology 56:117-137.

Tornadijo ME, Garcia MC, Fresno JM, Carballo J (2001) Study of Enterobacteriaceae during the manufacture and ripening of San Simon cheese. Food Microbiology 18:499-509.

Wertz JE, Riley MA (2004) Chimeric nature of two plasmids of *Hafnia alvei* encoding the bacteriocins alveicins A and B. Journal of Bacteriology 186:1598-1605.

Chapter 4

Adaptive nature of the growth rate variation in the genus *Hafnia*

Introduction

The growth characteristics of a bacterial population is a property of fundamental adaptive significance. The doubling time can vary among species from as little as 10 minutes to more than several days (Madigan 1997). In addition to the among-species variation, population growth rates will vary with temperature, nutrient levels and types, pH, water availability and oxygen concentration.

The gastrointestinal tract is a dynamic environment where intra-and inter-specific competition will occur among established populations as well as with cells attempting to invade the habitat. Cells of microbial population are lost from the gut at a constant rate due to cell death or with the passage of digested food material. In order to maintain a viable population, the population growth rate must be sufficient to counteract these losses. Gut transit time, nutrient levels and food types will differ substantially depending on the host species in which the bacteria finds itself. Faunivores have shorter guts with a lower intestinal surface to body weight ratio compared to herbivores, which need a longer gut with more surface area to digest plant material (Stevens and Hume 1995). Food retention times will be substantially longer for animals that consume plants as opposed to those that eat animal material (Stevens and Hume 1995). Among animals with similar diet, those with a small body size have shorter food retention times compared to larger animals (Karasov *et al.* 1986). Reptiles have lower metabolic rates than mammals, and their food retention time can be ten times longer (Karasov *et al.* 1986). Fish have varying food retention times and retention times are known to be temperature dependent (Stevens and Hume 1995). For freshwater fish in particular, herbivorous species seem to have shorter food retention times compared to

carnivorous fish (Stevens and Hume 1995). Therefore one would expect the growth characteristics of a commensal bacterial species inhabiting the gut would vary with the gastrointestinal environment of the host.

Laboratory experiments demonstrate that enteric bacteria adapt and maximise their growth rate in response to the environment. *Escherichia coli* strains maintained in minimal glucose media at 37 °C for 20,000 generations increased their maximum growth rate at 37 °C compared to the ancestral strain (Cooper *et al.* 2001). An increase in maximum growth rate at 45 °C compared to the ancestral strain was also observed in *Pseudomonas pseudoalcaligenes* maintained at 45 °C for 10 months (Shi and Xia 2003).

The growth characteristics of *E. coli* strains isolated from different environments have been found to vary. Commensal *E. coli* isolated from humans had higher growth rates at temperatures > 26 °C compared to those isolated from a septic tank. On the other hand, the septic tank strains grew better at lower temperatures compared to the human strains (Gordon *et al.* 2002). A shift in optimal temperature for growth in different *E. coli* populations has also been detected. The optimal temperature for growth of *E. coli* isolated from turtles was 2 °C lower compared to *E. coli* isolated from a herbivorous mammal (Bronikowski *et al.* 2001).

Hafnia alvei (Enterobacteriaceae) is a gut commensal distantly related to *E. coli* (Brenner 1981). It is a species of some concern to public health as it occasionally causes nosocomial infections (Gunthard and Pennekamp 1996; Ramos and Damaso 2000), and is frequently found as a contaminant of fish, dairy and meat products (Gamage *et al.* 1998; Gaya *et al.* 1987; Gonzalez-Rodriguez *et al.* 2001; Morales *et al.* 2003). *H. alvei* can be isolated from a variety of vertebrate species and is most prevalent in ectotherms (Okada and Gordon 2003). This species is known to consist of two genetic groups with distinct biochemical characteristics (Brenner 1981; Janda *et al.*

2002; Okada and Gordon 2003Ridell *et al.* 1995; Steigerwalt *et al.* 1976). A previous study characterising the thermal niche of *H. alvei* showed that the two genetic groups have significantly different thermal niches (Okada and Gordon 2001).

This study examined the growth rate-temperature relations of 90 *H. alvei* isolates from fish, reptiles and mammals. The aims were two-fold; to detect any differences in the growth characteristics between the two genetic groups of *H. alvei*, and to determine if aspects of host taxonomy and/or the biochemical traits of the individual *H. alvei* strains explained any of the variation in their growth characteristics.

Material and Methods

Bacterial strains

Ninety strains of *H. alvei* were used in this study: 14 from fish, 46 from mammals, and 30 from reptiles. Isolation and biochemical identification procedures are described in Gordon and Fitzgibbon (1999). All strains were genotyped by allozyme variation and assigned to one of the two previously described genetic groups (Okada and Gordon 2003). Three temporary freezer cultures were prepared for each strain set by culturing the strains overnight in Luria Bertani broth. One ml of overnight culture was mixed with 30 μl of glycerol and stored at -70 °C until required.

Growth characteristics

The growth characteristics of the 90 *H. alvei* strains were determined as follows, there were three replicates of each strain and each strain/replicate combination was assigned at random to the wells of three 96 well microtitre plates. An inoculum of the appropriate freezer cultures were grown overnight at 30 °C in a microtitre plate containing minimal glucose medium (per 1L; K_2HPO_4, 7 g; KH_2PO_4, 2 g; $(NH_4)_2SO_4$, 1 g; Na citrate, 0.5 g; Glucose, 0.5 g; 10 % $MgSO_4$/0.2 % VB1, 1 ml). Twenty five μl of

the overnight cultures were transferred to 225 μl of fresh minimal glucose medium to produce a starting density of about 5×10^6 cells ml^{-1}, and incubated at 5, 10, 14, 18, 22, 25, 28, 31, 33, 35, 37, 39, or 41 °C. The optical density (OD) at 640 nm was taken in regular intervals (30 minutes for 5 and 10 °C, 20 minutes for 14 °C, 10 minutes for 18 °C, and 5 minutes for remaining temperatures) using a Powerwave X Microplate Scanning Spectrophotometer (Bio-tek Instruments, Inc.) until the cultures reached stationary phase. The plates were shaken for one minute prior to every reading.

Calculation of the growth rate parameters

To determine the growth rate of a strain at a given temperature, the average rate of change in the OD versus time curve was calculated as the range of points over which the OD reading was increasing. A non-linear model ($\varphi_T h^{-1} = (B - C \times T) \times D^T$) was then fitted to the growth rates calculated at each temperature, where $\varphi_T h^{-1}$ is the maximum growth rate at a temperature, and B, C, and D are the parameters determining the shape of the curve. From this model the following four variables were estimated for each strain:

1) Lethal temperature, the temperature at which the growth rate is zero, calculated by B/C

2) Maximum growth rate, calculated by C/ln(D)*exp[{B*ln(D)/C} - 1]

3) Optimal temperature, defined as the temperature at which the maximum growth rate occurs, calculated by B/C - 1/ln(D)

4) The inflection temperature, defined as the temperature where the growth rate stops increasing with increasing temperature, calculated by B/C - 2/ln(D).

Analysis of variance was used to assess if any of the variation in the three model parameters and four derived variables were explained by the genetic group of the strains, host taxonomy, and biochemical properties of the strains. To determine if the biochemical characteristics of the strains explained any of the variation in the growth

rate-temperature relationship a stepwise regression analysis approach was taken (Miller, 2002) after adjusting for the variation due to genetic groups.

Results

Effect of genetic group

The genetic group of the strains explained a significant amount of the variation for the three model parameters and four derived variables (Table 1). Genetic group 1 strains had significantly lower values for all parameters and variables. The resulting growth curves showed that group 1 strains, on average, had a lower maximum growth rate compared to group 2 strains and that their optimal temperature for growth and their lethal temperature were 1 °C lower compared to genetic group 2 strains (Figure 1).

Table 1. Parameter values and calculated growth variables of *Hafnia alvei* strains by genetic group.

	Genetic group		P > F
	1	2	
No. of strains	46	44	
B	0.019	0.017	< 0.01
C	0.000	0.000	< 0.01
D	1.137	1.141	< 0.05
Maximum growth rate	0.206	0.218	< 0.05
Optimal temperature	31.052	32.096	< 0.01
Lethal temperature	38.851	39.716	< 0.01
Inflection temperature	23.253	24.476	< 0.01

Effect of host taxonomy

Nested analysis of variance was conducted to detect the effect of host taxonomic rank on the growth rate - temperature relations of the strains. After accounting for genetic group, there was a significant effect of host taxonomic class on the maximum growth rate (genetic group, $F_{(1,1)} = 6.01$, $P < 0.02$; host class within genetic group, $F_{(4,4)} = 3.86$, $P < 0.01$). The two genetic groups were then assessed separately for the effects of host

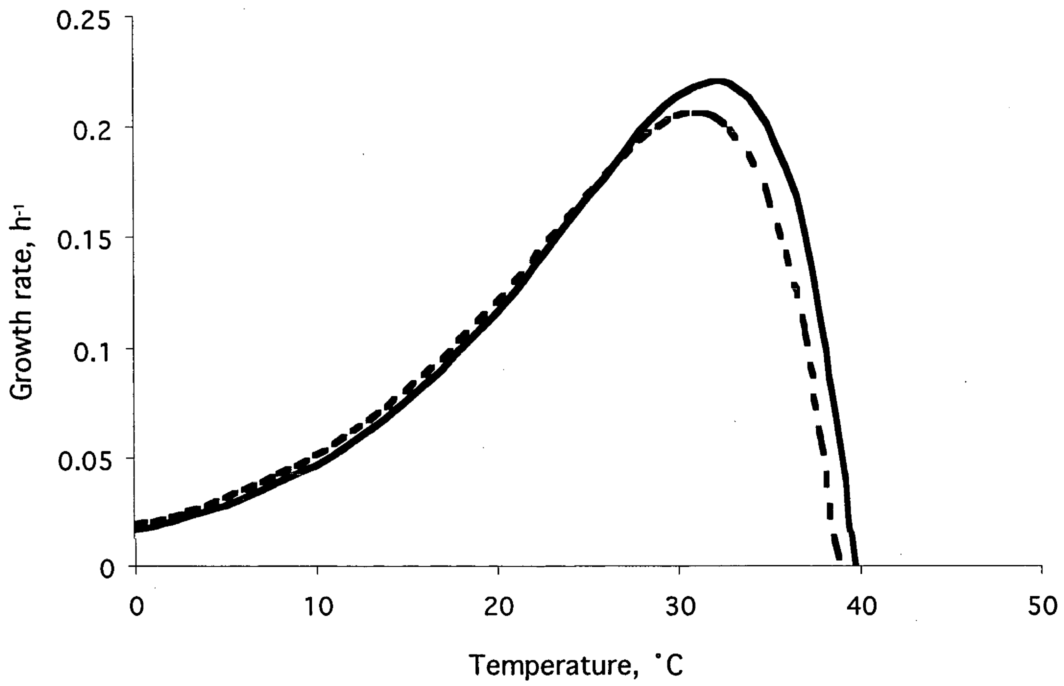


Figure 1. Growth curves of *Hafnia alvei* strains by genetic group. Dashed line, group 1; solid line, group 2.

taxonomic class. Host class explained a significant amount of the variation in the maximum growth rate for genetic group 2 strains, but not for those of group 1 (Table 2).

Table 2. Parameter values and calculated growth variables of *Hafnia alvei* strains by genetic group and host taxonomic class. F values for analysis of variance for the effect of host taxonomic class on the variables and significance is also shown. Max GR, maximum growth rate; Opt T, optimal temperature; Lethal T, lethal temperature; Inf T, inflection temperature.

Genetic group	Host class	N	B	C	D	Max GR	Opt T	Lethal T	Inf T
1	Fish	5	0.0212	0.0005	1.1331	0.21	30.99	39.00	22.97
	Mammal	24	0.0186	0.0005	1.1382	0.21	31.06	38.82	23.31
	Reptile	17	0.0188	0.0005	1.1371	0.20	31.06	38.86	23.26
	$F_{(2,43)}$		2.58	2.71	0.91	0.14	0.04	0.16	0.38
	P > F		0.09	0.08	0.41	0.87	0.96	0.86	0.69
2	Fish	9	0.0175	0.0004	1.1390	0.22	31.92	39.62	24.21
	Mammal	22	0.0170	0.0004	1.1422	0.23	32.23	39.78	24.68
	Reptile	13	0.0158	0.0004	1.1393	0.20	31.99	39.67	24.31
	$F_{(2,41)}$		1.49	1.88	0.70	7.08	0.77	0.11	1.91
	P > F		0.24	0.17	0.50	0.01	0.47	0.90	0.16

In genetic group 2, strains from reptiles had a lower maximum growth rate than those from fish and mammals (Figure 2, Table 2). After accounting for host taxonomic class and order, family had a significant effect on the variation in the optimal and inflection temperatures for group 2 strains (optimal temperature: class, $F_{(2,2)} = 1.69$, $P > 0.2$, order within class, $F_{(1,1)} = 1.30$, $P > 0.2$, family within order, $F_{(1,1)} = 12.27$, $P < 0.01$; inflection temperature: class, $F_{(2,2)} = 3.24$, $P > 0.05$, order within class, $F_{(1,1)} = 0.77$, $P > 0.3$, family within order, $F_{(1,1)} = 10.72$, $P < 0.01$). Strains from the teleost family Terapontidae (n = 5) had optimal and inflection temperatures that were 1 °C higher than those from the teleost family Percichthyidae (n = 3), which had the lowest optimal and inflection temperatures.

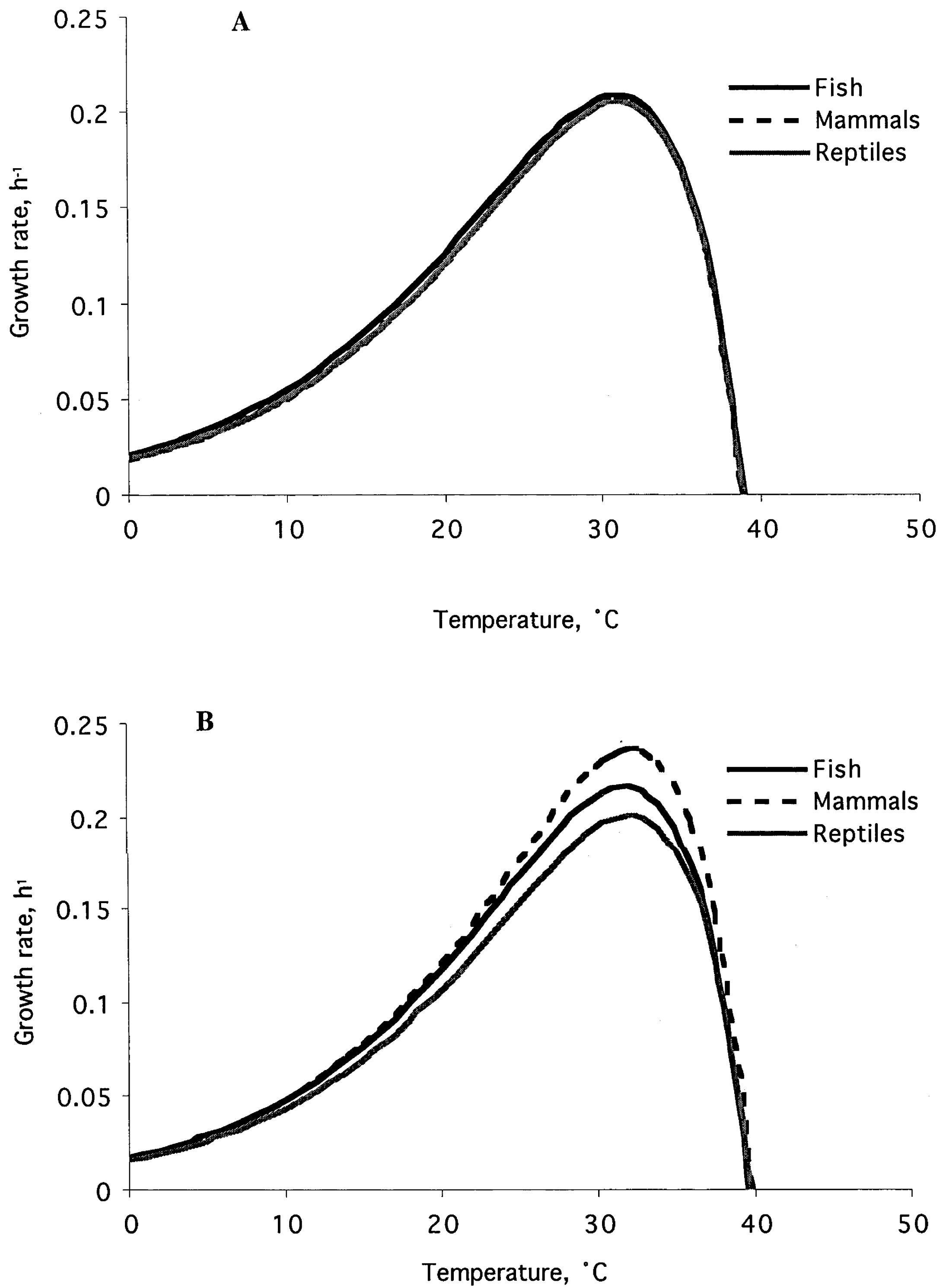


Figure 2. Growth curves of *Hafnia alvei* strains by host taxonomic class. A, genetic group 1; B, genetic group 2. Solid line, fish isolates; dashed lines, mammalian isolates; unevenly dashed lines, reptilian strains.

Effect of biochemical characteristics

A stepwise regression approach was taken to determine if any of the variation in the phenotypic profiles of the strains was responsible for the observed variation in the growth parameters. After accounting for genetic group and host taxonomic class, the ability of a strain to utilise melibiose as a carbon source, the ability to reduce tetrazolium, and the production of acetoin during aerobic glucose fermentation (assessed by the Voges-Proskauer Test) were found to explain a significant amount of the variation in some of the parameters and derived variables. Strains that were able to utilise melibiose had lower lethal temperatures, optimal and inflection temperatures and maximum growth rates in both genetic groups (Figure 3, Table 3).

Table 3. Multiple regression analysis. Effects of genetic group, melibiose utilisation, tetrazolium reduction and acetoin production on the three growth variables are presented as F values. TTC, tetrazolium.

	DF	F	P > F
Lethal temperature			
Genetic group	1	31.54	< 0.01
Melibiose utilisation	1	8.16	< 0.01
TTC reduction	1	8.39	< 0.01
Maximum growth rate			
Genetic group	1	7.55	< 0.01
Melibiose utilisation	1	33.9	< 0.01
Acetoin production	1	4.98	< 0.03
Inflection temperature			
Genetic group	1	79.87	< 0.01
Melibiose utilisation	1	23.18	< 0.01
Acetoin production	1	4.73	< 0.04

Strains that produced acetoin during fermentation of glucose had increased maximum growth rates and inflection temperature in both genetic groups (Table 3). Strains capable of reducing tetrazolium were more tolerant to high temperatures than those that could not (Table 3).

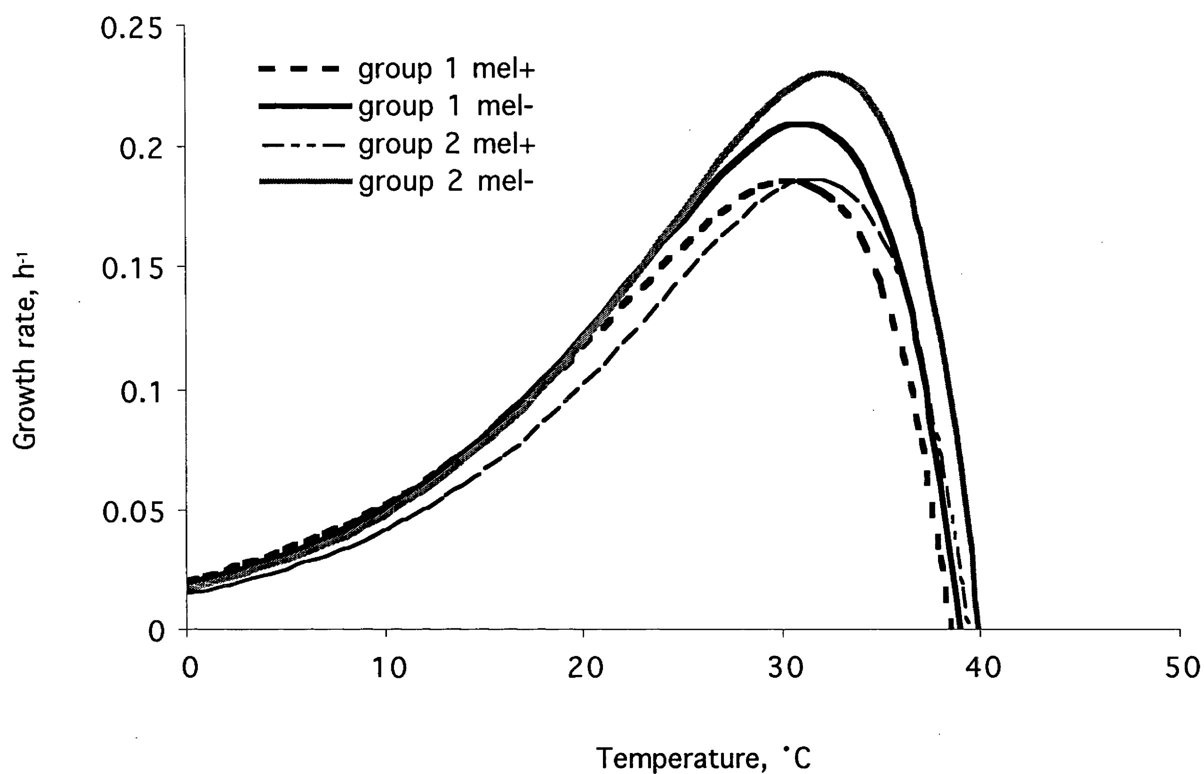


Figure 3. Growth curves of *Hafnia alvei* strains by genetic group (group 1 and 2) and their ability to utilise melibiose as a carbon source. Grey line, group 2 melibiose-negative strains; solid line, group 1 melibiose-negative strains; thin-dashed line, group 2 melibiose-positive strains; thick-dashed line, group 1 melibiose-positive strains.

Discussion

Adaptation of *H. alvei* group 2 strains to its host but not in group 1

Enteric bacteria (represented by *Escherichia coli*) cycle between two environments, the gastrointestinal tract of the host and environments external to the host such as soil, water and sediment (Savageau 1983). The two environments differ markedly in their biotic and abiotic characteristics such as the microbial community structure, nutrient availability and physical conditions. There is empirical evidence to suggest that the two environments harbour populations with different genetic backgrounds within a species, in other words there is ecological structure (Cohan 1994). For example, Whittam (1989) found that the *E. coli* population in the gastrointestinal tract of domestic birds were genetically distinct to ones found in their litter, water and soil in the shed. Gordon *et al.* (2002) found the *E. coli* population in a septic tank to be genetically different from those isolated from faeces of the humans that had produced the inputs into the septic tank. The lack of evidence for adaptation of genetic group 1 strains to their host in terms of growth characteristics is also correlated with lack of host effects on their genetic structure (based on multilocus sequence typing of six housekeeping genes; Okada, chapter 6). These results suggest that the host digestive tract may not be the primary habitat for group 1 strains. Rather, group 1 strains may prefer the external environment of the host as their main habitat.

The difference between the two genetic groups in their optimal temperature is also highly correlated with the results from our previous study investigating the thermal niche of six enteric bacterial species including *H. alvei* (Okada and Gordon 2001). Five genetic group 1 strains and nine genetic group 2 strains were subjected to serial dilution culturing for up to five days in minimal glucose medium for five days at temperatures between 7 and 42 °C, and the change in population density was monitored daily during the culturing period. Group 1 strains on average had an upper thermal niche limit for

growth that was 1 °C lower than that of group 2 strains. This is in agreement with the 1 °C downshift in the optimal temperature of group 1 strains compared to those of group 2. However, Okada and Gordon (2001) also detected a difference in the lower thermal niche limits between the two genetic groups of *H. alvei*. The population growth rate of group 1 strains was zero at temperatures below 16 °C, whereas the population growth rate of group 2 strains only stopped at temperatures below 11 °C. In this study we could not detect differences in the growth characteristics at the lower temperatures between the two genetic groups of *H. alvei* that would explain the difference that was observed in serial dilution culturing (Okada and Gordon 2001). During the five-day serial dilution culturing, the population growth rate over time is a function of the growth rate, cell death, and loss of cells due to daily dilution. The growth rate of the two genetic groups of *H. alvei* may initially be similar for lower temperature ranges (as was found in this study), but then group 2 strains may gradually decrease their growth rate over the course of five days of culturing in temperatures lower than 16 °C. On the other hand, group 1 strains may be capable of maintaining their maximum growth rate for five days in temperatures higher than 11 °C.

Correlation of some growth parameters to biochemical characteristics

It is not surprising to see certain characteristics of cellular metabolism correlating with the characteristics of reproductive physiology in bacteria. As mentioned in the introduction, growth characteristics can be influenced by factors such as temperature, amount of available nutrients and chemical toxicity.

In *E. coli*, melibiose utilisation is known to be temperature dependent; the melibiose transporter is activated at 30 °C but not at 37°C (Prestidge and Pardee 1965). The activation of the melibiose transporter at lower temperatures is controlled at the genetic level, by genes upstream of the transporter gene (Tamai *et al.* 1998). In *H. alvei* there may be a subtle but significant metabolic shift for more efficient sugar catabolism of

slightly lower temperatures for strains that can utilise melibiose, which may result in a lower optimal temperature for growth. The reasons for the decrease in maximum growth rate of strains that can utilise melibiose are unknown.

The increase in the maximum growth rate of *H. alvei* strains that produced acetoin may be due to their ability to degrade acetate into acetoin and hence escape the toxic effect of acetate. *E. coli* strains can not reduce acetate, and its intracellular accumulation results in a decreased growth rate (Luli and Strohl 1990). *E. coli* strains that produce less acetate had higher growth rates compared to strains that produced acetate in higher concentrations (Luli and Strohl 1990). Genetic engineering of *E. coli* with the *Bacillus subtilis* acetolactate synthase gene (the enzyme responsible for acetate reduction into acetoin) resulted in a higher cell yield under controlled pH conditions compared to the strain that was not genetically modified (Aristidou *et al.* 1995). *H. alvei* that produce acetoin (indicated by the positive response of the Voges - Proskauer test) may reduce acetate accumulation, which in turn will increase the growth rate relative to a strain not capable of reducing acetate.

The mechanism by which the ability to reduce tetrazolium leads to increased high temperature tolerance is unclear. This observed correlation between metabolistic and growth characteristics indicates that differences in the biochemical profiles between the two genetic groups of *H. alvei* has a significant effect on their rate of cell division.

Future directions

It would be of interest to further investigate the adaptive nature of the growth rate variation in the genetic group 1 strains of *H. alvei*. Variation in growth characteristics of bacteria is known to be influenced primarily by the efficiency of nutrient uptake and metabolism. Investigation of the glucose and melibiose utilisation cascade may reveal some of the molecular pathways and genes responsible for such adaptation.

Investigating the phenomenon of the 'viable but not culturable' state induction and resuscitation may shed further light on the preferred habitat of these two groups of *Hafnia*.

References

Aristidou AA, San KY, Bennett GN (1995) Metabolic engineering of *Escherichia coli* to enhance recombinant protein production through acetate reduction. Biotechnology Progress 11:475-478.

Brenner DJ (1981) Chapter 88. Introduction to the family Enterobacteriaceae. In: Starr MP, Stolp H, Trueper HG, Balows A, Schlegel HG (ed) *The Prokaryotes.* Springer-Verlag, Berlin, pp. 1105-1127.

Bronikowski AM, Bennett AF, Lenski RE (2001) Evolutionary adaptation to temperature. VIII. Effects of temperature on growth rate in natural isolates of *Escherichia coli* and *Salmonella enterica* from different thermal environments. Evolution 55:33-40.

Cohan FM (1994) Genetic exchange and evolutionary divergence in prokaryotes. Trends in Ecology and Evolution 9:175-180.

Cooper VS, Bennett AF, Lenski RE (2001) Evolution of thermal dependence of growth rate of *Escherichia coli* populations during 20,000 generations in a constant environment. Evolution 55:889-896.

Gamage SD, Luchansky JB, Ingham SC (1998) Pulsed-field gel electrophoresis typing of *Hafnia alvei* isolated from chub-packed and retail ground beef. Letters in Applied Microbiology 26:105-109.

Gaya P, Medina M, Nunez M (1987) Enterobacteriaceae, coliforms, fecal coliforms and *Salmonella*s in raw ewes milk. Journal of Applied Bacteriology 62:321-326.

Gonzalez-Rodriguez MN, Sanz JJ, Santos JA, Otero A, Garcia-Lopez ML (2001) Bacteriological quality of aquacultured freshwater fish portions in prepackaged trays stored at 3 degrees C. Journal of Food Protection 64:1399-1404.

Gordon DM, Bauer S, Johnson JR (2002) The genetic structure of *Escherichia coli* populations in primary and secondary habitats. Microbiology 148:1513-1522.

Gordon DM, FitzGibbon F (1999) The distribution of enteric bacteria from Australian mammals: host and geographical effects. Microbiology 145:2663-2671.

Gunthard H, Pennekamp A (1996) Clinical significance of extraintestinal *Hafnia alvei* isolates from 61 patients and review of the literature. Clinical Infectious Diseases 22:1040-1045.

Janda JM, Abbott SL, Khashe S, Probert W (2002) Phenotypic and genotypic properties of the genus *Hafnia*. Journal of Medical Microbiology 51:575-580.

Karasov WH, Petrossian E, Rosenberg L, Diamond JM (1986) How do food passage rate and assimilation differ between herbivorous lizards and nonruminant mammals? Journal of Comparative Physiology B-Biochemical Systemic and Environmental Physiology 156:599-609.

Luli GW, Strohl WR (1990) Comparison of growth, acetate production, and acetate inhibition of *Escherichia coli* strains in batch and fed-batch fermentations. Applied and Environmental Microbiology 56:1004-1011.

Madigan MT, Martinko JM, Parker J (1997) *Brock Biology of Microorganisms.* Prentice Hall, Upper Saddle River, NJ.

Morales P, Fernandez-Garcia E, Nunez M (2003) Caseinolysis in cheese by Enterobacteriaceae strains of dairy origin. Letters in Applied Microbiology 37:410-414.

Okada S, Gordon DM (2001) Host and geographical factors influence the thermal niche of enteric bacteria isolated from native Australian mammals. Molecular Ecology 10:2499-2513.

Okada S, Gordon DM (2003) Genetic and ecological structure of *Hafnia alvei* in Australia. Systematic and Applied Microbiology 26:585-594.

Prestidge LS, Pardee AB (1965) A second permease for methyl-thio-beta-D-galactoside in *Escherichia coli*. Biochimica Et Biophysica Acta 100:591-593.

Ramos A, Damaso D (2000) Extraintestinal infection due to *Hafnia alvei*. European Journal of Clinical Microbiology and Infectious Diseases 19:708-710.

Ridell J, Siitonen A, Paulin L, Lindroos O, Korkeala H, Albert MJ (1995) Characterization of *Hafnia alvei* by biochemical tests, random amplified polymorphic DNA PCR, and partial sequencing of 16S rRNA gene. Journal of Clinical Microbiology 33:2372-2376.

Savageau MA (1983) *Escherichia coli* habitats, cell types, and molecular mechanisms of gene control. American Naturalist 122:732-744.

Shi BH, Xia XH (2003) Changes in growth parameters of *Pseudomonas pseudoalcaligenes* after ten months culturing at increasing temperature. FEMS Microbiology Ecology 45:127-134.

Steigerwalt AG, Fanning GR, Fife-Asbury MA, Brenner DJ (1976) DNA relatedness among species of *Enterobacter* and *Serratia*. Canadian Journal of Microbiology 22:121-137.

Stevens CE, Hume D (1995) *Comparative Physiology of the Vertebrate Digestive System.* Cambridge University Press, New York.

Tamai E, Shimamoto T, Tsuda M, Mizushima T, Tsuchiya T (1998) Conversion of temperature-sensitive to -resistant gene expression due to mutations in the promoter region of the melibiose operon in *Escherichia coli*. Journal of Biological Chemistry 273:16860-16864.

Whittam TS (1989) Clonal dynamics of *Escherichia coli* in its natural habitat. Antonie Van Leeuwenhoek Journal of Microbiology 55:23-32.

Chapter 5

Genetic and phenotypic characterisation of the genus *Hafnia* and proposal of *Hafnia tempus* sp. nov.

Introduction

Hafnia alvei (Moeller 1954) is a commensal bacterial species that inhabits the gastrointestinal tract of vertebrates (Brenner 1984; Okada and Gordon 2003). It is frequently found as a contaminant of fish, dairy and meat products (Bruhn 2004; Lindberg *et al.* 1998), and is indicated as a cause of meat spoilage (Bruhn 2004; Dainty 1986; Kang *et al.* 2002). This species is also of veterinary significance. It is reported to cause haemorrhagic septicaemia in laying hens (Real *et al.* 1997) and trout species (Acosta *et al.* 2002; Gelev *et al.* 1990), and liver infections in commercial pullets (Proietti *et al.* 2004). In humans it is an opportunistic pathogen causing nosocomial and community-acquired extraintestinal infections (Gunthard and Pennekamp 1996; Ramos and Damaso 2000).

H. alvei is distantly related to *Escherichia coli* (relative DNA binding ratio of 20 % at 60 °C) (Steigerwalt *et al.* 1976), and several phylogenetic studies of the family Enterobacteriaceae suggest that its closest relatives is *Serratia* (Dauga 2002; Hedegaard 1999; Wertz *et al.* 2003). A DNA hybridisation study using four isolates of *H. alvei* found two genetically distinct groups (Steigerwalt *et al.* 1976). Since then random amplified polymorphic DNA-PCR (Ridell *et al.* 1995), partial 16S rDNA sequencing (Janda *et al.* 2002), and multilocus enzyme electrophoresis (Okada and Gordon 2003) have confirmed the existence of two genetic groups in this species. These two groups also have distinct biochemical and physiological properties (Brenner 1984; Janda *et al.* 2002; Okada and Gordon 2001; Okada and Gordon 2003).

The partitioning of strains of a bacterial species into largely distinct genetic subgroups is common among species of the family *Enterobacteriaceae*. *E. coli* and *Salmonella* are classic examples of species consisting of genetically distinct subgroups (Beltran *et al.* 1988; Kotetishvili *et al.* 2002; Ochman *et al.* 1983; Pupo *et al.* 1997; Whittam *et al.* 1983). How then does one distinguish sub-specific from species structure? In order to do so there needs to be a species concept for bacteria. Several eukaryotic species concepts have been embraced by prokaryotic taxonomists, including the ecological species concept, biological species concept, cohesion species concept, and evolutionary species concept (Cohan 2001). Among them, the biological species definition as applied to bacteria identifies a species as a cluster of strains that exchange genetic information among them (Dykhuizen and Green 1991). If there are multiple sequence clusters within a defined species (such as in *E. coli* and *Salmonella*), those clusters must share genetic information among them to be classified as one species. The genes of interest in the biological species concept refer to housekeeping genes that code for enzymes responsible for basic metabolic functions. They are assumed to be neutral in fitness and are not subjected to positive selection (Urwin and Maiden 2003). The multiple groups of *E. coli* are indeed one species because alleles of housekeeping genes are shared among groups (Wirth, unpublished results; website: http://www.web.mpiib-berlin.mpg.de). It is not known if the two genetic groups of *Hafnia* share any alleles at the genetic level. In order to identify the presence (or lack) of allele sharing between the two genetic groups of *H. alvei*, a nucleotide sequence-based comparison of six housekeeping genes (multilocus sequence typing, MLST) (Maiden *et al.* 1998) was conducted. Here we present evidence for lack of genetic recombination between the two groups, and propose a new species, *H. tempus* sp. nov., based on the biological species concept.

Material and methods

Strain selection

Ninety six *H. alvei* strains, 48 strains each from the two electrophoretic groups in the strain set used by Okada and Gordon (2003) were chosen at random. Strain designation, genetic group designation based on MLEE (Okada and Gordon 2003), host of origin and geographic locality are provided in Table 1.

Gene selection

Portions of six housekeeping genes were selected for sequencing. Because the complete genome sequence for *H. alvei* was not available, the distribution of the genes over the chromosome was assumed to reflect the distribution observed in *E. coli*. The six genes chosen for this study were *gapA*, *gnd*, *groEL*, *gyrA*, *gyrB* and *infB*. Information on the genes, primers, and PCR products are given in Table 2.

Sequencing procedure

Chromosomal DNA was extracted with DNAzol® (Molecular Research Center, Inc., Cincinatti, Ohio) from 50 μl of an overnight culture grown in Luria Bertani broth. The PCR reaction mix (40 μl) contained 20 pmol each forward and reverse primers, 2 mM $MgCl_2$, 1.25 U Taq polymerase and 2 μl DNA template in polymerisation buffer (67 mM Tris-HCl (pH8.8), 16.6 mM $(NH_4)_2SO_4$, 0.45 % Triton X-100, 0.2 mg gelatin ml-1 and 0.2 mM dNTPs; Fisher Biotec, Australia). PCR was carried out on a PC-960 aircooled thermal cycler (Cobrett Research Inc., Australia). Because five of the six primers were designed for *E. coli* and not *H. alvei*, a temperature gradient protocol was created as follows: 94 °C for 4 minutes, 2 cycles of 94 °C for 30 seconds, 60 °C for 20 seconds, 72 °C for 80 seconds, 2 cycles of 94 °C for 30 seconds, 55 °C for 20 seconds, 72 °C for 80 seconds, 2 cycles of 94 °C for 30 seconds, 50 °C for 20 seconds, 72 °C for 80 seconds, 2 cycles of 94 °C for 30 seconds, 45 °C for 20 seconds, 72 °C for 80

seconds, 35 cycles of 94 °C for 30 seconds, 40 °C for 20 seconds, 72 °C for 80 seconds, and 72 °C for 4 minutes. PCR products were purified with either ammonium acetate for single band products, or with the Ultra Clean 15 DNA purification kit (Mo Bio Laboratories, Inc., USA) for multiple band products. Five to 20 ng of the purified PCR product was added to the sequencing reaction mix and sequencing was performed using BigDye Terminator chemistry (Applied Biosystems Inc.) on an ABI Prism 3100 Genetic Analyzer (Applied Biosystems Inc.).

Phylogenetic analysis

Consensus sequences for each of the 6 genes were obtained by aligning forward and reverse sequences, editing ambiguous nucleotides, and trimming all 96 consensus sequences to the same length using Sequencher v3 (Gene Codes Corp.). Maximum likelihood trees of the six individual genes as well as the concatenated sequences were constructed using a heuristic search approach using PAUP v4 (Swofford 2003) Identical regions of the six genes for *Escherichia coli* K12 and *Salmonella enterica* subsp. *enterica* serovar Typhi Ty2 were used as out groups (Blattner *et al.* 1997; Deng *et al.* 2003). Optimised values for the maximum likelihood parameters were calculated using MODELTEST v3.06 (Posada and Crandall 1998). Bootstrap analysis (maximum likelihood approach, 100 replications) was conducted for the six genes and the concatenated sequences. The biochemical profiles of the individual strains were mapped onto the maximum likelihood tree of the concatenated sequence using MacClade v4.05 (Maddison 2002).

Table 1. The 96 isolates of *Hafnia* used in this study. Strain, strain identification; MLEE type, multilocus enzyme electrophoresis type. QLD, Queensland; NSW, New South Wales; TAS, Tasmania; SA, South Australia; WA, Western Australia; NT, Northern Territory.

Strain	MLEE group	Host class	Host species	Host locality
B348	1	Aves	*Acanthiza katherina*	QLD
B429	1	Aves	*Sericornis frontalis*	NSW
B459	1	Aves	*Malurus cyaneus*	TAS
B642	1	Aves	*Gymnorhina tibicen*	NSW
E072	1		water	NSW
F138	1	Actinopterygii	*Nematalosa erebi*	SA
F156	1	Actinopterygii	*Nematalosa erebi*	SA
F169	1	Actinopterygii	*Leiopotheropon unicolor*	SA
F182	1	Actinopterygii	*Macquaria ambigua*	SA
F187	1	Actinopterygii	*Macquaria ambigua*	SA
I040	1	Insecta	not identified	NSW
I128	1	Insecta	not identified	NSW
L009	1	Amphibia	*Litoria adelaidensis*	WA
L069	1	Amphibia	*Litoria ewingii*	NSW
M044	1	Mammalia	*Antechinus flavipes*	NSW
M051	1	Mammalia	*Antechinus flavipes*	NSW
M061	1	Mammalia	*Sminthopsis murina*	NSW
M082	1	Mammalia	*Rattus fuscipes*	NSW
M129	1	Mammalia	*Burramys parvus*	NSW
M172	1	Mammalia	*Rattus fuscipes*	NSW
M306	1	Mammalia	*Antechinus stuartii*	NSW
MI472	1	Mammalia	*Dasyurus viverrinus*	TAS
MI490	1	Mammalia	*Mastacomys fuscus*	NSW
MI524	1	Mammalia	*Antechinus swainsonii*	NSW
MI581	1	Mammalia	*Rattus fuscipes*	NSW
MI590	1	Mammalia	*Mastacomys fuscus*	NSW
MI777	1	Mammalia	*Pseudomys fumeus*	NSW
MI782	1	Mammalia	*Antechinus stuartii*	NSW
MI785	1	Mammalia	*Antechinus stuartii*	NSW
MI789	1	Mammalia	*Rattus fuscipes*	NSW
R012	1	Reptilia	*Eulamprus heatwolei*	NSW
R015	1	Reptilia	*Eulamprus heatwolei*	NSW
R021	1	Reptilia	*Eulamprus heatwolei*	NSW
R022	1	Reptilia	*Diplodactylus byrnei*	NSW
R023	1	Reptilia	*Tiliqua rugosa*	SA
R027	1	Reptilia	*Tiliqua rugosa*	SA
R044	1	Reptilia	*Eulamprus heatwolei*	NSW
R055	1	Reptilia	*Eulamprus heatwolei*	NSW
R067	1	Reptilia	*Eulamprus heatwolei*	NSW
R114	1	Reptilia	*Niveoscincus microlepidotus*	TAS
R124	1	Reptilia	*Niveoscincus microlepidotus*	TAS
R126	1	Reptilia	*Niveoscincus mettalicum*	TAS
R180	1	Reptilia	*Eulamprus tympanum*	NSW
R202	1	Reptilia	*Egernia saxatilis*	NSW
R211	1	Reptilia	*Egernia saxatilis*	NSW
R285	1	Reptilia	*Pseudemoia entrecasteauxii*	WA
R373	1	Reptilia	*Pseudechis porphyriacus*	NSW
R394	1	Reptilia	*Rhinoplocephalus nigrescens*	NSW

Table 1. Continued.

Strain	MLEE group	Host class	Host species	Host locality
B255	2	Aves	*Gymnorhina tibicen*	NSW
F118	2	Actinopterygii	*Leiopotheropon unicolor*	SA
F121	2	Actinopterygii	*Leiopotheropon unicolor*	SA
F152	2	Actinopterygii	*Nematalosa erebi*	SA
F231	2	Actinopterygii	*Bidyanus welchi*	SA
F292	2	Actinopterygii	*Macquaria ambigua*	SA
F296	2	Actinopterygii	*Macquaria ambigua*	SA
F298	2	Actinopterygii	*Macquaria ambigua*	SA
F311	2	Actinopterygii	*Macquaria ambigua*	SA
F321	2	Actinopterygii	*Macquaria ambigua*	SA
F348	2	Actinopterygii	*Macquaria ambigua*	SA
F364	2	Actinopterygii	*Macquaria ambigua*	SA
F397	2	Actinopterygii	*Macquaria ambigua*	SA
H698	2	Mammalia	*Homo sapiens*	NSW
L013	2	Amphibia	*Litoria moorie*	WA
M132	2	Mammalia	*Burramys parvus*	NSW
MI164	2	Mammalia	*Dasyurus geoffroii*	WA
MI177	2	Mammalia	*Notomys fuscus*	QLD
MI230	2	Mammalia	*Antechinus bellus*	NT
MI231	2	Mammalia	*Dasyurus hallucatus*	NT
MI240	2	Mammalia	*Antechinus stuartii*	NSW
MI293	2	Mammalia	*Rattus rattus*	NSW
MI310	2	Mammalia	*Rattus fuscipes*	NSW
MI319	2	Mammalia	*Rattus fuscipes*	NSW
MI387	2	Mammalia	*Dasyurus hallucatus*	NT
MI398	2	Mammalia	*Dasyurus viverrinus*	TAS
MI408	2	Mammalia	*Dasyurus viverrinus*	TAS
MI463	2	Mammalia	*Dasyurus viverrinus*	TAS
MI464	2	Mammalia	*Dasyurus viverrinus*	TAS
MI468	2	Mammalia	*Dasyurus viverrinus*	TAS
MI470	2	Mammalia	*Dasyurus viverrinus*	TAS
MI476	2	Mammalia	*Ornithorhynchus anatinus*	TAS
MI690	2	Mammalia	*Homo sapiens*	WA
MI761	2	Mammalia	*Trichosurus caninus*	NSW
MI763	2	Mammalia	*Lagorchestes hirsutus*	NT
MI765	2	Mammalia	*Potorous tridactylus*	NSW
R048	2	Reptilia	*Eulamprus heatwolei*	NSW
R053	2	Reptilia	*Eulamprus heatwolei*	NSW
R054	2	Reptilia	*Eulamprus heatwolei*	NSW
R057	2	Reptilia	*Eulamprus heatwolei*	NSW
R061	2	Reptilia	*Eulamprus heatwolei*	NSW
R117	2	Reptilia	*Niveoscincus microlepidotus*	TAS
R119	2	Reptilia	*Niveoscincus microlepidotus*	TAS
R122	2	Reptilia	*Niveoscincus microlepidotus*	TAS
R175	2	Reptilia	*Rhinocephalus bicolour*	WA
R184	2	Reptilia	*Eulamprus tympanum*	NSW
R231	2	Reptilia	*Ramphotyphlops sp.*	WA
R252	2	Reptilia	*Echopsis curta*	WA

Table 2. MLST primer information.

Gene	protein	*E. coli* location	Length	Forward primer	Reverse primer	reference
gapA	Glyceraldehyde-3-phosphate dehydrogenase	1860000	692	acgttctgacatcgagatcgttg	gctttagcatcgaacacggaagtgc	Wertz *et al.* 2004
gnd	Gluconate dehydrogenase	2100000	598	aacatcgaaagccgtggtta	tcttcatcyttyttggtgaa	this study
groEL	groEL (Hsp60)	4369000	561	gacgctcgygtraaaatgctac	tgttgatgaagtayggrgacagg	Wertz *et al.* 2004
gyrA	Gyrase A	2340000	711	atgagcgaccttgccagag	cgtgaacgatgatggtttcg	Wertz *et al.* 2004
gyrB	Gyrase B	3875000	577	tccaytaygaaggyggyatc	tccaccargtamagttcrga	this study
infB	Initiation factor 2	3310000	542	atyatggghcaygthgaycay gghaarac	tatccgacgccgaactccgrttncgca tngcncgnayncgncc	Hedegaard *et al.* 1999

Results

Genetic differentiation between the two genetic groups of *Hafnia*

Allele assignment revealed total lack of allele sharing between the two genetic groups of *Hafnia*. The clustering of the two genetic groups in the maximum likelihood tree using the concatenated data had over 95 % bootstrap support (Figure 1). The pairwise differences between the two genetic groups ranged between 4.1 % (*groEL*) and 11.5 % (*gnd*), with an overall average of 7.3 % (Table 3). Bootstrap support for the two genetic groups based on the individual genes was substantially lower, in particular for the clustering of group 1 strains for *gapA* and *groEL* (Figure 2). Maximum likelihood trees of all six gene trees with the exception of *gyrB* had equal branch lengths from the out group species *Escherichia coli* and *Salmonella enterica*, indicating the presence of a common ancestor of the two *Hafnia* groups (Figure 2). The gene tree for *gyrB* showed that genetic group 1 diverged from group 2 (Figure 2). In order to observe the fine scale relationship between the two genetic groups of *Hafnia*, maximum likelihood trees of the individual genes were reconstructed without the outgroup rooting of *E. coli* and *S. enterica*. All unrooted gene trees showed a similar overall structure with the rooted *gyrB* tree (Figure 3). Genetic group 1 of all six trees were connected to a certain clade of strains in genetic group 2 (Figure 3, 4). This group 2 clade consisted of seven isolates, three from mammals and four from fish (Figure 4). Of the 269 nucleotides that distinguished the two genetic groups, the group 2 clade of seven strains shared from 4.3 % (*gyrB*) to 13 % (*groEL*) of those nucleotides with group 1 strains.

Figure 1. Maximum likelihood tree of the six concatenated sequences of *Hafnia*. *Escherichia coli* K12 (K12) and *Salmonella enterica* (saen) are used as outgroups to root the tree. **1**, genetic group 1; **2**, genetic group 2. The branches marked with a solid line correspond to the clade in genetic group 2 that is genetically closest to group 1. Branches with bootstrap scores > 90 are indicated.

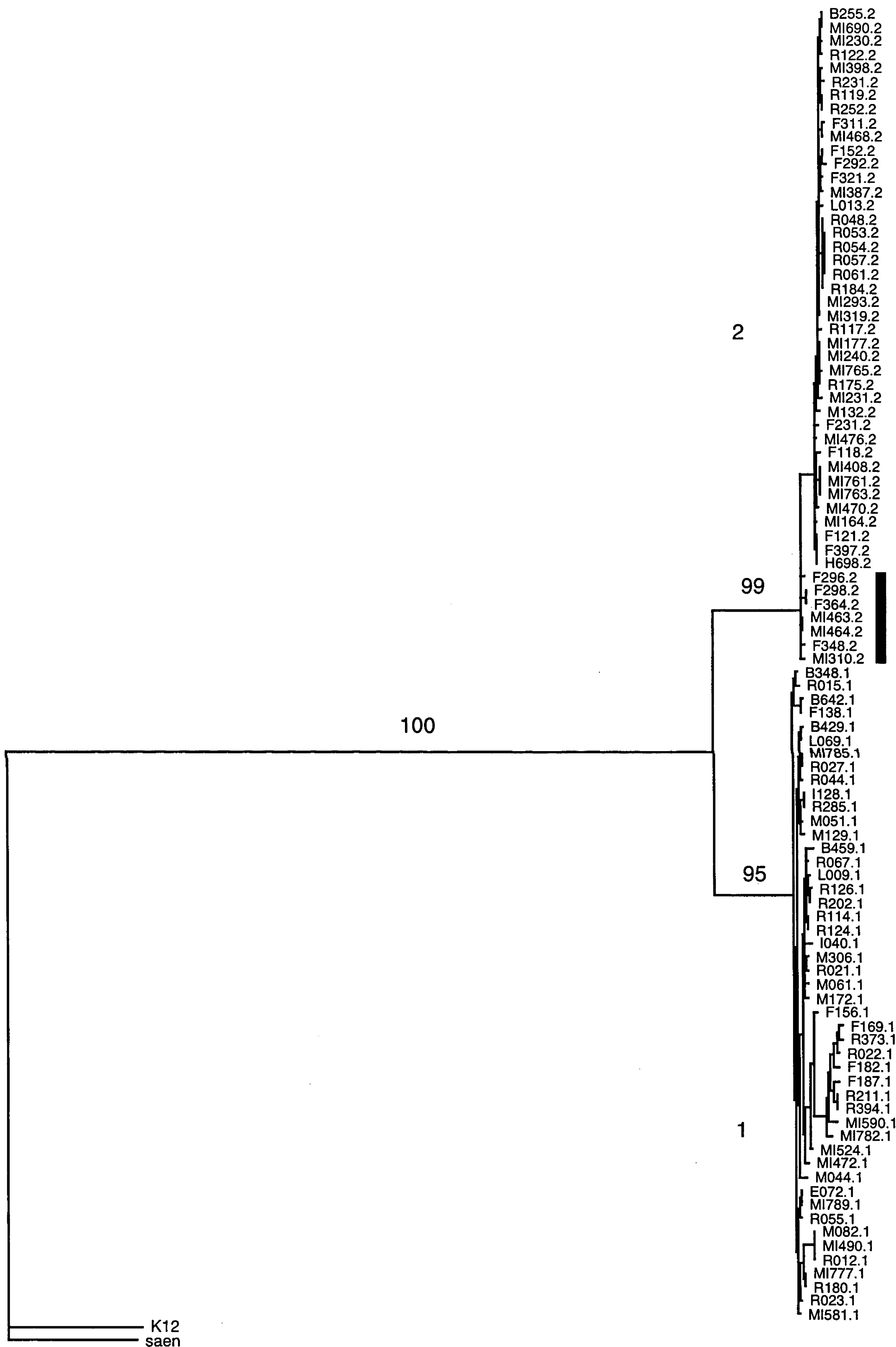


Figure 1.

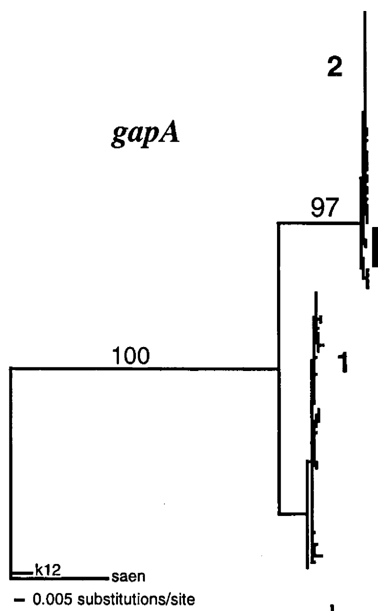
Table 3. Average percentage pairwise distances of the six partial sequences of *Hafnia*. K12, *Escherichia coli* K12; saen, *Salmonella enterica* subsp. *enterica* serovar Typhi Ty2.

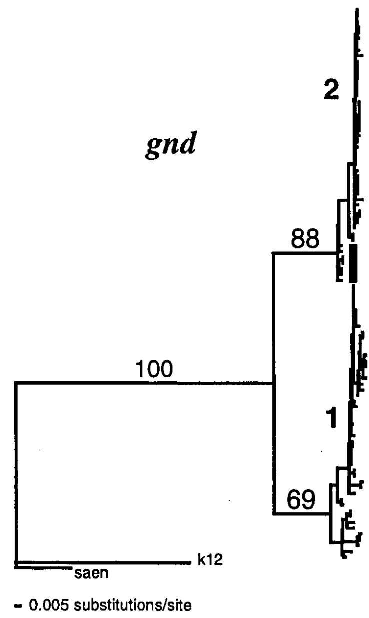
		group 1	group 2	K12	saen
gapA	group 1	0.36			
	group 2	6.78	0.23		
	K12	14.42	16.20	0.00	
	saen	16.71	17.51	6.50	0.00
gnd	group 1	2.26			
	group 2	11.53	1.19		
	K12	19.77	21.21	0.00	
	saen	19.38	18.73	14.88	0.00
groEL	group 1	0.75			
	group 2	4.14	0.39		
	K12	10.41	11.24	0.00	
	saen	12.28	12.58	6.95	0.00
gyrA	group 1	0.51			
	group 2	6.35	0.30		
	K12	18.54	18.16	0.00	
	saen	17.94	18.02	8.16	0.00
gyrB	group 1	1.74			
	group 2	8.17	1.04		
	K12	15.53	14.84	0.00	
	saen	15.97	14.92	7.97	0.00
infB	group 1	1.26			
	group 2	7.00	0.90		
	K12	15.39	16.07	0.00	
	saen	13.90	15.22	9.04	0.00

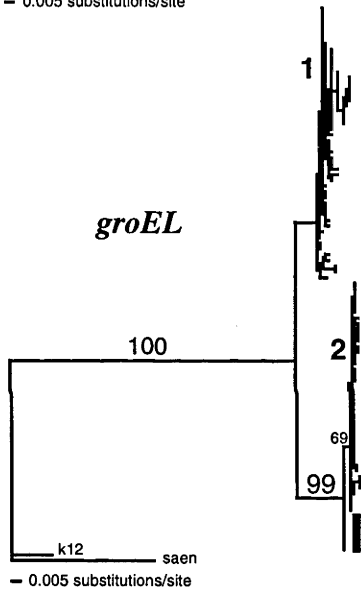
Figure 2. Maximum likelihood trees of the six genes of *Hafnia*. *Escherichia coli* K12 (k12) and *Salmonella enterica* (saen) are used as outgroups to root the tree. **1**, genetic group 1; **2**, genetic group 2. The branches marked with a solid line correspond to the clade in genetic group 2 that is genetically closest to group 1. Bootstrap scores > 60 are indicated.

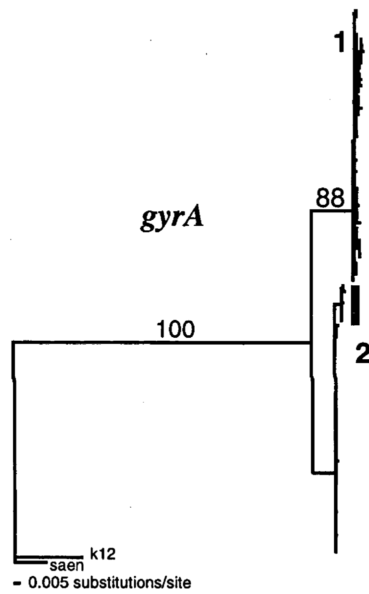
Figure 3. Maximum likelihood trees of the six genes of *Hafnia*. The trees are unrooted. **1**, genetic group 1; **2**, genetic group 2. The genetic group that is separated due to the positioning of the other genetic group in the individual trees are surrounded by a dashed rectangle. The branches marked with a solid line correspond to the clade in genetic group 2 that is genetically closest to group 1. Bootstrap scores for the clustering of the two Hafnia groups are indicated.

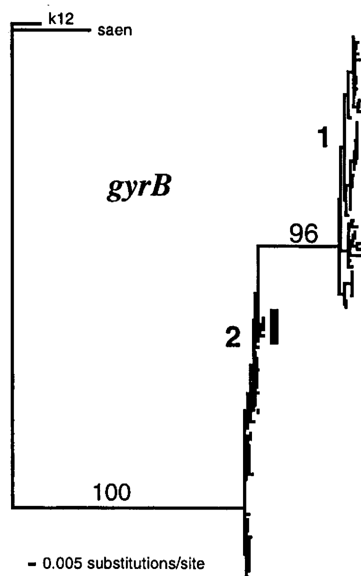
Figure 4. Maximum likelihood tree of the six concatenated sequences of *Hafnia*. The tree is unrooted. **1**, genetic group 1; **2**, genetic group 2. The branches marked with a solid line correspond to the clade in genetic group 2 that is genetically closest to group 1. Bootstrap scores > 80 are indicated.











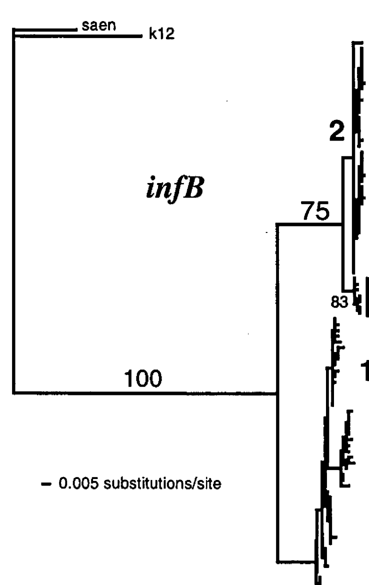
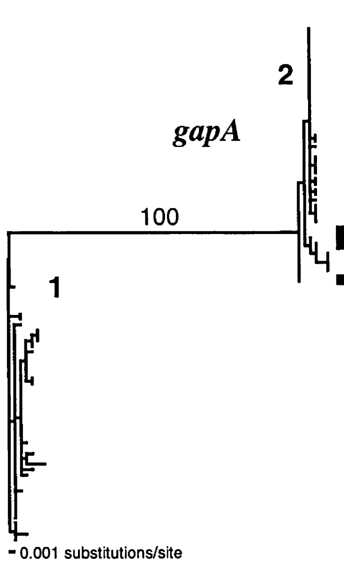
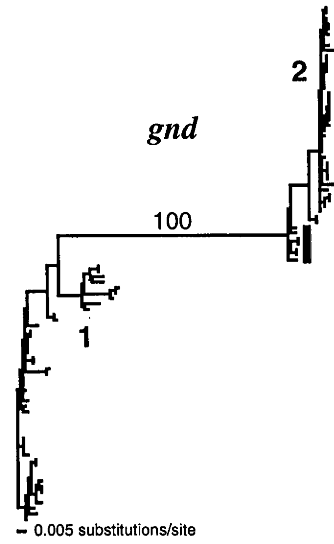
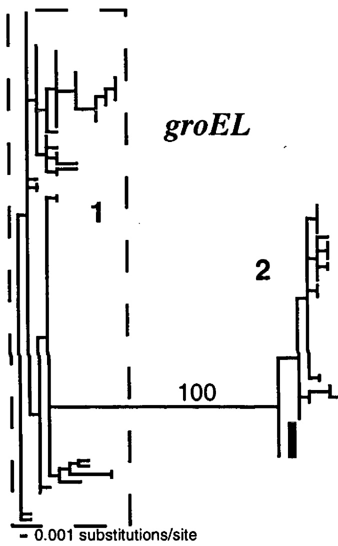
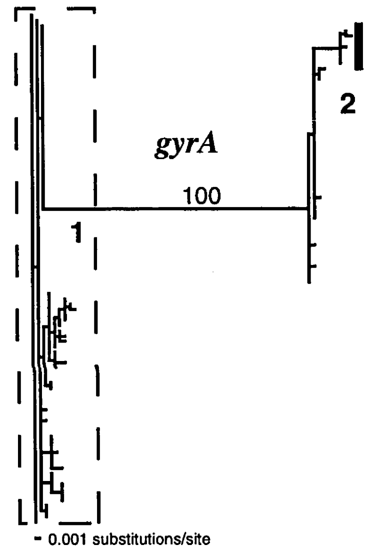


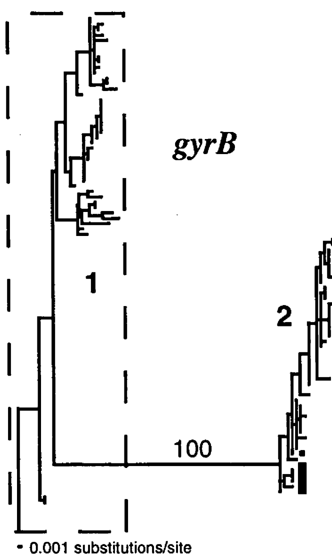
Figure 2.











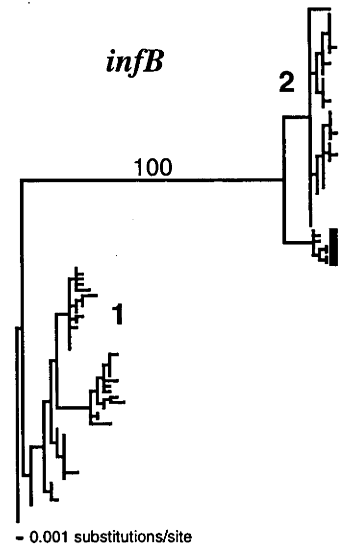


Figure 3.

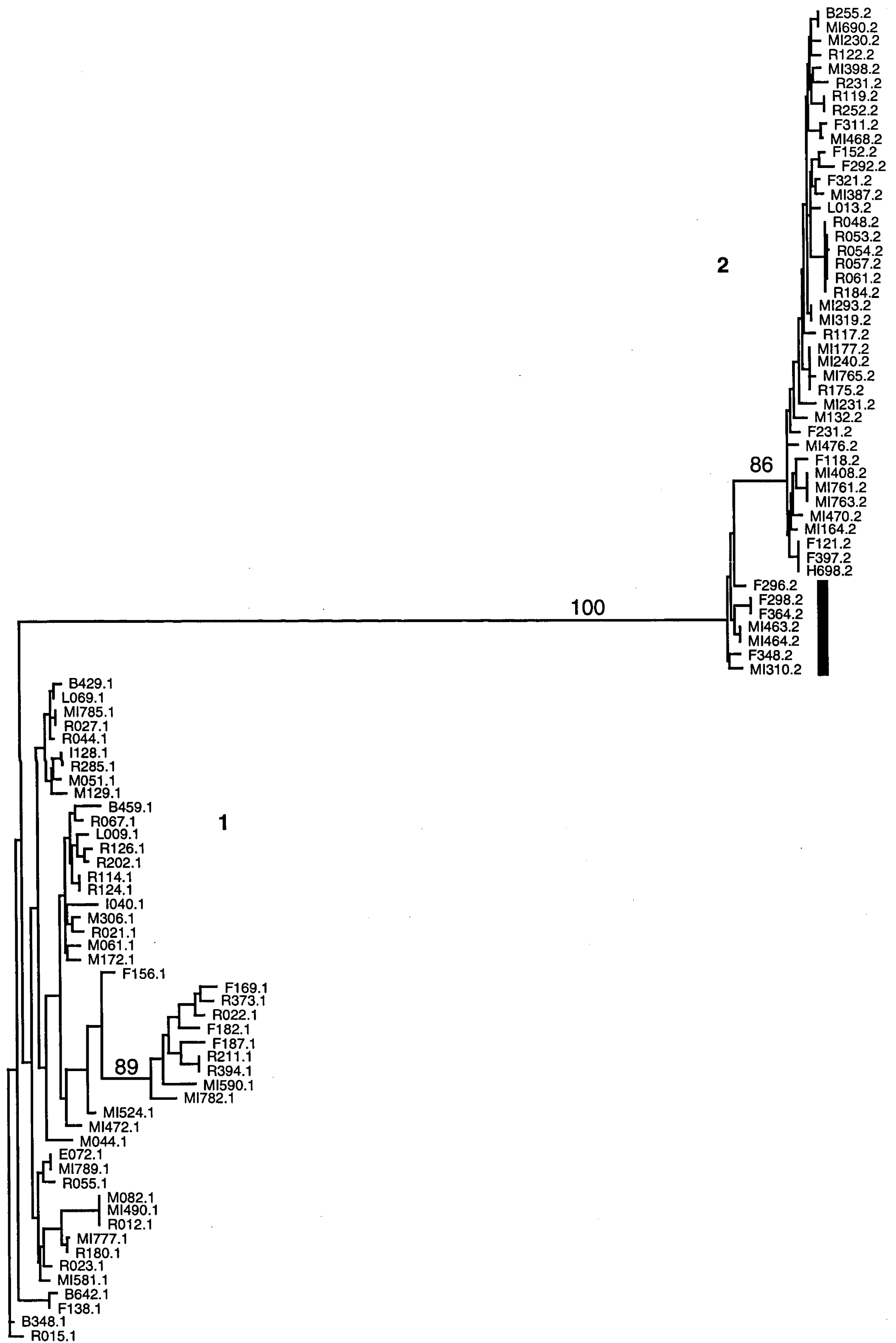


Figure 4.

Biochemical differentiation between the two genetic groups of *Hafnia*

The concatenated tree of the six genes was highly correlated with two biochemical traits that largely distinguished the two genetic groups. All genetic group 1 strains were able to utilise malonate as a carbon source and produce high amounts of acid during glucose fermentation (as indicated by the methyl red test). On the other hand the majority of genetic group 2 strains were negative for both of these characters (95 % negative for malonate utilisation and 80 % negative for acid production) except for the clade in group 2 that was genetically closest to group 1. Unlike the majority of group 2 strains, the seven strains in this clade were all positive for malonate utilisation and 70 % (five isolates out of seven) were positive for acid production during glucose fermentation (Table 4).

Although motility was one of the physiological properties that Brenner (1984) found to be a distinguishing characteristic of the two genetic groups of *Hafnia*, 33 % of the group 1 strains in this study were motile (compared to 56 % of group 2 strains; Table 4). There were eight other biochemical characteristics for which the proportions of strains with positive reactions were significantly different between the two *Hafnia* groups. Seven of these traits are presented in Table 4.

Table 4. Ten phenotypic characteristics of the 96 isolates of *Hafnia*. Characteristics are mapped onto the maximum likelihood concatenated sequence tree. The branches marked within a solid line correspond to the clade in genetic group 2 that is genetically closest to group 1. Motility was scored on a visual three-scale point system; -, no motility; + moderate motility; ++, strong motility. MLO, malonate; MR, methyl red test; MOT, motility; MEL, melibiose; RHA, rhamnose; NPG, p-n-p-β-galactoside; NAG, p-n-p-acetyl glucosaminide; ESC, esculin; ARG, arginine; TTC, triphenyl tetrazolium chloride. *Escherichia coli* K12 (k12) and *Salmonella enterica* (saen) are used to root the tree.

Table 4.

2

1

	MLO	MR	MOT	MEL	RHA	NPG	NAG	ESC	ARG	TTC
	-	-	-	+	+	+	+	-	-	-
	-	-	-	+	+	+	+	-	-	-
	-	-	-	+	+	+	+	-	+	-
	-	-	-	+	+	+	+	-	+	+
	-	-	+	+	+	-	+	-	+	-
	-	-	-	+	-	+	+	-	-	-
	-	-	+	-	+	+	+	-	+	+
	-	-	+	-	+	+	+	-	+	+
	-	-	-	-	+	+	+	-	+	+
	-	-	-	-	+	+	+	-	+	+
	-	-	-	-	+	+	+	-	+	-
	-	-	++	-	+	+	+	-	+	+
	-	-	++	-	+	+	+	-	+	+
	-	-	++	-	+	+	+	-	+	-
	-	-	-	+	+	+	+	-	+	+
	-	-	++	-	+	+	+	-	+	+
	-	+	-	-	+	+	-	-	-	-
	-	-	-	-	+	+	+	-	+	+
	-	-	++	-	+	+	+	-	+	+
	+	-	-	-	+	+	+	-	+	+
	-	-	++	-	+	+	+	-	+	+
	-	+	++	+	+	+	+	-	+	-
	-	+	++	+	+	+	+	-	+	-
	-	-	-	+	+	+	+	-	+	+
	-	-	++	-	+	+	+	-	+	-
	-	-	++	-	+	+	+	-	+	-
	-	+	++	-	+	+	+	-	+	-
	-	-	-	+	+	+	+	-	-	+
	-	+	-	-	+	+	+	-	+	-
	-	-	-	+	+	+	+	-	+	-
	-	+	+	-	+	+	+	-	+	+
	-	-	++	-	+	+	+	-	+	+
	-	-	++	-	+	+	+	-	+	-
	-	-	+	-	+	+	+	-	+	-
	-	-	++	-	+	+	+	-	+	-
	+	-	++	-	+	+	+	-	-	+
	-	-	++	+	+	+	+	-	+	+
	-	-	++	-	+	-	+	-	+	+
	-	-	+	-	+	+	+	-	+	+
	-	+	-	-	+	+	+	-	+	+
	-	+	+	-	+	+	+	-	+	+
	+	+	-	-	+	+	+	-	+	+
	+	-	-	-	+	+	+	-	+	+
	+	+	-	-	+	+	+	-	+	-
	+	+	++	-	+	-	+	-	+	-
	+	-	+	-	+	-	+	-	+	+
	+	+	-	-	+	+	+	-	+	+
	+	+	++	-	-	+	+	-	+	-
	+	+	-	+	+	+	+	-	-	+
	+	+	-	+	+	-	+	-	+	+
	+	+	-	-	+	+	+	-	+	+
	+	+	-	-	+	+	+	-	-	+
	+	+	-	-	+	+	+	-	-	+
	+	+	-	-	-	-	-	-	+	-
	+	+	-	-	+	+	+	-	-	+
	+	+	+	-	+	+	+	-	-	+
	+	+	-	-	+	+	+	-	-	+
	+	+	-	+	+	-	-	-	+	+
	+	+	-	-	-	-	+	-	+	-
	+	+	-	-	-	+	-	-	+	-
	+	+	-	-	+	+	+	+	+	+
	+	+	+	-	+	+	-	+	+	+
	+	+	-	-	+	+	-	+	-	+
	+	+	-	-	+	+	+	+	-	+
	+	+	-	-	+	-	+	-	+	+
	+	+	+	-	+	-	+	-	-	+
	+	+	-	-	+	-	+	+	-	+
	+	+	+	-	-	-	-	-	+	+
	+	+	+	-	-	-	+	-	+	+
	+	+	-	-	+	+	+	-	-	+
	+	+	-	-	-	-	+	-	+	-
	+	+	+	+	-	+	+	-	+	+
	+	+	-	-	+	-	-	-	-	-
	+	+	+	-	-	+	-	-	-	+
	+	+	-	-	+	+	+	-	+	+
	+	+	++	-	+	-	-	-	+	-
	+	+	-	-	+	+	-	-	+	+
	+	+	-	+	+	+	+	+	+	+
	+	+	-	-	+	+	+	+	+	+
	+	+	+	-	+	+	+	+	+	+
	+	+	-	-	+	+	-	+	+	+
	+	+	-	-	-	+	-	+	+	+
	+	+	-	-	-	-	-	-	+	+
	+	+	++	-	-	-	+	-	+	-
	+	+	+	-	+	-	+	-	+	+
	+	+	+	-	-	-	+	-	+	+
	+	+	-	-	+	+	+	-	-	+
	+	+	-	-	+	+	-	-	-	+
	+	+	+	-	-	-	+	-	-	+
	+	+	+	-	+	+	-	-	-	+
	+	+	+	-	-	-	+	-	+	-
	+	+	-	-	+	+	+	-	-	+
	+	+	-	-	-	+	+	+	+	+
	+	+	+	-	+	-	+	-	-	+
	-	+	-	-	-	+	-	-	-	+
	+	+	-	-	+	+	+	-	+	+
saen										
k12										

Discussion

Application of the biological species concept to bacterial speciation

The biological species concept as applied to organisms that reproduce sexually defines a species as the boundary of sexual reproduction (Mayr 1963). There are two major arguments against the application of the biological species concept to bacteria. The first argument is that bacteria reproduce asexually, therefore mutation is the driving force of genetic divergence and that recombination is too rare to have an effect on genetic diversity (Cohan 2002). However, there is strong evidence based on studies investigating sequence variation in housekeeping genes of bacterial populations that genetic exchange is frequent enough to have a significant influence on their population structure (Guttman and Dykhuizen 1994a; Spratt *et al.* 2001; Suerbaum *et al.* 1998; Thampapillai *et al.* 1994). It is also important to note that periodic selection (termed 'selective sweep') in local regions of the chromosome is an extreme consequence of intergenic recombination that results in allele sharing among strains of a species (Guttman and Dykhuizen 1994b; Majewski and Cohan 1999). The second argument against the application of the biological species concept is that genetic exchange in bacteria, unlike sexually reproducing organisms, occurs beyond the species boundary (Hoffmann *et al.* 1998). It has been shown that the rate of homologous recombination of chromosomal genes decreases with increasing sequence divergence (Vulic *et al.* 1997). This outcome will limit recombination among strains with a certain degree of genetic similarity. Lawrence (2002) suggests that bacterial speciation is a continuous process whereby the gradual accumulation of neutral mutations over time will decrease the rate of homologous recombination between two populations (each of which exploits a novel niche) until they reach total reproductive isolation. Genes acquired through interspecies recombination will gradually change over time to resemble the recipient chromosome because the foreign gene will be exposed to the mutational processes of the recipient DNA (Lawrence and

Ochman 1997), and therefore will consequently recombine with other members of the group. The biological species concept has been successfully applied to bacterial populations including *E. coli* (Dykhuizen and Green 1991), species in the family *Enterobacteriaceae* (Wertz *et al.* 2003), *Shigella* species (Lan and Reeves 2002), *Pseudomonas syringae* (Sarkar and Guttman 2004), and *Brucella* species (Moreno *et al.* 2002).

Subspecific versus species structure in bacterial populations

What makes a subspecies genetic clustering a distinct species as opposed to just a subspecies structure within a species? The most important distinction between species versus subspecies structure is whether or not the separate genetic clusters share a significant amount of genetic information. *E. coli* is an example of a species with extensive subspecies structure. It has long been known that the *E. coli* reference (ECOR) collection, which was compiled to represent the species' genetic diversity (Ochman 1984), consists of four major genetic clusters (Goullet 1989). Most pathogenic and non-pathogenic strains of *E. coli* fall more or less into one of the four genetic clusters (Pupo *et al.* 1997; Pupo *et al.* 2000). There is extensive evidence for intergenic recombination of housekeeping genes among the four genetic clusters (Dykhuizen and Green 1991; Guttman and Dykhuizen 1994a; Lecointre *et al.* 1998; Nelson and Selander 1994; Wirth, unpublished results). The presence of allele sharing in *E. coli* strongly suggests that although the genetic clusters are formed by some force of cohesion (e.g. ecological structure shaped by host biology) (Gordon and Cowling 2003), they all belong to one species. This is in contrast to the lack of allele sharing between the two *Hafnia* groups.

The extent of genetic differentiation between the two *Hafnia* groups

Wertz *et al.* (2003) looked at partial sequences of five housekeeping genes from seven species in the family Enterobacteriaceae. They found that the proportion of pairwise differences of the five genes between *Klebsiella pneumoniae* and *K. oxytoca* ranged from 6 to 11 %, while the average pairwise differences within each species varied from 0 to 5 %. By contrast, divergence among six enteric bacterial species of different genera exceeded 10 %. The pairwise genetic distance among and within the two *Hafnia* groups in this study are comparable to the pairwise distances found between the two *Klebsiella* species, which suggests that the two *Hafnia* groups should be considered two different species within the same genus.

The ancestor group of *Hafnia* may be genetic group 1

The presence of an outgroup taxa in the gene tree may provide information on the evolutionary relationships among subgroups of the ingroup taxa (Lecointre *et al.* 1998). However, in order to maintain the true genealogy of the ingroup taxa, the outgroup sequence must retain some of the same genetic signals that distinguish the ingroup populations (an illustrative example is presented in Figure 5) (Wheeler 1990). Potential outgroup species (e.g. *Serratia plymuthica*, *Enterobacter cloacae*) were selected for each individual gene tree based on BLAST searches for the individual genes. However, all of the selected outgroup species resulted in equal branch lengths of the two *Hafnia* groups from the outgroup species (data not shown). This observation can be interpreted in two ways; either that there was an ancestral species of the two *Hafnia* groups that no longer exists, or that the selected outgroup taxa were random roots that failed to provide any historical information on the genetic relationship among the ingroup taxa (Wheeler 1990).

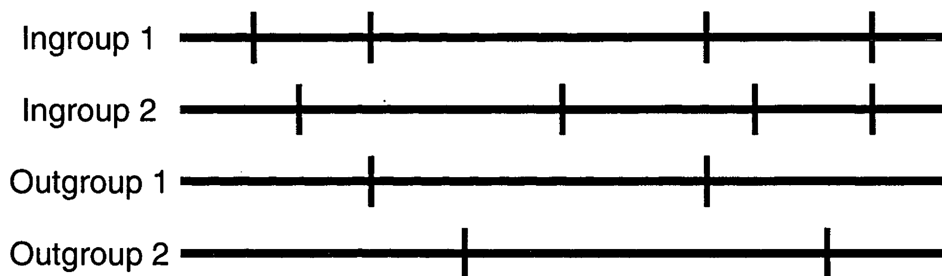


Figure 5. A hypothetical diagram of an ideal outgroup taxon. Vertical bars on the sequence depict informative bases that distinguish a taxon from others. Outgroup 1 is ideal as an outgroup taxon, as it shares two bases with ingroup 1 that idstinguishes it from ingroup 2. On the other hand, outgroup 2 is a random root that does not share any bases that distinguish the two ingroup taxa, and therefore is not ideal as an outgroup taxon.

Based on the phylogenetic order from *E. coli* and *S. enterica* to the two *Hafnia* groups of *gyrB*, it appears as though *Hafnia* group 1 evolved from group 2 (Figure 3). The structures of the six unrooted gene trees support the fact that the clade in genetic group 2, which shared two biochemical characteristics with group 1 strains, were genetically closest to group 1. If group 1 evolved from group 2, a convergent evolution must have occurred for the two characteristics after genetic divergence of the two groups (Figure 6a). On the other hand, if group 2 evolved from group 1, there is only one evolutionary step involved for the loss of the two phenotypic characteristics in group 2 (Figure 6b). Therefore, the more likely scenario is that the ancestral species is group 1, where all strains had the ability to utilise malonate and to produce acid during glucose fermentation, from which a population in group 2 after genetic divergence lost those two traits. The bootstrap values for each of the groups in the individual gene trees (except for *gyrB*) was over 75 % for group 2 strains, whilst *gnd* was the only gene where the clustering of group 1 strains had over 60 % bootstrap support (Figure 3). This observed relationship between the two *Hafnia* groups where there is less support for clustering of strains in one group (group 1) than the other (group 2) is analogous to the emergence of a new ecotype from an ancestral population (Cohan 2001). Because the origin of the derived population is a single mutant or recombinant, that population of strains would tend to have lower genetic diversity, and as a consequence will have higher bootstrap support.

Among the six genes investigated in this study, *gyrB* was the only gene where *E. coli* and *S. enterica* shared enough genetic information (i.e. parsimony informative nucleotides) with one of the two *Hafnia* groups. There is suggestive evidence for *Hafnia* group 1 (represented by *H. alvei* type strain ATCC 13337) acquiring the *gyrB* gene through interspecies horisontal gene transfer (Dauga 2002). Because of the more likely explanation for group 1 being the ancestral population, *Hafnia* group 1

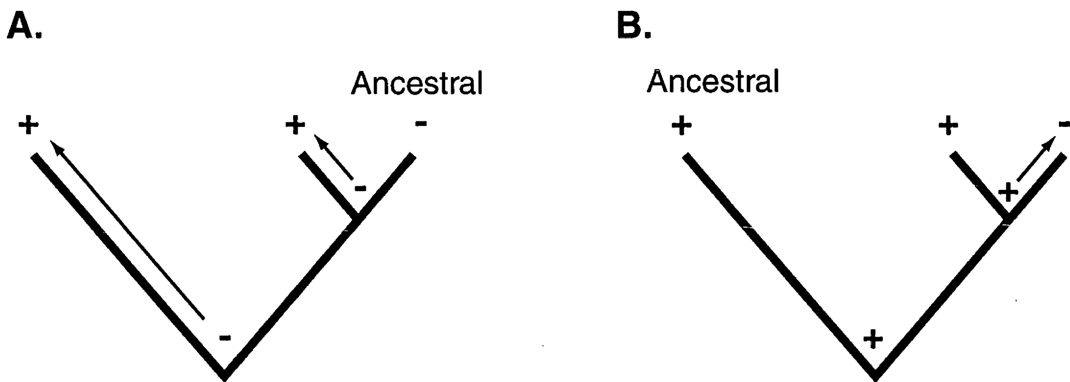


Figure 6. A hypothetical diagram of convergent evolution versus one-step evolution of a phenotypic trait. In case A, the ancestral population does not possess the trait that the evolved population has. If there is a cluster of strains within the ancestral population that also possesses the trait, then the trait would have had to arise twice (convergent evolution). In case B, the ancestral population possesses the trait, from which a population evolves. The majority of the evolved population loses the trait, but a fraction of the strains that are genetically closer to the ancestral population retains the trait. Therefore the evolution of the trait in case B is a one-step process.

may have acquired the *gyrB* gene from the group 2 lineage before the two groups diverged.

Polyphasic taxonomy requires not only genotypic differences to demarcate a species, but the phenotypic traits must also be distinct (Vandamme *et al.* 1996). The two *Hafnia* groups significantly differ in their host distribution. The majority of group 1 strains recovered from mammals were only isolated from hosts weighing less than 200 g. By contrast, group 2 strains from mammals were isolated from hosts of any body mass. The majority of fish hosts carried group 2 strains (Okada and Gordon 2003). Furthermore, Janda *et al.* (2002) found a higher proportion of the clinical isolates of *Hafnia* belonging to group 2. On the other hand, 70 % of the *Hafnia* strains isolated from minced meat belonged to group 1 identified by RAPD-PCR analysis (Ridell *et al.* 1995). This difference in distribution of the two *Hafnia* groups indicates that these two groups have different ecological niches. The two groups also differ in their phenotypic characteristics. Group 1 strains have a more narrow thermal niche (14 to 40 °C) compared to group 2 strains (10 to 40 °C; (Okada and Gordon 2001). Group 2 strains also had a significantly higher maximum growth rate at its optimal temperature, which was 1 °C higher than that of group 1 strains (Okada, Chapter 4). An isolate can be biochemically assigned to one of the two *Hafnia* groups with 95 % accuracy (Okada and Gordon 2003).

All of the genetic and phenotypic information that has been collected provides solid evidence for strains belonging to genetic group 2 being a separate species within the genus *Hafnia*, and we propose a new species, *Hafnia tempus* sp. nov. (*tempus* meaning opportunity or occasion). Although limited, available data suggests that *Hafnia* group 2 strains have clinical significance, and that group 1 strains are responsible for food contamination and are of veterinary significance (Gelev *et al.*

1990; Janda *et al.* 2002; Proietti *et al.* 2004; Ridell *et al.* 1995). A wider survey of the clinical, veterinary, and food isolates will provide better information on the prevention of disease and quality control of food products.

Description of *Hafnia tempus* sp. nov.

Hafnia tempus. Gram negative, rod shaped and motile. Forms white colonies on McConkey and Luria Bertani agar. Oxidase and indole negative, ferments D-glucose, D-mannose, D-mannitol, D-galactose. Negative for D-sucrose, D-adonitol, D-inositol fermentation and esculin. Positive for the Voges-Proskauer test. Phenotypic differentiation from *Hafnia alvei* (type strain ATCC13337) is based on the positive probability of a combination tests represented by malonate utilisation, production of acid during glucose fermentation (indicated by the methyl red test), and motility (Table 3). Type strain is ATCC 29927 (American Type Culture Collection, USA), which was assigned to *Hafnia* genetic group 2 by MLEE (Okada and Gordon 2003).

References

Acosta F, Real F, Caballero MJ, Sieiro C, Fernandez A, Rodriguez LA (2002) Evaluation of immunohistochemical and microbiological methods for the diagnosis of brown trout infected with *Hafnia alvei.* Journal of Aquatic Animal Health 14:77-83.

Beltran P, Musser JM, Helmuth R, Farmer JJ, Frerichs WM, Wachsmuth IK, Ferris K, McWhorter AC, Wells JG, Cravioto A, Selander RK (1988) Toward a population genetic analysis of *Salmonella* - genetic diversity and relationships among strains of serotypes *S. choleraesuis*, *S. derby, S. dublin*, *S. enteritidis*, *S. heidelberg*, *S. infantis*, *S. newport*, *and S. typhimurium.* Proceedings of the National Academy of Sciences of the United States of America 85:7753-7757.

Brenner DJ (1984) Family I. Enterobacteriaceae. In: Krieg NR (ed) *Bergey's Manual of Systematic Bacteriology.* Williams and Wilkins, Baltimore, pp. 408-420.

Bruhn JB, Christensen, A. B., Flodgaard, L. R., Nielsen, K. F., Larsen, T. O., Givskov, M., and Gram, L. (2004) Presence of acylated homoserine lactones (AHLs) and AHL-producing bacteria in meat and potential role of AHL in spoilage of meat. Applied and Environmental Microbiology 70:4293-4302.

Cohan FM (2001) Bacterial species and speciation. Systematic Biology 50:513-524.

Cohan FM (2002) What are bacterial species? Annual Review of Microbiology 56:457-487.

Dainty RH, Edwards RA, Hibbard CM, Ramantanis SV (1986) Bacterial sources of putrescine and cadaverine in chill stored vacuum-packaged beef. Journal of Applied Bacteriology 61:117-123.

Dauga C (2002) Evolution of the *gyrB* gene and the molecular phylogeny of Enterobacteriaceae: a model molecule for molecular systematic studies. International Journal of Systematic and Evolutionary Microbiology 52:531-547.

Dykhuizen DE, Green L (1991) Recombination in *Escherichia coli* and the definition of biological species. Journal of Bacteriology 173:7257-7268.

Gelev I, Gelev E, Steigerwalt AG, Carter GP, Brenner DJ (1990) Identification of the bacterium associated with haemorrhagic septicaemia in rainbow trout as *Hafnia alvei.* Research in Microbiology 141:573-576.

Gordon DM, Cowling A (2003) The distribution and genetic structure of *Escherichia coli* in Australian vertebrates: host and geographic effects. Microbiology 149:3575-3586.

Goullet P, Picard B (1989) Comparative electrophoretic polymorphism of esterases and other enzymes in *Escherichia coli.* Journal of General Microbiology 135:135-143.

Gunthard H, Pennekamp A (1996) Clinical significance of extraintestinal *Hafnia alvei* isolates from 61 patients and review of the literature. Clinical Infectious Diseases 22:1040-1045.

Guttman DS, Dykhuizen DE (1994a) Clonal divergence in *Escherichia coli* as a result of recombination, not mutation. Science 266:1380-1383.

Guttman DS, Dykhuizen DE (1994b) Detecting selective sweeps in naturally-occurring *Escherichia coli.* Genetics 138:993-1003.

Hedegaard J, Steffensen SA, Norskov-Lauritsen N, Mortensen KK, Sperling-Petersen HU (1999) Identification of Enterobacteriaceae by partial sequencing of the gene encoding translation initiation factor 2. International Journal of Systematic Bacteriology 49:1531-1538.

Hoffmann A, Thimm T, Droge M, Moore ERB, Munch JC, Tebbe CC (1998) Intergeneric transfer of conjugative and mobilizable plasmids harbored by *Escherichia coli* in the gut of the soil microarthropod *Folsomia candida* (Collembola). Applied and Environmental Microbiology 64:2652-2659.

Janda JM, Abbott SL, Khashe S, Probert W (2002) Phenotypic and genotypic properties of the genus *Hafnia.* Journal of Medical Microbiology 51:575-580.

Kang DH, Arthur TM, Siragusa GR (2002) Gas formation in ground beef chubs due to *Hafnia alvei* is reduced by multiple applications of antimicrobial interventions to artificially inoculated beef trim stock. Journal of Food Protection 65:1651-1655.

Kotetishvili M, Stine OC, Kreger A, Morris JG, Sulakvelidze A (2002) Multilocus sequence typing for characterization of clinical and environmental *Salmonella* strains. Journal of Clinical Microbiology 40:1626-1635.

Lan RT, Reeves PR (2002) *Escherichia coli* in disguise: molecular origins of *Shigella.* Microbes and Infection 4:1125-1132.

Lawrence JG (2002) Gene transfer in bacteria: Speciation without species? Theoretical Population Biology 61:449-460.

Lawrence JG, Ochman H (1997) Amelioration of bacterial genomes - rates of change and exchange. Journal of Molecular Evolution 44:383-397.

Lecointre G, Rachdi L, Darlu P, Denamur E (1998) *Escherichia coli* molecular phylogeny using the incongruence length difference test. Molecular Biology and Evolution 15:1685-1695.

Lindberg AM, Ljungh A, Ahrne S, Lofdahl S, Molin G (1998) Enterobacteriaceae found in high numbers in fish, minced meat and pasteurised milk or cream and the presence of toxin encoding genes. International Journal of Food Microbiology 39:11-17.

Maddison DR, Maddison WP (2002) MacClade 4.05. Sinauer Associates, Sunderland, Massachusetts.

Maiden MCJ, Bygraves JA, Feil E, Morelli G, Russell JE, Urwin R, Zhang Q, Zhou JJ, Zurth K, Caugant DA, Feavers IM, Achtman M, Spratt BG (1998) Multilocus sequence typing - a portable approach to the identification of clones within populations of pathogenic microorganisms. Proceedings of the National Academy of Sciences of the United States of America 95:3140-3145.

Majewski J, Cohan FM (1999) Adapt globally, act locally: The effect of selective sweeps on bacterial sequence diversity. Genetics 152:1459-1474.

Mayr E (1963) *Animal Species and Evolution.* Harvard University Press, Cambridge, Massachusetts.

Moeller V (1954) Distribution of amino acid decarboxylases in Enterobacteriaceae. Acta Pathologica Microbiologica Scandinavica 35:259-277.

Moreno E, Cloeckaert A, Moriyon I (2002) *Brucella* evolution and taxonomy. Veterinary Microbiology 90:209-227.

Nelson K, Selander RK (1994) Intergeneric transfer and recombination of the 6-phosphogluconate dehydrogenase gene (*gnd*) in enteric bacteria. Proceedings of the National Academy of Sciences of the United States of America 91:10227-10231.

Ochman H, Selander RK (1984) Standard reference strains of *Escherichia coli* from natural populations. Journal of Bacteriology 157:690-693.

Ochman H, Whittam TS, Caugant DA, Selander RK (1983) Enzyme polymorphism and genetic population structure in *Escherichia coli* and *Shigella*. Journal of General Microbiology 129:2715-2726.

Okada S, Gordon DM (2001) Host and geographical factors influence the thermal niche of enteric bacteria isolated from native Australian mammals. Molecular Ecology 10:2499-2513.

Okada S, Gordon DM (2003) Genetic and ecological structure of *Hafnia alvei* in Australia. Systematic and Applied Microbiology 26:585-594.

Posada D, Crandall KA (1998) MODELTEST: testing the model of DNA substitution. Bioinformatics 14:817-818.

Proietti PC, Passamonti F, Franciosini MP, Asdrubali G (2004) *Hafnia alvei* infection in pullets in Italy. Avian Pathology 33:200-204.

Pupo GM, Karaolis DKR, Lan RT, Reeves PR (1997) Evolutionary relationships among pathogenic and nonpathogenic *Escherichia coli* strains inferred from multilocus enzyme electrophoresis and *mdh* sequence studies. Infection and Immunity 65:2685-2692.

Pupo GM, Lan RT, Reeves PR, Baverstock PR (2000) Population genetics of *Escherichia coli* in a natural population of native Australian rats. Environmental Microbiology 2:594-610.

Ramos A, Damaso D (2000) Extraintestinal infection due to *Hafnia alvei*. European Journal of Clinical Microbiology and Infectious Diseases 19:708-710.

Real F, Fernandez A, Acosta F, Acosta B, Castro P, Deniz S, Oros J (1997) Septicemia associated with *Hafnia alvei* in laying hens. Avian Diseases 41:741-747.

Ridell J, Siitonen A, Paulin L, Lindroos O, Korkeala H, Albert MJ (1995) Characterization of *Hafnia alvei* by biochemical tests, random amplified polymorphic DNA PCR, and partial sequencing of 16S rRNA gene. Journal of Clinical Microbiology 33:2372-2376.

Sarkar SF, Guttman DS (2004) Evolution of the core genome of *Pseudomonas syringae*, a highly clonal, endemic plant pathogen. Applied and Environmental Microbiology 70:1999-2012.

Spratt BG, Hanage WP, Feil EJ (2001) The relative contributions of recombination and point mutation to the diversification of bacterial clones. Current Opinion in Microbiology 4:602-606.

Steigerwalt AG, Fanning GR, Fife-Asbury MA, Brenner DJ (1976) DNA relatedness among species of *Enterobacter* and *Serratia*. Canadian Journal of Microbiology 22:121-137.

Suerbaum S, Smith JM, Bapumia K, Morelli G, Smith NH, Kunstmann E, Dyrek I, Achtman M (1998) Free recombination within *Helicobacter pylori*. Proceedings of the National Academy of Sciences of the United States of America 95:12619-12624.

Swofford DL (2003) PAUP*. Phylogenetic Analysis Using Parsimony (*and Other Methods). Sinauer Associates, Sunderland, Massachusetts.

Thampapillai G, Lan RT, Reeves PR (1994) Molecular evolution in the *gnd* locus of *Salmonella enterica*. Molecular Biology and Evolution 11:813-828.

Urwin R, Maiden MCJ (2003) Multi-locus sequence typing: a tool for global epidemiology. Trends in Microbiology 11:479-487.

Vandamme P, Pot B, Gillis M, Devos P, Kersters K, Swings J (1996) Polyphasic taxonomy, a consensus approach to bacterial systematics. Microbiological Reviews 60:407-438.

Vulic M, Dionisio F, Taddei F, Radman M (1997) Molecular keys to speciation - DNA polymorphism and the control of genetic exchange in enterobacteria. Proceedings of the National Academy of Sciences of the United States of America 94:9763-9767.

Wertz JE, Goldstone C, Gordon DM, Riley MA (2003) A molecular phylogeny of enteric bacteria and implications for a bacterial species concept. Journal of Evolutionary Biology 16:1236-1248.

Wheeler WC (1990) Nucleic acid sequence phylogeny and random outgroups. Cladistics - the International Journal of the Willi Hennig Society 6:363-367.

Whittam TS, Ochman H, Selander RK (1983) Multilocus genetic structure in natural populations of *Escherichia coli*. Proceedings of the National Academy of Sciences of the United States of America 80:1751-1755.

Chapter 6

Comparative analyses on the population genetics of the two *Hafnia* groups

Introduction

The evolutionary mechanism of species diversity in bacteria is unique in space and time. A broad spectrum of genetic diversity can be seen in the bacterial world; from species with a highly clonal structure like *Salmonella enterica* (Beltran *et al.* 1991) to others with a panmictic structure such as *Helicobacter pylori* (Suerbaum *et al.* 1998). This happens because different species have different modes of creating genetic diversity. A totally clonal population happens when mutation is the driving force of divergence. On the other hand, a panmictic population structure is created by promiscuous intra-species recombination (Maynard Smith *et al.* 1993). However, the amount of genetic diversity that is observed at any one time in a bacterial population due to mutation can be as high (or low) as that caused by recombination.

Escherichia coli has been extensively studied in the context of bacterial population biology. Multilocus enzyme electrophoresis studies showed a highly clonal population structure in this species (Ochman *et al.* 1983; Whittam *et al.* 1983). With the advance in molecular genetic techniques in the early 1990s, many studies have detected evidence for frequent intra- and inter-genic recombination (Dykhuizen and Green 1991; Guttman and Dykhuizen 1994; Nelson and Selander 1994). How can a species with frequent recombination maintain a clonal population structure? Cohan (2001) suggested the presence of 'ecological structure' in bacterial species. He argues that bacterial species are comprised of sequence clusters, each representing an ecotype. Each ecotype occupies a unique niche that strains in other niches can not, and therefore homologous recombination and periodic selection are limited to strains within a given niche.

Okada and Gordon (2003) recently reinvestigated the population genetics and ecology of *Hafnia alvei* (family Enterobacteriaceae) (Moeller 1954). This species is distantly related to *E. coli* (Brenner 1984). It is found in the faeces of a wide range of vertebrate and invertebrate hosts as well as in the external environment of the host (Allen *et al.* 1983; Brenner 1984; Okada and Gordon 2003; Shirey and Bissonnette 1992). *H. alvei* consists of two genetically distinct groups that can be distinguished by several biochemical properties (Janda *et al.* 2002; Okada and Gordon 2003; Ridell *et al.* 1995; Steigerwalt *et al.* 1976). Multilocus sequence typing with partial sequences of six housekeeping genes found that there were no alleles shared between the two genetic groups (Okada chapter 5). Along with the fact that the two genetic groups were phenotypically different (Janda *et al.* 2002; Okada and Gordon 2003), we concluded that the two *Hafnia* groups represent separate species (Okada chapter 5).

Both allozyme analysis (Okada and Gordon 2003) and sequence variation of the housekeeping genes (Okada chapter 5) showed that genetic group 1 strains were more genetically diverse compared to group 2 strains. However, the evolutionary forces that create the observed structure are unknown. Is it recombination or mutation in these populations that drive genetic diversification? Can the ecological structure that was found in allozyme polymorphism be detected in neutral sequence variation? What forces of cohesion can be seen that maintain species integrity? We conducted a series of phylogenetic analyses with the sequences collected for multilocus sequence typing (MLST) (Maiden *et al.* 1998) to find the answers to these question and contrast it to what is currently known concerning the evolution of *E. coli*. This paper covers the evolutionary aspect from the results of MLST; the systematics aspect of the genus *Hafnia* is described elsewhere (Okada chapter 5).

Material and methods

DNA sequences

The data set comprised ninety-six isolates of *Hafnia*, 48 each from the two species, (Okada chapter 5) chosen at random from a larger set of strains. Partial sequences of six housekeeping genes (*gapA*, *gnd*, *groEL*, *gyrA*, *gyrB*, and *infB*) were used for the following analyses. Strain information is provided elsewhere (Okada chapter 5). Sequence type and allele designations of the individual strains are given in Tables 1 and 2.

Phylogenetic analyses

Each unique sequence was given an allele designation, from which allelic and haplotype diversity estimates were calculated using GenAlEx v5 (Peakall and Smouse 2001). A standardised index of association (stnd.I_A) was calculated using LIAN 3.0 (Haubold and Hudson 2000). This association index quantifies the degree to which alleles at different loci are non-randomly associated. G+C content was calculated by S.T.A.R.T. (Jolley *et al.* 2001).

Recombination estimates were calculated using three separate methods. The homoplasy test (Maynard Smith and Smith 1998) as implemented in the S.T.A.R.T. package calculates a probability of recombination that is relatively accurate for sequences that have low diversity (1 - 5 %) when compared to other algorithms. It compares values between the number of steps in a most-parsimonious tree and the number of polymorphic sites in the third codon of a set of sequences to find

Table 1. Strain identification and allele designations for the six genes of *Hafnia* group 1. ID, identification; ST, sequence type.

ID	*Hafnia* group	Host class	ST	*gapA*	*gnd*	*groEL*	*gyrA*	*gyrB*	*infB*
MI590	1	Mammalia	36	102	103	118	107	101	119
F169	1	Actinopterygii	37	120	106	115	108	102	124
R373	1	Reptilia	38	104	107	109	116	103	127
R022	1	Reptilia	39	120	108	122	115	104	125
F138	1	Actinopterygii	40	113	121	112	104	105	116
B642	1	Aves	41	114	122	112	104	105	117
F182	1	Actinopterygii	42	120	109	122	109	106	126
R015	1	Reptilia	43	118	113	114	102	107	129
B348	1	Aves	44	111	120	112	105	107	129
M082	1	Mammalia	45	119	114	107	106	107	122
MI490	1	Mammalia	45	119	114	107	106	107	122
R012	1	Reptilia	45	119	114	107	106	107	122
MI782	1	Mammalia	46	106	104	105	121	108	121
R044	1	Reptilia	47	115	129	103	111	109	102
MI785	1	Mammalia	48	115	129	101	120	110	102
R027	1	Reptilia	48	115	129	101	120	110	102
L069	1	Amphibia	49	115	128	104	111	111	102
I128	1	Insecta	50	122	131	104	111	111	102
R285	1	Reptilia	50	122	131	104	111	111	102
B429	1	Aves	51	116	128	104	112	111	102
M129	1	Mammalia	52	121	111	101	110	112	103
E072	1	water	53	113	120	107	104	113	116
MI789	1	Mammalia	53	113	120	107	104	113	116
R055	1	Reptilia	54	111	132	116	104	113	129
R023	1	Reptilia	55	111	133	116	105	113	105
MI777	1	Mammalia	56	109	119	110	118	113	123
R180	1	Reptilia	57	109	119	110	118	113	122
MI581	1	Mammalia	58	108	119	102	120	113	115
M051	1	Mammalia	59	120	130	104	111	114	102
I040	1	Insecta	60	105	125	121	117	115	107
B459	1	Aves	61	107	112	126	112	116	108
R126	1	Reptilia	62	117	123	110	111	117	113
R202	1	Reptilia	63	102	123	110	122	117	110
L009	1	Amphibia	64	115	116	110	114	118	106
F156	1	Actinopterygii	65	101	117	106	113	119	104
R021	1	Reptilia	66	102	126	110	124	120	110
R067	1	Reptilia	67	102	127	123	112	121	111
R114	1	Reptilia	68	115	101	120	119	121	110
R124	1	Reptilia	68	115	101	120	119	121	110
M172	1	Mammalia	69	102	102	117	114	122	114
MI524	1	Mammalia	70	102	118	119	101	123	112
MI472	1	Mammalia	71	102	115	119	103	123	107
M044	1	Mammalia	72	111	120	113	106	124	101
M306	1	Mammalia	73	102	125	111	111	125	109
M061	1	Mammalia	74	103	124	124	111	126	128
R211	1	Reptilia	75	102	105	125	123	127	118
R394	1	Reptilia	75	102	105	125	123	127	118
F187	1	Actinopterygii	76	102	110	108	123	128	120

Table 2. Strain identification and allele designations for the six genes of *Hafnia* group 2. ID, identification; ST, sequence type.

ID	*Hafnia* group	Host class	ST	*gapA*	*gnd*	*groEL*	*gyrA*	*gyrB*	*infB*
MI231	2	Mammalia	1	4	23	3	4	1	3
F311	2	Actinopterygii	2	4	8	4	4	2	9
MI765	2	Mammalia	3	4	11	8	4	2	14
MI177	2	Mammalia	4	4	11	8	4	2	15
MI240	2	Mammalia	4	4	11	8	4	2	15
R175	2	Reptilia	4	4	11	8	4	2	15
R048	2	Reptilia	5	4	14	6	4	3	19
R184	2	Reptilia	5	4	14	6	4	3	19
R054	2	Reptilia	6	12	14	6	4	3	8
R053	2	Reptilia	7	12	14	6	4	3	19
R057	2	Reptilia	7	12	14	6	4	3	19
R061	2	Reptilia	7	12	14	6	4	3	19
L013	2	Amphibia	8	13	18	11	4	4	2
MI468	2	Mammalia	9	4	8	4	4	5	12
MI293	2	Mammalia	10	4	15	13	4	5	12
MI319	2	Mammalia	10	4	15	13	4	5	12
R117	2	Reptilia	11	4	1	9	2	6	12
R122	2	Reptilia	12	4	9	4	4	7	18
R119	2	Reptilia	13	3	16	7	5	8	10
R252	2	Reptilia	13	3	16	7	5	8	10
MI230	2	Mammalia	14	1	17	11	5	9	18
R231	2	Reptilia	15	4	20	11	4	10	11
B255	2	Aves	16	8	21	4	3	11	18
MI690	2	Mammalia	16	8	21	4	3	11	18
MI398	2	Mammalia	17	4	26	1	4	11	10
F321	2	Actinopterygii	18	4	13	4	5	11	20
MI387	2	Mammalia	19	4	7	4	6	11	9
MI408	2	Mammalia	20	7	27	4	5	12	2
MI761	2	Mammalia	20	7	27	4	5	12	2
MI763	2	Mammalia	20	7	27	4	5	12	2
F292	2	Actinopterygii	21	4	24	10	1	13	2
F348	2	Actinopterygii	22	6	3	12	9	14	5
F296	2	Actinopterygii	23	11	5	12	7	15	4
M132	2	Mammalia	24	7	12	3	4	16	16
F152	2	Actinopterygii	25	4	25	4	5	17	2
MI476	2	Mammalia	26	4	19	4	11	18	12
MI164	2	Mammalia	27	4	22	4	5	19	2
F397	2	Actinopterygii	28	13	10	2	4	20	15
F121	2	Actinopterygii	29	13	10	5	4	20	15
H698	2	Mammalia	29	13	10	5	4	20	15
F118	2	Actinopterygii	30	2	27	9	4	20	18
MI470	2	Mammalia	31	9	9	13	4	20	9
MI463	2	Mammalia	32	5	2	12	10	21	6
MI464	2	Mammalia	32	5	2	12	10	21	6
F231	2	Actinopterygii	33	9	15	4	5	22	17
MI310	2	Mammalia	34	6	4	12	8	23	1
F298	2	Actinopterygii	35	10	6	12	10	23	7
F364	2	Actinopterygii	35	10	6	12	10	23	7

significantly higher numbers of apparent homoplasies. If the estimated distribution of homoplasies with no recombination does not include the number of apparent homoplasies, then this would suggest the occurrence of recombination. The second program, developed by Kuhner *et al.* (RECOMBINE) (Kuhner *et al.* 2000) uses a maximum likelihood approach to estimate $\theta = 2N_e\mu$ (the per-site mutation rate, μ, times the effective population size, N_e) and $r = C/\mu$ (the per-site recombination to mutation ratio). The third program, LDhat (McVean *et al.* 2002) uses a permutation approach, based on coalescent theory, to detect recombination in genes that are known to have relatively high rates of mutation. LDhat requires an estimate for θ, and the estimate used was that calculated by RECOMBINE. In order to visualise conflicts in phylogenetic signals, split decomposition analysis (Bandelt 1992) was conducted for each of the individual gene trees and the tree based on the concatenated sequences using Splitstree v3.2 (website: http://www.bibiserv.,techfak.uni-bielefeld.de/splits; (Huson 1998)). Genetic variation of strains that is not supported by a unique tree will result in a network-like relationship.

Congruence between gene genealogies was assessed by three methods. The first was the incongruence length difference (ILD) test (Farris *et al.* 1994). This test determines if there are significant differences between the branch lengths of maximum parsimony trees based on two different genes. All pair-wise comparisons of individual gene trees were tested using the partition homogeneity test as implemented in PAUP*. Identical sequences were omitted from the data sets, and 100 replicates with a maximum of 1000 saved trees were compared. Because there are arguments concerning the excessive sensitivity resulting in type 1 error of the ILD test (Barker and Lutzoni 2002), the Shimodaira-Hasegawa test (SH test; in PAUP*) was also conducted as a comparison. In this test, the maximum likelihood trees of each of the six genes were compared to the sequence data for the other five genes in order to assess the extent to which the topology

of each tree matched the sequence variation in the other genes (Shimodaira and Hasegawa 1999). Finally, bootstrap analyses (n = 100) of the maximum likelihood gene trees were conducted using PAUP*. The strain membership of clusters with over 70 % bootstrap support were compared among the 6 gene trees in order to identify those strains that did not consistently cluster in each of the gene trees.

For each of the genes and the concatenated data set, analysis of molecular variance (AMOVA) was used to assess the extent to which host taxonomic class explained the observed among strain variation. The AMOVA routine implemented in GenAlEx v5 (Peakall and Smouse 2001) was used for these analyses. Due to the small sample sizes for some host groups, only isolates from fish, mammals, reptiles and birds (bird isolates were used in the analyses for group 1 isolates only) were included in the analyses. The relationships among isolates from different taxonomic groups were visualised using principal co-ordinates analysis as implemented in GenAlEx v5.

Results

Allelic variation in *Hafnia* group 1

The number of alleles per locus for group 1 isolates ranged from 20 (*gapA*) to 33 (*gnd*), and alleles were non-randomly associated among loci (stnd $I_A = 0.24$, $p < 0.000$)(Tables 3, 4). There were 41 unique allelic profiles among the 48 strains, indicating a high degree of haplotype diversity ($G = 0.992 \pm 0.0007$; Table 4). The number of polymorphic sites per locus varied from 22 for *gapA* and *gyrA* to 71 for *gnd* (Table 3). The average G+C content of the six genes was 49.8 %.

Table 3. Nucleotide polymorphism of the six genes in the two *Hafnia* groups.

Gene	Sequence length	variable sites		informative sites		no. alleles		dn/ds	
		group 1	group 2	group 1	group 2	group 1	group 2	group1	group2
gapA	692	22	11	10	9	20	13	0.019	0.023
gnd	598	71	43	54	28	33	27	0.006	0.009
groEL	561	24	12	19	10	26	13	0.016	0.019
gyrA	711	22	14	13	8	24	11	0.013	0.000
gyrB	577	47	25	33	20	28	23	0.000	0.000
infB	542	35	24	21	15	29	19	0.021	0.016

Table 4. Genetic diversity of the six genes and standardised index of association in the two *Hafnia* groups. P, significance for the test of the null hypothesis H_0:$V_D = V_e$ by a parametric approach.

	group 1		group 2	
n	48	41	48	35
stnd.IA	0.24	0.11	0.20	0.07
P	0.00	0.00	0.00	0.00
genetic diversity				
mean	0.95	0.96	0.87	0.87
gapA	0.91	0.92	0.79	0.76
gnd	0.98	0.99	0.97	0.98
groEL	0.96	0.96	0.87	0.87
gyrA	0.95	0.96	0.69	0.71
gyrB	0.96	0.97	0.95	0.97
infB	0.96	0.97	0.93	0.95

Allelic variation in *Hafnia* group 2

Compared to the group 1 strains, there were fewer alleles per locus for group 2 isolates (11 (*gapA*) to 27 (*gnd*); Table 3), but as in group 1, alleles were non-randomly associated among loci (stnd $I_A = 0.20$, $p < 0.000$)(Table 4). There were 35 unique allelic profiles among the 48 strains, indicating a lower degree of haplotype diversity compared to group 1 strains ($G = 0.986 \pm 0.001$; Table 4). The number of polymorphic sites per locus varied from 11 for *gapA* to 43 for *gnd* (Table 3). The average G+C content of the six genes was 50.4 % and this value was significantly different from the G+C content of the group 1 strains (AOV:, $F_{(1,1)} = 408.7$, $p < 0.0001$)

Intragenic recombination estimates of the six genes in the two *Hafnia* groups

Comparison of the r/μ estimates (RECOMBINE) between genetic group 1 and group 2, revealed that the r/μ estimates for each of the genes were greater for group 2 isolates, with *gyrA* being the sole exception. The per region recombination rate estimates, ρ (LDhat) showed a similar pattern, with generally higher estimates of ρ for group 2 isolates (exceptions *gyrA* and *gyrB*). Homoplasy ratio estimates were determined only for genes with more than 10 informative sites (Table 3). Significant homoplasy estimates suggested intragenic recombination occurs in all genes (Table 5). Contrasting the estimates derived for group 1 and group 2 strains showed evidence for intragenic recombination in *gnd* for group 2 but not group 1 strains, in *gyrB* for both groups, and in *infB* for group 1 but not group 2 strains (Table 5).

Table 5. Intragenic recombination rate estimates and probability of recombination of the two *Hafnia* groups. Homoplasy ratios were calculated for genes with more than 10 parsimony informative sites. Watterson, Watterson estimate of theta. Significant homoplasy ratios ($P < 0.05$) are indicated in bold-face.

gene	genetic group	Watterson	Theta#	r/mu#	2Ner#	LDhat rho	homoplasy
gapA	1	0.007	0.008	1.494	0.012	16.33	na
	2	0.004	0.004	2.994	0.013	42.52	na
gnd	1	0.027	0.061	0.175	0.011	8.50	0.03
	2	0.016	0.030	0.531	0.016	11.56	**0.31**
groEL	1	0.010	0.014	1.416	0.020	36.73	**0.55**
	2	0.005	0.005	1.970	0.010	17.35	na
gyrA	1	0.007	0.011	1.326	0.014	21.43	**0.32**
	2	0.004	0.004	0.000	0.000	0.00	na
gyrB	1	0.018	0.040	0.318	0.013	17.35	**0.24**
	2	0.010	0.016	1.129	0.018	8.16	**0.49**
infB	1	0.015	0.026	0.026	0.001	3.06	**0.13**
	2	0.010	0.014	0.967	0.013	4.08	0.06

#, estimates calculated by RECOMBINE.

Split decomposition analysis of the six genes suggested potential intragenic recombination events (as indicated by the network-like relationship between strains) in

all genes but *infB* for genetic group 1 strains, whereas for group 2 strains *gapA* and *groEL* were the only genes for which recombination events were suggested by the analysis (Figures 1, 2).

Intergenic recombination estimates of the two genetic groups of *Hafnia*

Bootstrap scores for the individual maximum likelihood gene trees using only the unique allele set were calculated for each *Hafnia* group. There were two group 1 strains (F156, MI524) that occurred in different clades with > 70 % bootstrap support in the maximum likelihood trees for *gnd* and *gyrA* (Figure 3). Among group 2 strains, there was no evidence for conflict among the allelic clusters that had >70 % bootstrap support for any of the six genes (Figure 4).

Congruence between gene trees within each *Hafnia* group was assessed by pairwise comparisons of the six gene trees using the incongruence length difference (ILD) test. The ILD test results indicated significant incongruence between all gene pairs for genetic group 1 strains, whereas over half of the gene tree comparisons were congruent for the group 2 strains (Table 6). The Shimodaira-Hasegawa (SH) test also indicated significant incongruence for all gene and gene tree comparisons involving group 1 strains. Comparison of the individual genes with the concatenated gene tree for group 1 strains indicated significant incongruence for *groEL*, *gyrA* and *infB*. although the probability values for *groEL* and *infB* were marginal (Table 7). For group 2, all gene - gene tree comparisons were incongruent, with the exception of most comparisons involving *gyrA* (Table 8). Group 2 gene – concatenated tree comparisons only revealed significant incongruence for the *gnd* data and statistical borderline incongruence for the *groEL* gene data (Table 8).

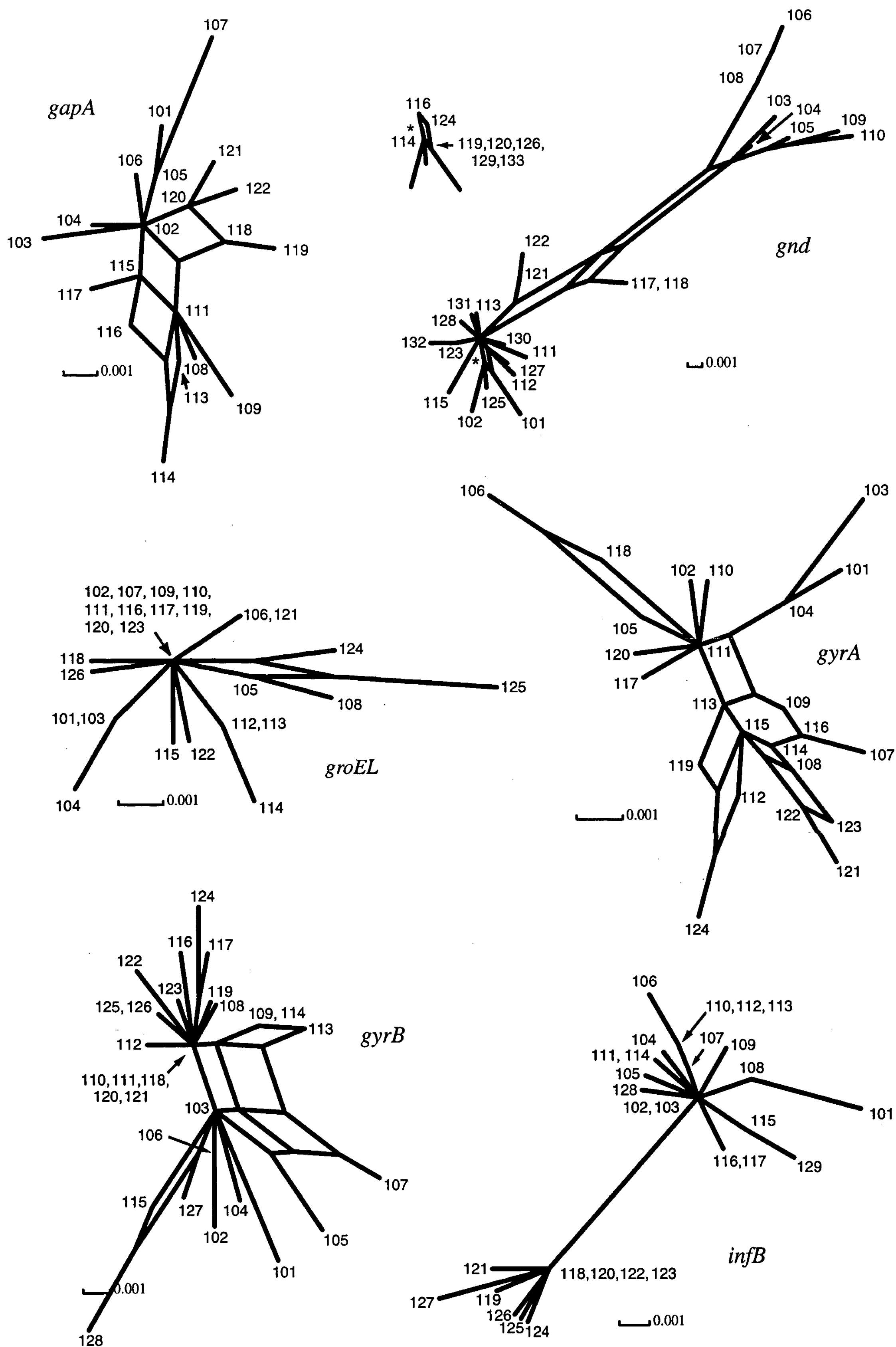


Figure 1. Split decomposition analysis of the six genes of *Hafnia* group 1 strains (n = 48). Parsimony, hamming distances. Numbers on each node refer to the designated allele numbers as in Table 1.

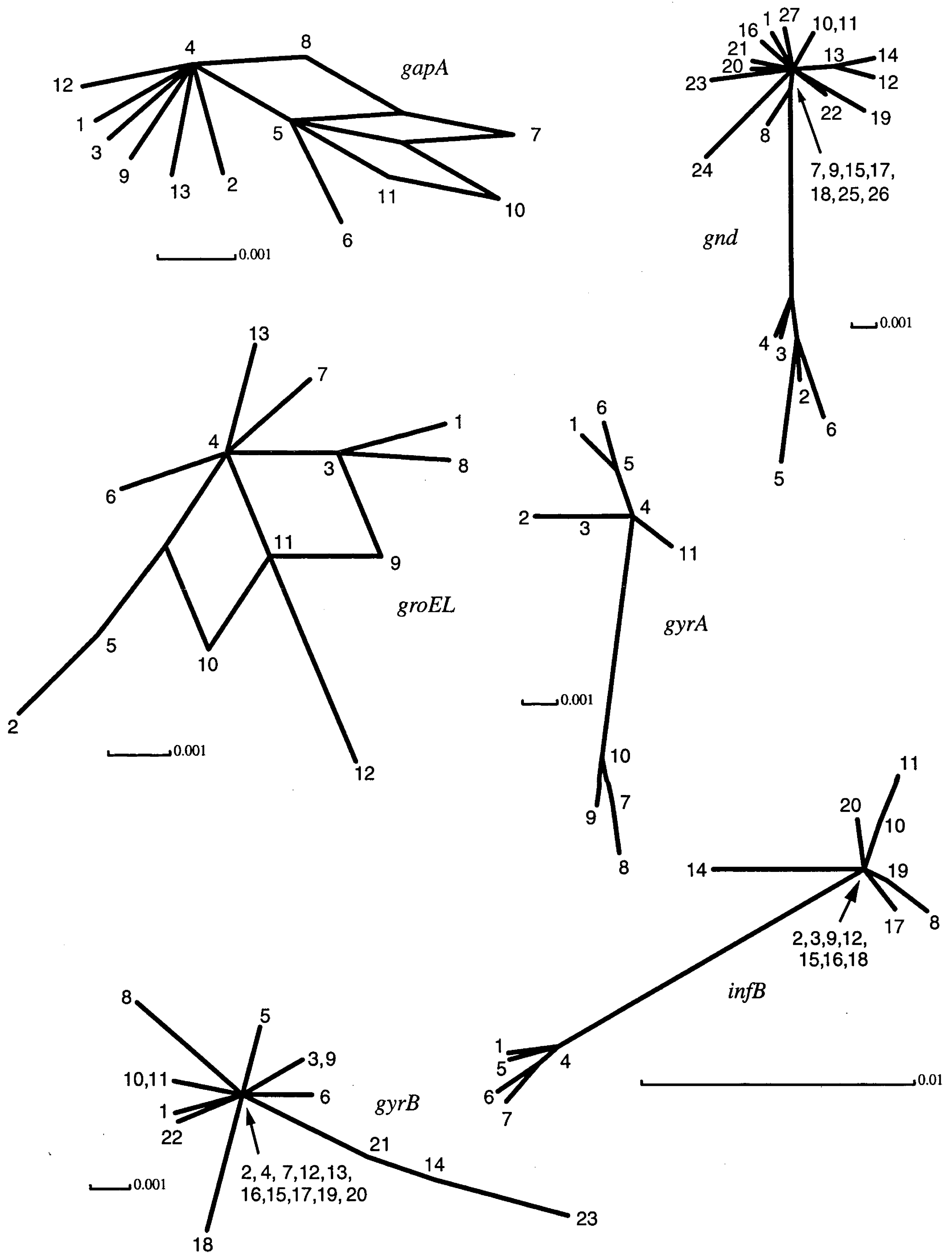


Figure 2. Split decomposition analysis of the six genes of *Hafnia* group 2 strains (n=48). Parsimony, hamming distances. Numbers on each node refer to the designated allele numbers as in Table 2.

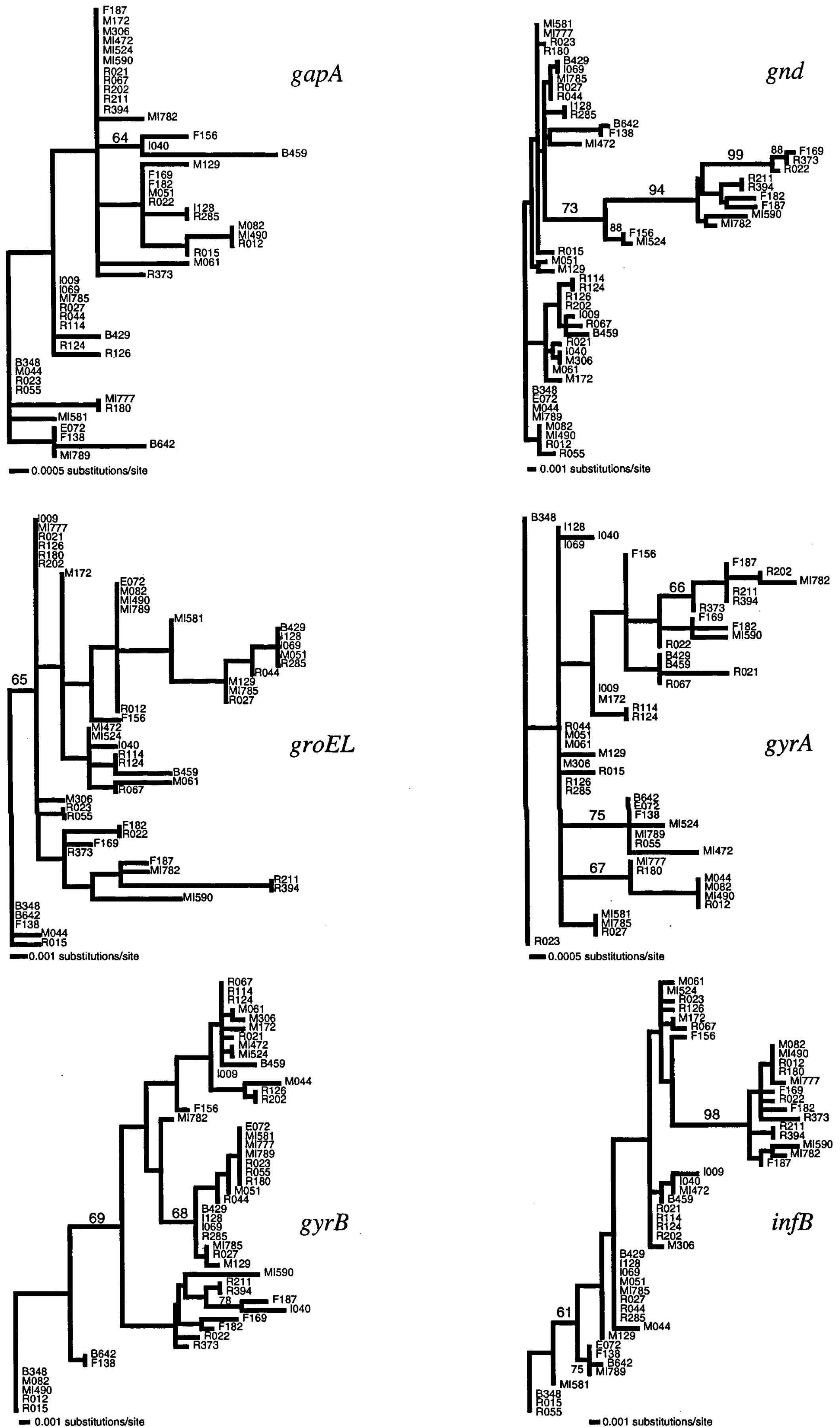


Figure 3. Maximum likelihood trees of the six genes for *Hafnia* group 1 strains. Bootstrap scores > 60 are indicated.

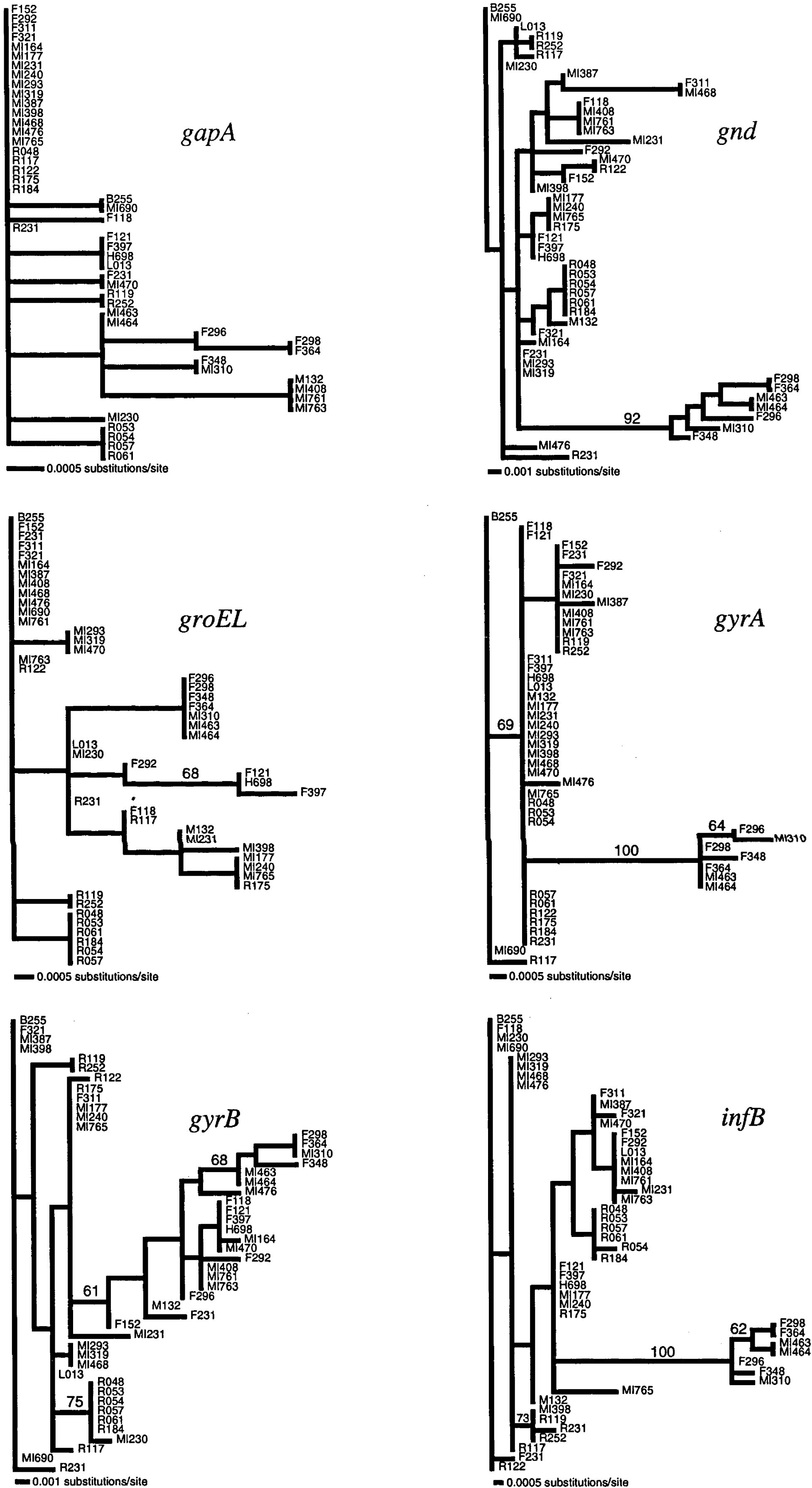


Figure 4. Maximum likelihood trees of the six genes for *Hafnia* group 2 strains. Bootstrap scores > 60 are indicated.

Split decomposition analysis using the concatenated sequences provided little evidence for the presence of recombination events. Both *Hafnia* groups were similar in overall structure, with one outgroup that was genetically distinct from the other strains that composed a star phylogeny (Figure 5).

Table 6. Incongruence length difference (ILD) test of the two *Hafnia* groups. Probability values above the diagonal are the pairwise comparisons of group 1 strains, and those below are of group 2 strains.

gene	*gapA*	*gnd*	*groEL*	*gyrA*	*gyrB*	*infB*
gapA	-	0.04	0.05	0.01	0.01	0.01
gnd	0.24	-	0.01	0.01	0.01	0.01
groEL	0.38	0.23	-	0.01	0.01	0.01
gyrA	0.04	0.60	0.51	-	0.01	0.01
gyrB	0.22	0.01	0.03	0.23	-	0.01
infB	0.02	0.01	0.01	0.14	0.01	-

Table 7. Shimodaira-Hasegawa (SH) test of *Hafnia* group 1 strains. Significant incongruence between a gene and maximum likelihood gene tree (P > 0.05) are bold-faced.

	maximum likelihood tree						
gene	*gapA*	*gnd*	*groEL*	*gyrA*	*gyrB*	*infB*	concatenated
gapA	-	0.020	0.017	0.017	0.021	0.025	0.193
gnd	0.00	-	0.000	0.000	0.000	0.000	0.214
groEL	0.00	0.000	-	0.000	0.000	0.000	0.032
gyrA	0.00	0.000	0.000	-	0.001	0.000	0.002
gyrB	0.00	0.000	0.000	0.000	-	0.000	0.096
infB	0.00	0.000	0.000	0.000	0.000	-	0.036

Table 8. Shimodaira-Hasegawa (SH) test of *Hafnia* group 2 strains. Significant incongruence between a gene and maximum likelihood gene tree (P > 0.05) are bold-faced.

	maximum likelihood tree						
gene	*gapA*	*gnd*	*groEL*	*gyrA*	*gyrB*	*infB*	concatenated
gapA	-	**0.009**	**0.012**	**0.007**	**0.011**	**0.011**	0.144
gnd	**0.000**	-	**0.001**	**0.000**	**0.001**	**0.000**	**0.026**
groEL	**0.001**	**0.005**	-	**0.000**	**0.002**	**0.003**	**0.035**
gyrA	0.062	0.083	0.087	-	**0.047**	0.087	0.228
gyrB	**0.000**	**0.000**	**0.000**	**0.000**	-	**0.000**	0.405
infB	**0.000**	**0.003**	**0.000**	**0.000**	**0.002**	-	0.060

Effect of host taxonomic class on the genetic structure of the two *Hafnia* groups

Analysis of molecular variance (AMOVA) revealed a striking difference between the two genetic groups in the effect the host had on their genetic structures (Table 9). Among *Hafnia* group 1 strains, the taxonomic class of the host explained a significant amount of the among strain variation for three of the six genes and extent of the explained variation ranged from 5 % (*gyrA*, *gyrB*) to 12 % (*gnd*). For *Hafnia* group 2 strains host taxonomic class accounted for a significant fraction of the variation in five of the genes and the extent of the variation ranged from 8 % (*infB*) to 20 % (*gyrB*). Overall, host class explained a consistently greater amount of the observed genetic variation in group 2 strains, as compared to group 1 strains.

Table 9. Analysis of molecular variance. The percentage of the genetic variation explained by host taxonomic class was calculated for each gene within each of the *Hafnia* groups. Significant percentages (P < 0.05) are bold-faced.

gene	% group 1 (P > F)	% group 2 (P > F)
gapA	2 (0.12)	**11 (0.005)**
gnd	**12 (0.001)**	**11 (0.004)**
groEL	0 (0.45)	**12 (0.002)**
gyrA	**5 (0.02)**	6 (0.08)
gyrB	**5 (0.02)**	**20 (0.001)**
infB	2 (0.20)	**8 (0.04)**
concatenated	**6 (0.002)**	**12 (0.004)**

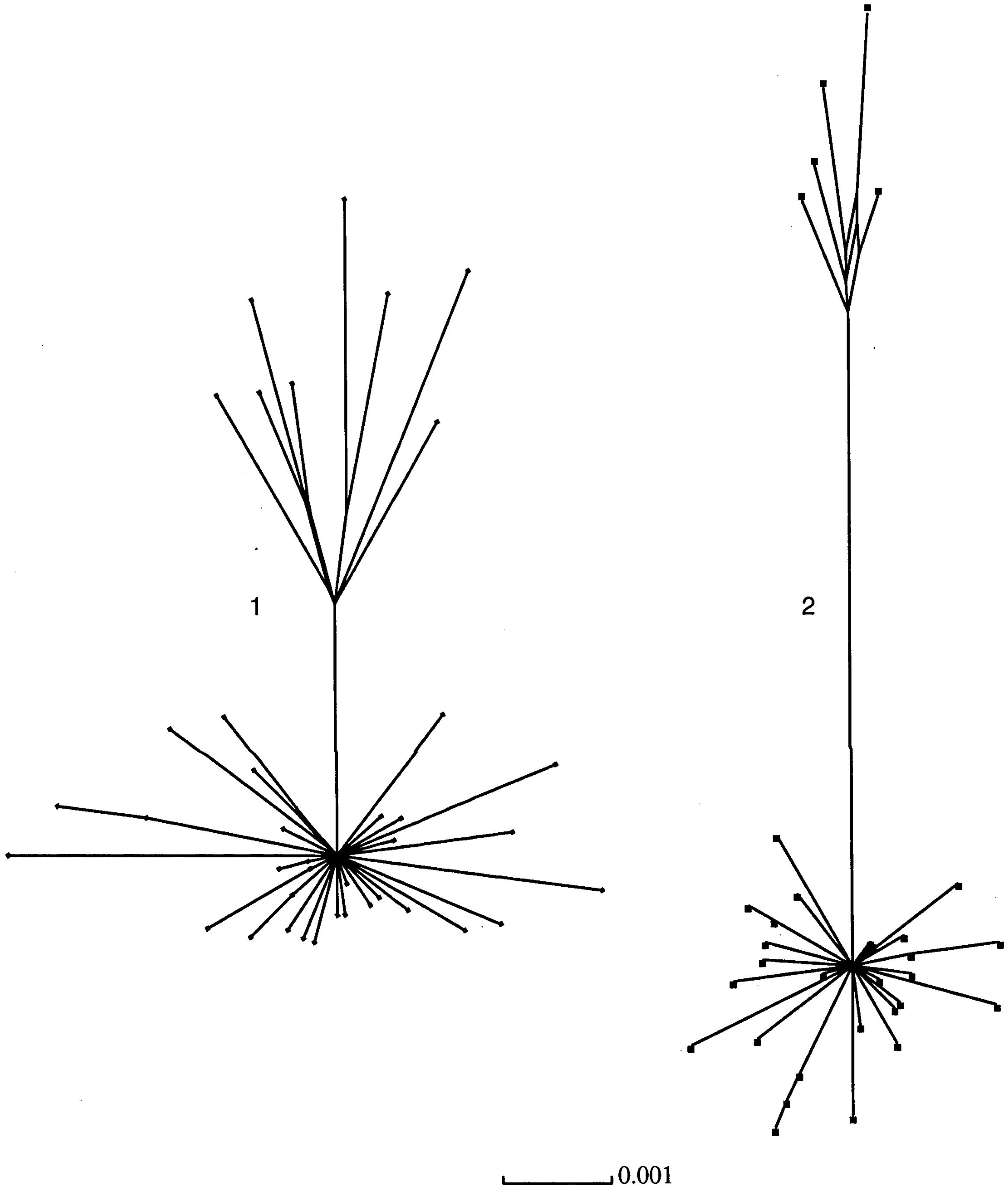


Figure 5. Split decomposition analysis of the concatenated sequence of *Hafnia* group 1 (1) and group 2 (2). Parsimony, hamming distances.

In order to eliminate a bias due to the presence of multiple clones in each of the six loci, strains with identical sequences were excluded and AMOVA was run on the non-redundant sequence set (Table 10). Host class explained twice as much of the genetic variation in *gyrB* among group 1 strains in comparison to the complete sequence set. Among the group 2 strains, significant host class effects were only detected for *groEL* and *gyrB* (Table 10). These results indicate that for *Hafnia* group 1 strains, the same *gyrB* alleles are distributed across different host classes, whereas for group 2 strains, the distribution of identical alleles tended to be limited to within host classes for *gapA*, *gnd*, *gyrA* and *infB*.

Table 10. Analysis of molecular variance with the non-redundant data set for each *Hafnia* group. The percentage of the genetic variation explained by host taxonomic class was calculated for each gene within each of the *Hafnia* groups. Significant percentages ($P < 0.05$) are bold-faced.

gene	% group 1 (P > F)	% group 2 (P > F)
gapA	0 (0.67)	0 (0.43)
gnd	**9 (0.02)**	0 (0.38)
groEL	0 (0.99)	**20 (0.05)**
gyrA	5 (0.09)	0# (0.81)
gyrB	**10 (0.02)**	**21 (0.019)**
infB	7 (0.13)	0 (0.42)
concatenated	**7 (0.01)**	6 (0.08)

#, only one sequence type for reptilian strains

The concatenated data set was analysed using AMOVA and the pair wise population (host class) differentiation estimates used in a principal coordinates analysis (PCA). Both *Hafnia* groups had a significant amount of their overall genetic variation explained by host class (Table 9). However, the host effect on the overall genetic variation was insignificant in the non-redundant data set for group 2 strains, again illustrating the

distribution of identical haplotypes within the same host class among strains in *Hafnia* group 2 (Table 10). A plot of the first two PCA axes showed that for group 2, strains from reptiles were different from strains from mammals or fish, whilst for group 1, strains from mammals and reptiles were similar but different from those isolated from birds or fish (Figure 6).

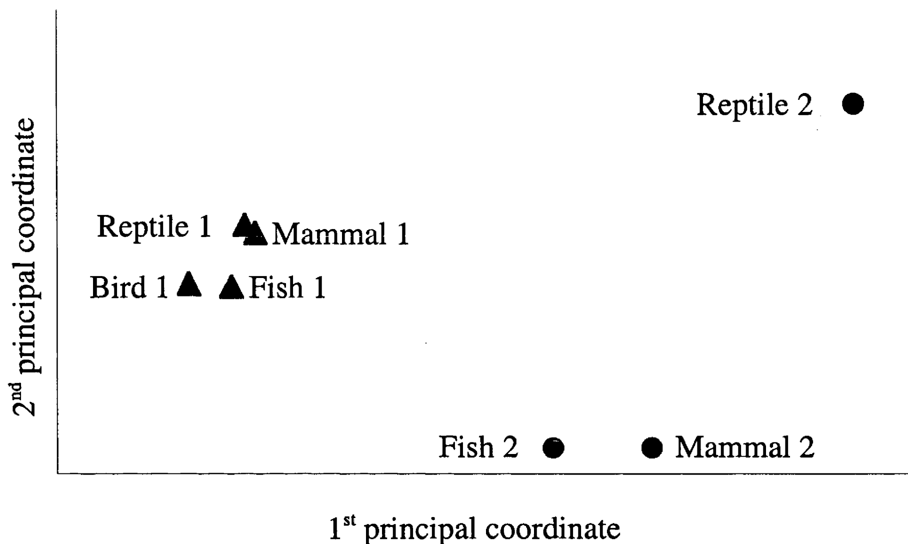


Figure 6. Principal coordinates analysis of the genetic distance matrix for *Hafnia* groups 1 and 2 by host taxonomic class. The distance between each host class by genetic group indicates their relative degree of genetic relatedness. Each legend next to the plots indicates the host followed by the *Hafnia* group.

Discussion

Contributions of recombination and mutation in the two genetic groups of the genus *Hafnia*

Genome diversity and haplotype diversity in the two groups of *Hafnia* were considerably higher compared to other studies that have used MLST to observe the extent of genetic diversity in bacterial populations (Urwin and Maiden 2003). This is

partially due to the fact that the strains in this study were commensal isolates that were recovered from a diverse range of non-domesticated hosts collected across a continent. In contrast, the majority of MLST studies have investigated clinical isolates, which in many cases, only represent a single clonal population of the species (Feil and Enright 2004).

The different approaches used to examine the extent to which recombination shapes the genetic structure of the two *Hafnia* groups yielded conflicting results. The intragenic recombination to mutation ratio estimates calculated by two different algorithms (RECOMBINE and LDhat) suggest that recombination is more predominant within group 2 strains in comparison to group 1 strains. However, these results must be interpreted with caution, as estimates of r/μ are only valid when certain assumptions are met (Stumpf and McVean 2003). One of the crucial assumptions is that the variation observed in the genes under study is not a consequence of selection (McVean *et al.* 2002). The six genes examined in the present study are housekeeping genes and therefore it is generally thought that variation in these genes should reflect largely neutral variation. However, the observation that the type of host from which the strains were isolated explained a significant amount of the variation in most of these genes suggests that selection is shaping the observed sequence variation (Tables 9, 10). Given that significant host effects were detected more frequently in group 2 strains compared to group 1 strains, it is difficult to determine if the greater r/μ estimates for group 2 strains reflect higher rates of intragenic recombination or if it is a bias due to the consequences of selection. Although there was a lack of data for some genes, the results of the homoplasy ratio tests, whilst suggesting that recombination has occurred, did not indicate that the rate of intragenic recombination differed between *Hafnia* groups. Compared to the r/μ estimates, the split decomposition analysis suggested more intragenic recombination among strains of group 1 than those of group 2.

Analyses investigating intergenic recombination also gave rise to conflicting outcomes. The results of both the ILD tests and the SH tests are consistent with the idea that recombination shapes the genetic structure of *Hafnia*. Whilst the ILD test results suggests that group 2 strains have a more clonal genetic structure than do group 1 strains, the results of the SH tests provide little evidence that the importance of intergenic recombination differs among *Hafnia* groups.

The repeated isolation of identical haplotypes from different hosts or geographic locations is also a hallmark of a clonal genetic structure. For *Hafnia* group 2 strains, identical haplotypes were more common and often collected from different hosts or localities than was the case for group 1 strains. Thus the balance of the results would suggest that *Hafnia* group 2 strains are more clonal than *Hafnia* group 1 strains.

How does the genetic structure of the two *Hafnia* groups compare with other members of the Enterobacteriaceae? The degree of clonality in these two *Hafnia* groups, when compared using the standardised index of association, is similar to that of *E. coli* (stnd $I_A = 0.29$ with 197 non-pathogenic isolates) and $<$ that of *S. enterica* (stnd $I_A = 0.49$ with 249 isolates; based on the allelic profile provided by M. Achtman; website: http://web.mpiib-berlin.mpg.de/mlst/). Thus the two *Hafnia* species would be said to have a moderately clonal genetic structure (Maynard Smith *et al.* 1993).

Conclusion

The two *Hafnia* species were found to differ in their population genetic structure; group 2 strains constitute a more clonal structure in comparison to those of group 1. The degree of selective pressure that is created by the host is acting differently upon the two *Hafnia* groups, with group 2 strains sharing a greater amount of host adaptation than

group 1 strains. At present the importance of recombination versus mutation in shaping the genetic structure of these two *Hafnia* groups can not be determined. To better address the question, longer sequences containing more neutral genetic variation (i.e. regions that have a certain amount of functional constraint on the coding protein) is needed.

References

Allen DA, Austin B, Colwell RR (1983) Numerical taxonomy of bacterial isolates associated with a fresh water fishery. Journal of General Microbiology 129:2043-2062.

Bandelt H-J, Dress AWM (1992) Split decomposition: A new and useful approach to phylogenetic analysis of distance data. Molecular Phylogenetics and Evolution 1:242-252.

Barker FK, Lutzoni FM (2002) The utility of the incongruence length difference test. Systematic Biology 51:625-637.

Beltran P, Plock SA, Smith NH, Whittam TS, Old DC, Selander RK (1991) Reference collection of strains of the *Salmonella typhimurium* complex from natural populations. Journal of General Microbiology 137:601-606.

Brenner DJ (1984) Family I. Enterobacteriaceae. In: Krieg NR (ed) *Bergey's Manual of Systematic Bacteriology*. Williams and Wilkins, Baltimore, pp. 408-420.

Cohan FM (2001) Bacterial species and speciation. Systematic Biology 50:513-524.

Dykhuizen DE, Green L (1991) Recombination in *Escherichia coli* and the definition of biological species. Journal of Bacteriology 173:7257-7268.

Farris JS, Kallersjo M, Kluge AG, Bult C (1994) Testing significance of incongruence. Cladistics - the International Journal of the Willi Hennig Society 10:315-319.

Feil EJ, Enright MC (2004) Analyses of clonality and the evolution of bacterial pathogens. Current Opinion in Microbiology 7:308-313.

Guttman DS, Dykhuizen DE (1994) Clonal Divergence in *Escherichia coli* as a result of recombination, not mutation. Science 266:1380-1383.

Haubold B, Hudson RR (2000) LIAN 3.0: detecting linkage disequilibrium in multilocus data. Bioinformatics 16:847-848.

Huson DH (1998) Splitstree - analyzing and visualizing evolutionary data. Bioinformatics 14:68-73.

Janda JM, Abbott SL, Khashe S, Probert W (2002) Phenotypic and genotypic properties of the genus *Hafnia*. Journal of Medical Microbiology 51:575-580.

Jolley KA, Feil EJ, Chan MS, Maiden MCJ (2001) Sequence type analysis and recombinational tests (START). Bioinformatics 17:1230-1231.

Kuhner MK, Yamato J, Felsenstein J (2000) Maximum likelihood estimation of recombination rates from population data. Genetics 156:1393-1401.

Maiden MCJ, Bygraves JA, Feil E, Morelli G, Russell JE, Urwin R, Zhang Q, Zhou JJ, Zurth K, Caugant DA, Feavers IM, Achtman M, Spratt BG (1998) Multilocus sequence typing - a portable approach to the identification of clones within populations of pathogenic microorganisms. Proceedings of the National Academy of Sciences of the United States of America 95:3140-3145.

Maynard Smith J, Smith NH (1998) Detecting recombination from gene trees. Molecular Biology and Evolution 15:590-599.

Maynard Smith J, Smith NH, Orourke M, Spratt BG (1993) How clonal are bacteria? Proceedings of the National Academy of Sciences of the United States of America 90:4384-4388.

McVean G, Awadalla P, Fearnhead P (2002) A coalescent-based method for detecting and estimating recombination from gene sequences. Genetics 160:1231-1241.

Moeller V (1954) Distribution of amino acid decarboxylases in Enterobacteriaceae. Acta Pathologica Microbiologica Scandinavica 35:259-277.

Nelson K, Selander RK (1994) Intergeneric transfer and recombination of the 6-phosphogluconate dehydrogenase gene (*gnd*) in enteric bacteria. Proceedings of the National Academy of Sciences of the United States of America 91:10227-10231.

Ochman H, Whittam TS, Caugant DA, Selander RK (1983) Enzyme polymorphism and genetic population structure in *Escherichia coli* and *Shigella*. Journal of General Microbiology 129:2715-2726.

Okada S, Gordon DM (2003) Genetic and ecological structure of *Hafnia alvei* in Australia. Systematic and Applied Microbiology 26:585-594.

Peakall R, Smouse PE (2001) GenAlEx V5: Genetic Analysis in Excel. Population genetic software for teaching and research.

Ridell J, Siitonen A, Paulin L, Lindroos O, Korkeala H, Albert MJ (1995) Characterization of *Hafnia alvei* by biochemical tests, random amplified polymorphic DNA PCR, and partial sequencing of 16S rRNA gene. Journal of Clinical Microbiology 33:2372-2376.

Shimodaira H, Hasegawa M (1999) Multiple comparisons of log-likelihoods with applications to phylogenetic inference. Molecular Biology and Evolution 16:1114-1116.

Shirey JJ, Bissonnette GK (1992) Sheen formation and growth-response of groundwater bacteria to reduced oxygen concentrations during incubation of m-endo medium. Canadian Journal of Microbiology 38:261-266.

Steigerwalt AG, Fanning GR, Fife-Asbury MA, Brenner DJ (1976) DNA relatedness among species of *Enterobacter* and *Serratia*. Canadian Journal of Microbiology 22:121-137.

Stumpf MPH, McVean GAT (2003) Estimating recombination rates from population-genetic data. Nature Reviews Genetics 4:959-968.

Suerbaum S, Smith JM, Bapumia K, Morelli G, Smith NH, Kunstmann E, Dyrek I, Achtman M (1998) Free recombination within *Helicobacter pylori.* Proceedings of the National Academy of Sciences of the United States of America 95:12619-12624.

Urwin R, Maiden MCJ (2003) Multi-locus sequence typing: a tool for global epidemiology. Trends in Microbiology 11:479-487.

Whittam TS, Ochman H, Selander RK (1983) Multilocus genetic structure in natural populations of *Escherichia coli.* Proceedings of the National Academy of Sciences of the United States of America 80:1751-1755.

Chapter 7

General discussion

Life history of the two *Hafnia* species

This study has shown that the two genetic groups of *Hafnia* are indeed distinct species. The results of the phylogenetic analyses suggest that *Hafnia* group 2 evolved from group 1. *Hafnia* group 2 appears to exhibit a more clonal genetic structure than does *Hafnia* group 1, but there is no compelling evidence to suggest that the relative importance of mutation versus recombination differs substantially among the two groups of strains. Overall, the degree of clonality in both *Hafnia* groups is comparable to that of non-pathogenic *Escherichia coli*.

A comparison of the ecological characteristics of the two *Hafnia* species is presented in Table 1. Both *Hafnia* species are most commonly isolate from hosts living in temperate regions, and rarely recovered from hosts in desert or tropical regions (Gordon and FitzGibbon 1999; D.M. Gordon, personal communication). Both *Hafnia* species are most likely to be isolated from ectotherms and are rare in birds (D.M. Gordon, personal communication). The majority of *Hafnia* group 1 strains isolated from mammals were recovered from hosts with a body mass of < 200 g, whereas group 2 strains from mammals were found in hosts of any body size. Group 2 strains were quite common in fish, but group 1 strains were not. For group 2 strains host effects explained a substantial fraction of the observed phenotypic and genotypic variation, but accounted for a smaller fraction of the variation observed in group 1 strains. Overall the two *Hafnia* species are isolated at equal frequencies from the intestinal tracts of vertebrates and appear to be equally frequent in water samples (J. Wyer, personal communication).

Table 1. Summary of the differences between the two *Hafnia* species.

	Group 1	Group 2
Distribution in:		
Mammals	small mammals	any body size
Reptiles	any	any
Fish	rare	any (most abundant)
Host effect in:		
Biochemical traits	strong	strong
Sequence variation	moderate	strong
Growth rate variation	none	strong
Bacteriocin production	rare	frequent
Others	meat contaminant veterinary significance	human clinical significance

What are the ecological niches of these two *Hafnia* species? The life-history characteristics of members of the Enterobacteriaceae indicate that the relative importance of the host versus external environment varies among species. Species such as *E. coli* are predominantly host adapted, whist other species such as *Rahnella aquatilis* appear to exhibit a more free-living lifestyle (Leclerc *et al.* 2001). All of the evidence indicates that aspects of host biology shape the phenotypic and genotypic characteristics of group 2 strains to a greater extent than for group 1 strains. Bacteriocin production is more prevalent in group 2 strains and bacteriocin production has been shown to mediate intra-strains interactions in the mouse intestinal tract (Kirkup and Riley 2004). Do group 1 strains then exhibit a more free-living life style? Whilst group 1 strains are apparently as likely to occur in a host as group 2 strains perhaps they are less capable of persisting in a host than are group 2 strains. If this were true, then it would imply that group 1 strains are more capable of persisting in external environments than are group 2 strains (Davis and Gordon 2002). The observation that group 1 and 2 strains are detected at similar frequencies in water

samples is at odds with this hypothesis, but perhaps the survival of the two groups is different in soils and sediments.

Future directions

The relative fitness of the two *Hafnia* species in the gastrointestinal tract may be investigated using *in vivo* competition experiments with animal models such as rats, skinks and guppies (Kirkup and Riley 2004). Based on the observation of a stronger host effect in shaping the genetic and phenotypic variation in group 2, the expected outcome will be that group 2 strains will win more often than group 1 strains. By contrast, competition trials in soil, water and sediment will provide information on the relative fitness of the two *Hafnia* groups to conditions in the external environment of the host (Bogosian *et al.* 1996). It is expected that group 1 strains will exhibit better survival in these trials, but that the extent of the fitness advantage may vary depending on the nature of the environmental conditions (eg. soil versus water).

Other molecular genetic approaches may provide further information on the evolutionary origin and subsequent speciation of *Hafnia* group 2 from group 1. Genome subtraction may shed some light into the genes responsible for adaptation to the gut environment. Investigation of the genetic relationship of isolates from human-related sources (clinical samples, food, domesticated animals and humans faeces) to the commensal isolates may identify potential reservoirs of these strains. Such information may contribute to better measures for controlling disease and food contamination. Virulence factor identification and subsequent development of a standard screening protocol may aid in clinical typing of *Hafnia* strains. in particular, virulence factors responsible for septicemia (e.g. adhesin gene cluster, Babai *et al.* 1997) and peritonitis (Gunthard and Pennekamp 1996; e.g. hemoglobin protease, Otto *et al.* 2002; *degS*, Redford *et al.* 2003).

The methods used in this study to elucidate the ecological and genetic structure of *Hafnia* may be applied to other bacterial populations of which their structures are currently poorly understood (e.g. *Klebsiella pneumoniae*, *Enterobacter cloacae*). An understanding of the intricate relationship of bacterial populations can only be achieved by extensive investigation using a set of samples that best represents the distribution of the species as a whole. Not until more bacterial species are studied may we start to understand the true nature of bacterial population structures and how they evolve.

References

Babai R, Blum-Oehler G, Stern BE, Hacker J, Ron EZ (1997) Virulence patterns from septicemic *Escherichia coli* O78 strains. FEMS Microbiology Letters 149:99-105.

Bogosian G, Sammons LE, Morris PJL, Oneil JP, Heitkamp MA, Weber DB (1996) Death of the *Escherichia coli* K-12 strain W3110 in soil and water. Applied and Environmental Microbiology 62:4114-4120.

Davis SA, Gordon DM (2002) The influence of host dynamics on the clonal composition of *Escherichia coli* populations. Environmental Microbiology 4:306-313.

Gordon DM, FitzGibbon F (1999) The distribution of enteric bacteria from Australian mammals: host and geographical effects. Microbiology 145:2663-2671.

Gunthard H, Pennekamp A (1996) Clinical significance of extraintestinal *Hafnia alvei* isolates from 61 patients and review of the literature. Clinical Infectious Diseases 22:1040-1045.

Kirkup BC, Riley MA (2004) Antibiotic-mediated antagonism leads to a bacterial game of rock-paper-scissors in vivo. Nature 428:412-414.

Leclerc H, Mossel DAA, Edberg SC, Struijk CB (2001) Advances in the bacteriology of the coliform group: Their suitability as markers of microbial water safety. Annual Review of Microbiology 55:201-234.

Otto BR, van Dooren SJM, Dozois CM, Luirink J, Oudega B (2002) *Escherichia coli* hemoglobin protease autotransporter contributes to synergistic abscess formation and heme-dependent growth of *Bacteroides fragilis*. Infection and Immunity 70:5-10.

Redford P, Roesch PL, Welch RA (2003) *degS* is necessary for virulence and is among extraintestinal *Escherichia coli* genes induced in murine peritonitis. Infection and Immunity 71:3088-3096.